

MEMOIRE DE SYNTHESE
Présenté en vue de l'obtention de
l'Habilitation à Diriger des Recherches

Volume II

TRAVAUX, CV et LISTE DE PUBLICATIONS

**Description grammaticale,
comparaison typologique et approche théorique.
De l'otomi à d'autres langues otomangues**

Enrique L. PALANCAR

Chargée de Recherche 1^{re} classe au CNRS
SeDyL (CNRS, IRD, INALCO)

Coordinatrice :

Claudine CHAMOREAU (CNRS, SeDyL)

Jury :

Judith Aissen	(Université de Californie à Santa Cruz)
Marianne Mithun	(Université de Californie à Santa Barbara)
Martine Vanhove	(CNRS, LLACAN)
Olivier Bonami	(Université Paris Diderot-Paris 7)
Roberto Zavala	(CIESAS-Sureste)

TABLE DE MATIÈRES

Table de Matières	2	
Avant-propos	5	
CURRICULUM VITAE	6	
LISTE DE PUBLICATIONS	15	
REPRODUCTION DES TRAVAUX	31	
1. Travaux la syntaxe otomie	32	
TRAVAIL 1.	Palancar, Enrique L. 2016. On the tightrope between infinitives and action nouns : The case of Otomi nominalizations. Claudine Chamoreau et Zarina Estrada (eds.), <i>Finiteness and nominalization</i> , pp. 141 – 170. Amsterdam : John Benjamins.	33
TRAVAIL 2.	Palancar, Enrique L. 2013. The evolution of number in Otomi : The many faces of the dual. <i>Studies in Language</i> 37(1), 94 – 143.	
TRAVAIL 3.	Palancar, Enrique L. 2012. Subordination and juxtaposition. Volker Gast et Holger Diesel (eds.), <i>Clause Combining in Cross-linguistic perspective</i> , pp. 37 – 76. Berlin : Mouton de Gruyter.	
TRAVAIL 4.	Palancar, Enrique L. 2010. Semántica depictiva sin predicación secundaria. Judith Aissen et Roberto Zavala Maldonado (eds.), <i>La Predicación Secundaria en Lenguas de Mesoamérica</i> , pp. 251 – 279. Mexico : CIESAS.	
TRAVAIL 5.	Palancar, Enrique L. 2008. Juxtaposed adjunct clauses in Otomi : Expressing both depictive and adverbial semantics. <i>International Journal of American Linguistics</i> 74(3), 365 – 92.	
TRAVAIL 6.	Palancar, Enrique L. 2004. Verbal morphology and prosody in Otomi. <i>International Journal of American Linguistics</i> 70(3) : 251 – 278.	
2. Travaux sur le lexique otomi		
TRAVAIL 7.	Peralta, Valentin et Enrique L. Palancar. 2015. Actividades rutinarias en náhuatl y otomí. <i>Amerindia</i> 37:2, 259 – 284.	
TRAVAIL 8.	Palancar, Enrique L. 2008. The emergence of active/stative alignment in Otomi. Mark Donohue et Søren Wichmann (eds.), <i>The Typology of Semantic Alignment</i> , pp. 357 – 79. Oxford: Oxford University Press.	
TRAVAIL 9.	Palancar, Enrique L. 2007. Cutting and breaking in Otomi. <i>Cognitive Linguistics</i> 18:2, 307 – 17.	

- TRAVAIL 10. Palancar, Enrique L. 2006. Property concepts in Otomi: A language with no adjectives. *International Journal of American Linguistics* 72:3, 325 – 66.
- TRAVAIL 11. Palancar, Enrique L. 2006. Middle voice and intransitivity in Otomi. *Linguistics* 43:1: 613 – 43.
- TRAVAIL 12. Palancar, Enrique L. 2004. Middle voice in Otomi. *International Journal of American Linguistics* 70:1, 52 – 85.

3. Travaux sur aspects de la phonologie otomie

- TRAVAIL 13. Palancar, Enrique L. 2013. Aspiration in Northern Otomi and the relic of an old fortis/lenis contrast. Jean Léo Léonard et Samia Naïm (eds.), *Base Articulatoire Arrière. Backing and Backness*. Munich : LINCOM Europa.

4. Travaux de typologie

- TRAVAIL 14. Palancar, Enrique L. A paraître. Oto-Pamean. Soeren Wichmann (ed.), *The Languages and Linguistics of Middle and Central America: A Comprehensive Guide*. Berlin: Mouton de Gruyter.
- TRAVAIL 15. Palancar, Enrique L. 2016. A typology of tone and inflection: A view from the Oto-Manguean languages of Mexico. Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection: New data and new perspectives*, pp. 109 – 140. Berlin: De Gruyter Mouton.
- TRAVAIL 16. Palancar, Enrique L. et Roberto Zavala Maldonado. 2013. Clases léxicas, posesión y estructuras de cláusula compleja en Mesoamérica: Una introducción. Enrique L. Palancar et Zavala Maldonado (eds.). *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica*, pp. 17 – 48. Mexico: CIESAS.
- TRAVAIL 17. Palancar, Enrique L. 2012. A typology of split conjunction. *Linguistic Typology* 16: 265 – 320.
- TRAVAIL 18. Palancar, Enrique L. 2008. Varieties of ergatives. Andrey Malchukov et Andrew Spencer (eds.), *The Handbook of Case*, Chapter 37. Oxford: Oxford University Press.

5. Travaux sur la morphologie de langues otomangues

- TRAVAIL 19. Palancar, Enrique L. et Heriberto Avelino. A paraître. Inflectional complexity and verb classes in Chichimec. Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguean languages*. (Morphological Investigations Series 1). Berlin: Language Science Press

- TRAVAIL 20. Palancar, Enrique L. et Leonardo Carranza. A paraître. Inflectional class interaction in Matlatzinca. Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguean languages*. (Morphological Investigations Series 1). Berlin: Language Science Press
- TRAVAIL 21. Timothy Feist et Enrique L. Palancar. 2016. Tracing the emergence of inflectional tone in Cuicatec. Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection: New data and new perspectives*, pp. 267 – 294. Berlin: De Gruyter Mouton
- TRAVAIL 22. Palancar, Enrique L., Jonathan Amith et Rey Castillo García. 2016. Verbal inflection in Yoloxóchitl Mixtec. Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection: New data under new perspectives*, pp. 295 – 336. Berlin: De Gruyter Mouton
- TRAVAIL 23. Palancar, Enrique L. 2015. A mixed system of agreement in the suffix classes of Lealao Chinantec. *Morphology* 25, 29 – 62.
- TRAVAIL 24. Palancar, Enrique L. et Timothy Feist. 2015. Agreeing with subjects in number: The rare split of Amuzgo inflection *Linguistic Typology* 19:3, 337 – 383.
- TRAVAIL 25. Baerman, Matthew et Enrique L. Palancar. 2014. The organization of Chinantec tone paradigms. *Carnets de Grammaire* 22: 46 – 59.
- TRAVAIL 26. Palancar, Enrique L. 2014. Revisiting the complexity of the Chinantecan verb conjugation classes. Jean Léo Léonard et Alain Kihm (eds.), *Patterns in Mesoamerican morphology*, pp. 77 – 102. Paris : Michel Houdiard.
- TRAVAIL 27. Palancar, Enrique L. 2012. The conjugations classes of Tilapa Otomi: An approach from canonical typology. *Linguistics* 50 – 4: 783 – 832.
- TRAVAIL 28. Palancar, Enrique L. 2011. The conjugations of Colonial Otomi. *Transactions of the Philosophical Society* 109: 1 – 19.

6. Travaux sur l'espagnol

- TRAVAIL 29. 2011 Palancar, Enrique L. 2011. Oblique reflexives in Spanish: A constructional account of *sí* and *sí mismo*. *Probus: International Journal of Latin and Romance Linguistics* 23: 57 – 101.
- TRAVAIL 30. 2008 Palancar, Enrique L. et Luisa J. Alarcón. 2008. Predicación secundaria depictiva en español. *Revista de la Sociedad Española de Lingüística* 37, 337 – 370.

AVANT-PROPOS

Ce deuxième volume contient un CV détaillé, une liste complète de publications et présentations à colloques, et une sélection de 30 études conçue comme étant un appui au document de synthèse du premier volume. Les différents travaux, articles et chapitres d'ouvrages, publiés entre 2004 et 2016, sont classés thématiquement et je considère comme étant les plus signifiants pour montrer ma contribution scientifique au domaine de la linguistique ma thèse de Doctorat.

LISTE DES PUBLICATIONS

1. Ouvrages scientifiques

1.1. Ouvrages

- 2009 *Gramática y Textos del Hñöñhö : Otomí de San Ildefonso Tultepec, Queretaro*
Volumen I : Gramática, Mexico : Plaza y Valdés, ISBN 978 – 607 – 402 – 146
– 2 (xxiii + 597pp.)
- 2009 *Gramática y Textos del Hñöñhö : Otomí de San Ildefonso Tultepec, Queretaro*
Volumen II : Textos, Mexico City : Plaza y Valdés, ISBN 978 – 607 – 402 – 147
– 9 (v + 146pp. + CD)

Prix :

—Prix 2010 Wigberto Jiménez pour la meilleure recherche en linguistique par l'Institut national pour l'histoire et l'anthropologie du Mexique

Comptes – rendus :

—Levy, Paulette. 2012. « *Gramática Y Textos Del Hñöñhö, Otomí De San Ildefonso Tultepec, Querétaro* by Enrique L. Palancar ». *International Journal of American Linguistics* 78 :2, 281 – 284

- 2002 *The Origin of Agent Markers*, [Studia Typologica 5], Berlin : Akademie Verlag, ISBN 3 – 05 – 003767 – 9 (310 pp.)

Comptes – rendus :

—Holton, Gary. 2007. « The Origin of Agent Markers ». *Language* 83 :1, 226 – 227

—Maldonado, Ricardo. 2005. « Enrique L. Palancar, The origin of agent markers ». *Studies in Language* 29 :2, 490 – 495

—Luraghi, Silvia. 2004. « Review of The Origin of Agent Markers, by Enrique L. Palancar ». *Language Typology* 8 :1, 140 – 143

1.2. Directions d'ouvrages

- A paraître *Inflectional class complexity in the Oto-Manguean languages*
(2017) (Morphological Investigations Series 1). Berlin : Language Science Press
(Avec Timothy Feist et Matthew Baerman)

- 2016 *Tone and inflection : New data and new perspectives*
(Trends in Linguistics Series 296). Berlin : De Gruyter Mouton
(Avec Jean Léo Léonard)

- 2013 *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica*
Mexico : CIESAS
(Avec Roberto Zavala)

Comptes – rendus :

—Chamoreau, Claudine et Francisco Arellanes. 2016. « Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica. Edited by Enrique Palancar and Roberto Zavala Maldonado. México: CIESAS, 2013. Pp. 436. \$15.00; 220 pesos (paper) ». *International Journal of American Linguistics* 82:2, 250 – 253.

—Vázquez Rojas Maldonado, Violeta. 2015. « Enrique L. Palancar & Roberto Zavala (eds.). Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica. México: Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS), 2013. 440 págs. ». *Estudios de Lingüística Aplicada* 33:62, 245–261

2. Articles dans des revues avec comité de lecture :

2.1. REPERTORIEES dans les bases de données internationales

- 2015 « Agreeing with subjects in number : The rare split of Amuzgo inflection » *Linguistic Typology* 19 :3, 337 – 383
(Avec Timothy Feist)
- 2015 « A mixed system of agreement in the suffix classes of Lealao Chinantec ». *Morphology* 25, 29 – 62
- 2015 « Actividades rutinarias en náhuatl y otomí ». *Amerindia* 37 :2, 259 – 284
(Avec Valentín Peralta Ramírez)
- 2015 « The Contact Diffusion of Linguistic Practices : Reference Frames in Mesoamerica » *Language dynamics and change* 5 :2, 169 – 201
(Avec Jürgen Bohnemeyer; Katharine T. Donelson; Randi E. Moore; Elena Benedicto; Alyson Eggleston; Carolyn K. O’Meara; Gabriela Pérez; Alejandra Capistrán; Néstor H. Green; Selene Hernández Gómez; Samuel Herrera; Gilles Polian et Rodrigo Romero)
- 2013 « The evolution of number in Otomi : The many faces of the dual ». *Studies in Language* 37 :1, 94 – 143
- 2012 « A typology of split conjunction ». *Linguistic Typology* 16 : 265 – 320
- 2012 « The conjugations classes of Tilapa Otomi : An approach from canonical typology ». *Linguistics* 50 – 4 : 783 – 832
- 2011 « The Spanish loanword *lado* in Otomi ». *Language Sciences* 33 : 961 – 980
(Avec Néstor H. Green et Selene Hernández)
- 2011 « The conjugations of Colonial Otomi ». *Transactions of the Philosophical Society* 109 : 1 – 19
- 2011 « Revisiting the conjugation classes of Eastern Highlands Otomi ». *Language Typology and Universals* 64 – 3 : 213 – 236
- 2011 « Oblique reflexives in Spanish : A constructional account of *sí* and *sí mismo* ». *Probus : International Journal of Latin and Romance Linguistics* 23 : 57 – 101

- 2008 « Juxtaposed adjunct clauses in Otomi : Expressing both depictive and adverbial semantics ». *International Journal of American Linguistics* 74 :3, 365 – 92
- 2007 « Cutting and Breaking in Otomi ». *Cognitive Linguistics* 18 :2, 307 – 17
- 2006 « Property Concepts in Otomi : A language with no adjectives ». *International Journal of American Linguistics* 72 :3, 325 – 66
- 2006 « Middle Voice and Intransitivity in Otomi ». *Linguistics* 43 :1 : 613 – 43
- 2004 « Verbal morphology and prosody in Otomi ». *International Journal of American Linguistics* 70 :3, 251 – 78
- 2004 « Middle Voice in Otomi ». *International Journal of American Linguistics* 70 :1, 52 – 85
- 2001 « Emergent markers : The relation of instruments, causes, and agents in grammatical constructions ». *Language Typology and Universals* 54/4 : 365 – 384
- 1999 « Instrumental Prefixes in Amerindian languages : An overview to their meanings, origin, and functions ». *Language Typology and Universals* 52/2 : 151 – 166
- 1999 « What do we give in Spanish when we hit? A constructionist account of hitting expressions ». *Cognitive Linguistics* 10/1 : 56 – 91

2.2. NON-Répertoriées dans les bases de données internationales

- 2014 « The organization of Chinantec tone paradigms ». *Carnets de Grammaire* 22 : 46 – 59
(Avec Matthew Baerman)
- 2009 « Nor Kamalio : El Camaleón ». *Tlalocan : Revista de fuentes para el conocimiento de las culturas indígenas de México* 16 : 51 – 84. Mexico City : UNAM
- 2008 « Predicación secundaria depictiva en español ». *Revista de la Sociedad Española de Lingüística* 37, 337 – 370
(Avec Luisa J. Alarcón)
- 2008 « Posesión externa en otomí ». *Estudios de Cultura Otopame* 6, 255 – 271
- 2007 « Split coordination in Otomi and the expression of comitativity ». *Proceedings to the 33rd Meeting of the Berkeley Linguistics Society*. Berkeley, CA : Berkeley Linguistics Society
- 2004 « Datividad en Otomí ». *Estudios de Cultura Otopame* 4, 171 – 96
- 2001 « Agent encoding as an areal phenomenon? : Exploring the possibilities of case syncretisms ». *Proceedings of the First Seoul International Conference on Discourse and Cognitive Linguistics*, pp. 452 – 465. Seoul : Hankook Publishing Company

- 1997 « A private history of the preposition by in English : From instrumental to passive agent ». *Cuadernos de Filología Inglesa* 6/2 : 123 – 46. Murcia : Université de Murcia
- 1996 « Effectors : More than mere instruments ». *Revista de Filología de la Universidad Complutense* 4 : 67 – 90

3. Chapitres dans des ouvrages

- A paraître (2017) « Inflectional class complexity in the Oto-Manguéan languages : An introduction ». Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguéan languages*. (Morphological Investigations Series 1). Berlin : Language Science Press (Avec Timothy Feist et Matthew Baerman)
- A paraître (2017) « Inflectional complexity and verb classes in Chichimec ». Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguéan languages*. (Morphological Investigations Series 1). Berlin : Language Science Press (Avec Heriberto Avelino)
- A paraître (2017) « Inflectional class interaction in Matlatzinca ». Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguéan languages*. (Morphological Investigations Series 1). Berlin : Language Science Press (Avec Leonardo Carranza)
- A paraître (2017) « Oto – Pamean ». Soeren Wichmann (ed.), *The Languages and Linguistics of Middle and Central America : A Comprehensive Guide*. Berlin : Mouton de Gruyter
- 2016 « On the tightrope between infinitives and action nouns : The case of Otomi nominalizations ». Claudine Chamoreau et Zarina Estrada (eds.), *Finiteness and nominalization*, pp. 141 – 170. Amsterdam : John Benjamins
- 2016 « Tone and inflection : New data and new perspectives ». Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection : New data and new perspectives*, pp. 1 – 10. Berlin : De Gruyter Mouton (Avec Jean Léo Léonard)
- 2016 « A typology of tone and inflection : A view from the Oto-Manguéan languages of Mexico ». Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection : New data and new perspectives*, pp. 109 – 140. Berlin : De Gruyter Mouton
- 2016 « Tracing the emergence of inflectional tone in Cuicatec ». Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection : New data and new perspectives*, pp. 267 – 294. Berlin : De Gruyter Mouton (Avec Timothy Feist)

- 2016 « Verbal inflection in YoloXóchitl Mixtec ». Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection : New data under new perspectives*, pp. 295 – 336. Berlin : De Gruyter Mouton
(Avec Jonathan Amith et Rey Castillo García)
- 2014 « Revisiting the complexity of the Chinantecan verb conjugation classes ». Jean – Léo Léonard et Alain Kihm (eds.), *Patterns in Mesoamerican morphology*, pp. 77 – 102. Paris : Michel Houdiard
- 2013 « Aspiration in Northern Otomi and the relic of an old fortis/lenis contrast ». Jean – Léo Léonard et Samia Naïm (eds.), *Base Articulatoire Arrière. Backing and Backness*. Munich : LINCOM Europa
- 2013 « Clases léxicas, posesión y estructuras de cláusula compleja en Mesoamérica : Una introducción ». Enrique L. Palancar et Zavala Maldonado (eds.). *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica* pp. 17 – 48. Mexico : CIESAS
(Avec Roberto Zavala)
- 2013 « Nominalización y complementación en otomí ». Enrique L. Palancar et Zavala Maldonado (eds.). *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica*, pp. 381 – 412. Mexico : CIESAS
- 2012 « Clausal juxtaposition and subordination : New light from Otomi. ». Volker Gast et Holger Diesel (eds.), *Clause Combining in Cross – linguistic perspective*, pp. 37 – 76. Berlin : Mouton de Gruyter
- 2011 « Encadenamiento clausal en otomí : propósito como estrategia de complementación ». Ana Lidia Mungía Duarte (ed.), *Fonología, morfología y tipología semántico – sintáctica* [Colección Estudios Lingüísticos 1]. Hermosillo, Sonora, Mexique : Editorial UniSon
- 2011 « Semántica depictiva vs. semántica adverbial en textos narrativos de la escuela primaria ». María del Carmen Morúa et Gerardo López Cruz (eds.), *Estudios morfológicos, sintácticos, semánticos y de lingüística aplicada* [Colección Estudios Lingüísticos 3], pp. 169 – 191. Hermosillo, Sonora, Mexique : Editorial UniSon
(Avec Luisa J. Alarcón)
- 2010 « Semántica depictiva sin predicación secundaria ». Judith Aissen et Roberto Zavala Maldonado (eds.), *La Predicación Secundaria en Lenguas de Mesoamérica*, pp. 251 – 279. Mexico City : CIESAS
- 2008 « Varieties of ergatives ». Andrey Malchukov et Andrew Spencer (eds.), *The Handbook of Case*, Chapter 37. Oxford : Oxford University Press
- 2008 « The emergence of active/stative alignment in Otomi ». Mark Donohue et Søren Wichmann (eds.), *The Typology of Semantic Alignment*, pp. 357 – 79. Oxford : Oxford University Press

- 2008 « Predicación secundaria depictiva en el discurso de los niños mexicanos en la escuela primaria ». Mercedes García Cuesta (ed.), *Estudios de Desarrollo y Lenguaje y Educación*, pp. 65 – 74. Oviedo : Université de Oviedo (Avec Luisa J. Alarcón)
- 2005 « Los verbos saber y conocer en español : aspecto léxico y morfodinámica ». Margaret Lubbers et Ricardo Maldonado (eds.), *Dimensiones del Aspecto en Español*, pp. 17 – 54. Mexico : UNAM
- 2004 « Transitividad en el otomí de San Ildefonso : una incursión tipológica en lenguas otomangués ». Isabel Barreras et Mirna Castro (eds.), *Memoria del VII Encuentro Internacional de Lingüística en el Noroeste 7*, Volume I, pp. 173 – 196. Sonora : UniSon
- 2003 « La polisemia dei verbi *dar*, *meter* e *pegar* in spagnolo ». Livio Gaeta et Silvia Luraghi (eds.), *Introduzione alla Linguistica Cognitiva*, pp. 197 – 211. Roma : Carocci Editore
- 2002 « Demostrativos en el otomí de San Ildefonso ». Zarina Estrada et Rosa María Ortiz (eds.), *Memorias del VI Encuentro Internacional de Lingüística en el Noroeste 6*, pp. 361 – 375. Sonora : UniSon
- 2001 « Del verbo conocer en español : Un estudio morfodinámico del aspecto léxico ». Ricardo Maldonado (ed.), *Estudios Cognoscitivos del Español*, pp. 339 – 54. Madrid/Mexico City : Revista de la Asociación Española de Lingüística Aplicada

4. Autres productions :

4.1. Articles de vulgarisation

- 2014 « The Oto-Manguéan languages of Mexico ». *OGMIOS Newsletter 55* : 7 – 8, Foundation for Endangered Languages

4.2. Bases de données

- 2015 « The Oto-Manguéan Inflectional Class Database. Université de Surrey ». <<http://dx.doi.org/10.15126/SMG.28/1>> (Avec Timothy Feist)

4.3. Comptes -rendus d'ouvrages

- 2017 « Natalie Operstein and Aaron Huey Sonnenschein (eds.) 2015. *Valence changes in Zapotec : Synchrony, diachrony, typology* [Typological Studies in Language 110]. Amsterdam : John Benjamins ». *Journal of Historical Linguistics 7* : 124 – 127
- 2005 « Del Cora al Maya Yucateco : *Estudios Lingüísticos sobre Algunas. Lenguas Indígenas Mexicanas*. Edited by Paulette Levy, Mexico, DF : Universidad Nacional Autónoma de México ». *International Journal of American Linguistics 71* : 118 – 120

4.4. Publications pour l'enseignement

- 2009 « Materiales didácticos para asignaturas de Tipología Lingüística ». Eugenio R. Luján, *Seminario Aplicaciones de las TIC en la docencia en Lingüística Histórica y Comparada*. Universidad Complutense de Madrid.
<eprints.ucm.es/8443/1/TICLHistyComp.pdf>

5. Présentations :

5.1. Conférences données à l'invitation du Comité d'organisation dans un congrès national ou international

- 2016 Invité au Colloque *AnaMorphoSys : Analyzing Morphological Systems*
Université de Lyon
Lyon, 20 – 22 Juin
Conférence : Some comments to “Understanding multiple exponence : How can diachronic linguistics contribute to understanding morphology?” by Alice C. Harris
- 2016 Invité au Workshop Syntaxe de la relativisation III
CIESAS – Sudeste, San Cristóbal de las Casas, Mexique, 23 – 27 Mai
Conférence : De relativas libres y otras en el otomí de Tilapa
- 2016 Invité à la Journée d'études EFL : *Journée sur des relatives*
Paris Diderot
Paris, 12 Février
Conférence : Stratégies de relativisation en otomi de Tilapa
- 2015 Invité au Workshop Syntaxe de la relativisation I
CIESAS – Sudeste, San Cristóbal de las Casas, Mexique, 31 Aout – 11 Septembre
Conférence : Estrategias de relativización en el otomí de Tilapa
- 2015 Invité au Workshop Syntaxe de la relativisation I
CIESAS – Sudeste, San Cristóbal de las Casas, Mexique, 31 Aout – 11 Septembre
Conférence : De construcciones relativas sin núcleo en el otomí de Tilapa y su relación con estructuras de foco
- 2015 Invité à la Journée d'études : *Typologie aréale et stratégies de relativisation*
INALCO
Paris, 10 – 11 Décembre
Conférence : Focus and relative clauses in Otomi
- 2015 Invité à la Journée d'études EFL : *Approches formelles, quantitatives et descriptives des systèmes morphologiques*
University of Chicago Center in Paris, Université Paris Diderot
Paris, 13 Octobre
Conférence : The rare plural subject split of Amuzgo verbs
- 2014 Invité au Colloque *State of the Art of Mesoamerican Linguistics*
Max Plank Institute for Evolutionary Anthropology, Leipzig, 6 – 7 Décembre
Conférence : Oto – Pamean

- 2014 Invité au *Séminaire de Typologie linguistique "TypUlm"*
 École Normale Supérieure de Paris, 6 Juin
 Conférence : When two agreement systems compete : The suffix classes of Lealao Chinantec
- 2014 Invité au Journée d'études EFL : *Documentation des langues en danger : approches qualitatives, expérimentales et participatives*
 Institut National de Langues et Cultures Orientales, 6 Février
 Conférence : The on – going documentation of Tilapa Otomi : A critically endangered language of Mexico"
- 2013 Invité au Colloque *Congreso de idiomas indígenas de Latinoamérica VI*
 Université de Texas at Austin, USA, 24 – 26 Octobre
 Conférence : Evolución de la complejidad morfológica : La diacronía de las conjugaciones verbales del otomí
- 2013 Invité au Colloque *Tons et Paradigmes Flexionnels : Modélisation et Parcimonie*
 Maison de la Recherche, Université de Paris 3, 28 Juin
 Conférence : Disentangling the tonal inflection of Cuicatec
 (Avec Timothy Feist)
- 2011 Invité au Département de linguistique de l'université Karl – Franz
 Graz, 9 Jun
 Conférence : A typology of split conjunction
- 2011 Invité au Colloque *III Morphology Meeting of the Southeast*
 Université du Surrey, Guildford, RU, 8 Avril
 Conférence : Shape conditions in Otomi
- 2011 Invité au International workshop *Finitude et Nominalisation*
 INALCO, Paris, 12 – 14 Octobre
 Conférence : Where nominalization ends and derivation begins
- 2011 Invité au International workshop *Mesoamerican Inflectional Morphology*
 Paris 3, 1 Février
 Conférence : Reducing morphological complexity
- 2011 Invité au *Séminaire de présentations du SeDyL*
 Paris, 2 Février
 Conférence : Complementation and nominalization : infinitives in Spanish and Otomi
- 2010 Invité au *Surrey Linguistics Club*
 Université de Surrey, 23 Septembre
 Conférence : Verb inflectional classes in Tlapeuzco Chinantec
- 2010 Invité au *VIII Coloquio de Lingüística de la ENAH*
 Mexico, 22 Avril
 Conférence : Formas no – finitas y pseudo – finitas en la familia otomí

- 2009 Invité à l'Institut de recherches philologiques
Université National Autonome du Mexique, Mexico, 27 Novembre
Conférence : Aspectos críticos del libro : Studies in Role and Reference Grammar,
editado por Valeria Belloro y Lilián Guerrero
- 2009 Invité au *Surrey Linguistics Club*
Université de Surrey, 3 Juin
Conférence : TAM morphology in Colonial Otomi
- 2008 Invité au Département d'études grecques et indo – européens
Université Complutense de Madrid, 12 Décembre
Conférence : Métodos de trabajo de campo y documentación lingüística
- 2008 Invité au Département d'études grecques et indo – européennes
Université Complutense de Madrid, 11 Décembre
Conférence : Aplicación de las TIC a la formación de profesores de Lingüística
Histórica y Comparada
- 2007 Invité au *I Coloquio de Etnografía Maya*
École National d'Anthropologie et Histoire, Mexico, 1 – 3, Octobre
Conférence : Morfología verbal del otomí de Tilapa
- 2007 Invité au *II Taller sobre la Complementación*
Association Oxlajuuj Keej Maya' Ajtz' iib'
Antigua, Guatemala, 30 Juin – 11 Août
Conférence : Aspectos de integración clausal en el otomí de San Ildefonso Tultepec
- 2007 Invité au Journée d'Études : *Master en linguistique indoaméricaine*
CIESAS, Mexico, 4 Avril
Conférence : Comitativos en otomí
- 2006 Invité au Département d'études grecques et indo – européennes
Université Complutense de Madrid, 31 Mai
Conférence : Sobre las estructuras de coordinación
- 2006 Invité au Département d'études grecques et indo – européens
Université Complutense de Madrid, 31 Mai
Conférence : Algunas características tipológicas del otomí
- 2006 Invité au Workshop *I Taller de Complementación*
CIESAS – San Cristóbal de las Casas, 31 Juillet – 11 Août
Conférence : Complementación en el otomí de San Ildefonso Tultepec
- 2005 Invité à l'École d'été *Diplomado en lingüística descriptiva y empatía cultural*
Université de Madero, Puebla, Mexique, 14 Juillet
Conférence : Construcciones posesivas externas en otomí
- 2004 Invité au Colloque *Conférence of Otomanguan and Oaxacan Languages*
Université de California à Berkeley, 19 – 21 Mars
Conférence : Minimizing morphology, maximazing function

- 2003 Invité à l'Institut de recherches philologiques
 Université National Autonome du Mexique
 Mexico, 20, Février
 Conférence : Aspectos críticos del libro : *Del cora al maya yucateco : Estudios lingüísticos recientes sobre algunas lenguas indígenas mexicanas*, editado por Paulette Levy
- 2003 Invité au *Language and Cognition Group*
 Max Plank Institute for Psycholinguistics, 20 Juin
 Conférence : Otomi middle voice and other verb classes
- 2002 Invité au Département de linguistique
 CIESAS, Mexico, 23, Octobre
 Conférence : Aspectos críticos del libro : *La adquisición de la lengua materna : Español, lenguas mayas y euskera*, editado por Cecilia Rojas y Lourdes de León
- 2001 Invité au Colloque *III Simposio de Psicolingüística*
 Université Métropolitane du Mexique, 25 – 26 Avril
 Conférence : De cómo codificamos eventos de movimiento : indagaciones en la relatividad (psico –)lingüística
- 2001 Invité au *Language and Cognition Group*
 Max Plank Institute for Psycholinguistics, 13 Juin
 Conférence : Mysterious ways in San Ildefonso Otomi, Language and Cognition Group
- 2000 Invité au Séminaire général de recherche
 Faculté de Lettres, Université de Quérétraro, 12 Octobre
 Conférence : Dinamismo y representación en la caracterización del rol semántico del instrumento
- 1999 Invité au *Winter Symposium : The Role of Language in Cognition and Culturalization*
 Semiotic Research Centre, Aarhus, 28 – 30, Janvier
 Conférence : Hitting events in Spanish, or the case for the emergence of a novel grammatical construction
- 1998 Invité au Département de linguistique
 Université de Bremen, 11 Novembre
 Conférence : Instrumental Prefixes in the languages of the Americas
- 1996 Invité au Lingvistklubben
 Université d'Oslo
 Oslo, 29 Mai
 Conférence : How differently do we HIT in Spanish, English and Norwegian? : A typology of hitting constructions

5.2 Communications avec actes dans un congrès international

- 2008 Participation au *X Encuentro Internacional de Lingüística en el Noroeste*
Université de Sonora, Mexique. 11 – 14 Novembre
Communication : Propósito como estrategia de complementación
- 2004 Participation au *VIII Encuentro Internacional de Lingüística en el Noroeste*
Université de Sonora, Mexique, 17 – 20, Novembre
Communication : La Posesión Externa con dativo
- 2002 Participation au *VII Encuentro Internacional de Lingüística del Noroeste*
Université de Sonora, Mexique, 13 – 15, Novembre
Communication : Transitividad en el otomí de San Ildefonso : una incursión tipológica en lenguas otomangués
- 2000 Participation au *VI Encuentro Internacional de Lingüística del Noroeste*
Université de Sonora, Mexique, 29 – 30, Novembre
Communication : Diciembre. Sistema demostrativo en el Otomí de Amealco
- 2001 Participation au *1st Seoul Conference in Discourse and Cognitive Linguistics*
Université de Seoul, Corée, 8 – 10, Juin
Communication : Agent encoding as an areal phenomenon?

5.3 Communications orales sans actes dans un congrès international/national

- 2015 Participation au *Colloque international de morphologie Décembrettes 9*
Université Toulouse Jean Jaurès, Toulouse, 3 – 4 Décembre
Communication : Cross – classification in the inflection of Amuzgo verbs
- 2014 Participation au *16th International Morphology Meeting*
Session : *Synchrony and diachrony of Inflectional Classes : Theoretical and Empirical Considerations*
Budapest, 29 Mai – 1 Juin
Communication : Reaching out beyond the word form : periphrasis and inflectional classes in Otomi
(Avec Matthew Baerman)
- 2014 Participation au *16th International Morphology Meeting*
Session : *Synchrony and diachrony of Inflectional Classes : Theoretical and Empirical Considerations*
Budapest, 29 Mai – 1 Juin
Communication : The singular/plural split and the marking of verb inflectional classes in San Pedro Amuzgos Amuzgo
(Avec Timothy Feist)
- 2014 Participation au *Coloquio sobre lenguas oto – mangués y vecinas 6*
Oaxaca, Mexique. 27 – 27 Avril
Communication : Perífrasis y clases flexivas en otomí
(Avec Matthew Baerman)

- 2014 Participation au *Coloquio sobre lenguas oto – mangués y vecinas 6*
Oaxaca, Mexique. 27 – 27 Avril
Communication : Patronés en la flexión tonal del cuicateco
(Avec Timothy Feist)
- 2013 Participation au *Linguistics Association of Great Britain Annual Meeting*
Session : *Comparative Morphology & Morphological Theory*
SOAS, Londres, 28 – 31 Août
Communication : Choosing what bit? Tone and Inflection in Cuicatec
(Avec Timothy Feist)
- 2013 Participation au *Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas*
Session : *Inflectional classes in the languages of the Americas*
Boston, 3 – 6 Janvier
Communication : Inflectional Classes in the languages of the Americas : An overview
(Avec Matthew Baerman, Greville G. Corbett, et Dunstan Brown)
- 2013 Participation au *Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas*
Session : *Inflectional classes in the languages of the Americas*
Boston, 3 – 6 Janvier
Communication : Revisiting the conditioning and distribution of the subject suffixes in Lealao Chinantec
- 2010 Participation au *2010 Meeting of the Society for the Study of the Indigenous Languages of the Americas*
Linguistics Society of America, Baltimore, 7 – 11 Janvier
Communication : The many interpretations of the loanword *lado* in the Grammar of Space of Otomi
(Avec Néstor H. Green et Selene Hernández)
- 2010 Participation au colloque *Language, Text & History : Linguistics and Philology in the 21st Century*
Murray Edwards College, Université de Cambridge, Cambridge, 19 – 20 Mars
Communication : Revisiting Colonial Otomi verb conjugations in Cárceres' 1580 grammar
- 2010 Participation au *Coloquio Thomas – Smith Stark : IV Congreso sobre Lenguas Otomangués*
Biblioteca Francisco de Burgoa, Oaxaca, Mexique, 16 – 18 Avril
Communication : Clases verbales en la familia otomí : El otomí de la Sierra
- 2009 Participation au *Congreso de la Asociación Mexicana de Lingüística*
Toluca, Mexique, 5 – 7 Oct
Communication : El uso de los artículos en los niños con trastorno específico del lenguaje
(Avec Alejandra Auza)

- 2009 Participation au *III International Conférence of the Languages of Latinamerica*, Université de Texas à Austin, 29 – 31 Oct
Communication : Reconstructing verbal conjugations in Colonial Otomi
- 2009 Participation au *Morphology of the World's Languages*, Leipzig, 11 – 13 Juin
Communication : The TAM morphology in Tilapa Otomi verb conjugations
- 2008 Participation au *13th International Morphology Meeting*
Université de Vienna, 3 – 6 Février
Communication : Variation in Tilapa Otomi Verb Conjugations
- 2008 Participation au *Syntax of the World's Languages III*
Free Université de Berlin, 25 – 28 Sept
Communication : Linking clauses by juxtaposition in Otomi
- 2007 Participation au *III International Conférence of the Languages of Latinamerica*
Université de Texas à Austin, 25 – 27 Octobre
Communication : Clausal complementation in Otomi
- 2007 Participation au *III International Conférence of the Languages of Latinamerica*
Université de Texas à Austin, 25 – 27 Octobre
Communication : Object and Dative marking in San Ildefonso Tultepec Otomi
(Avec Selene Hernández)
- 2007 Participation au *V Congreso Internacional de Adquisición del Lenguaje*
Université de Oviedo, 26 – 28 Septembre
Communication : Predicación Secundaria Depictiva en el Discurso Narrativo de los niños en edad escolar
(Avec Luisa J. Alarcón)
- 2007 Participation au *IX Congreso Nacional de la Asociación Mexicana de Lingüística Aplicada*
Université de Zacatecas, Mexique, 2 – 5 Octobre
Communication : Predicación Secundaria en el español medieval de Calilla e Dimna
(Avec M^a Luisa Álvarez)
- 2007 Participation au *33rd Meeting of the Berkeley Linguistics Society*
Université de California à Berkeley, 9 – 10 Février
Communication : Split coordination in Otomi and the expression of comitativity
- 2006 Participation au *I Coloquio Transdisciplinario de Literatura, Salvador Elizondo In Memoriam*
Université Autonome de Querétaro, 25 – 26 Mai
Communication : Deor : ejemplo atípico de la poesía elegíaca anglosajona
- 2006 Participation au *Syntax of the World's Languages II*
Université de Lancaster, 14 – 17 Septembre
Communication : Depictive semantics without secondary predication

- 2005 Participation au *The Typology of Stative – Active Languages*
Max Planck Institute for Evolutionary Anthropology, Leipzig, 20 – 22 Mai
Communication : Otomi Split Intransitivity and other related constructions
- 2005 Participation au *II International Conférence of the Languages of Latinamerica*
Université de Texas à Austin, 22 – 23 Octobre
Communication : Reciprocals in Otomi
- 2004 Participation au *Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas*
Boston, 7 – 11, Janvier
Communication : An initial approach to property concepts in Otomi
- 2004 Participation au *V Coloquio de Lingüística en la ENAH*
Mexico, 28 – 30, Avril
Communication : Morfología verbal y prosodia en otomí
- 2003 Participation au *I International Conférence of the Languages of Latinamerica*
Université de Texas à Austin, Texas, 23 – 25, Octobre
Communication : Middle voice and intransitivity in Otomi
- 2003 Participation au *6th Annual Workshop on American Indigenous Languages*
Université de California à Santa Barbara, 25 – 27, Avril
Communication : Middle voice in Otomi
- 2003 Participation au *Child Language Seminar*
Université de Newcastle upon Tyne, 9 – 11, Juillet
Communication : The role of Instruments in the acquisition of agent words
(Avec Alejandra Auza)
- 2002 Participation au *Reviewing Linguistic Thought : Perspectives into the 21st Century*, Université de Athènes, 21 – 24, Mai
Communication : Absolute and energetic intransitives
(Avec Ricardo Maldonado)
- 2002 Participation au *III Congreso de la Asociación Española de Lingüística Cognitiva*
Université de Valencia, 15 – 17, Mai
Communication : La voz media en un continuo de in/transitividad
- 2002 Participation au *5th Annual Workshop on American Indigenous Languages*
Université de California à Santa Barbara, 26 – 28, Avril
Communication : San Ildefonso Otomi verbs : verbal allomorphy sensitive to syntactic environment
- 2001 Participation au *VI Congreso Nacional de la Asociación Mexicana de Lingüística Aplicada*
Université de Yucatán, Mexique, 26 – 29, Octobre
Communication : El caso de los verbos compactos en el Otomí de San Ildefonso

- 2001 Participation au 7th *International Cognitive Linguistics Conference*
Université de California à Santa Barbara, 22 – 27, Juillet
Communication : Ergativity and dynamism : a conceptual introspection into the origin and evolution of ergative markers
- 2001 Participation au *Jornadas del curso Voz Sintáctica*
Université Autonome de Queretaro, 12 – 17, Novembre
Communication : The Dative in Spanish and in Otomí
- 2000 Participation au *II Congreso de la Asociación Española de Lingüística Cognitiva*, Université complutense de Madrid, 17 – 20, Mai
Communication : A new perspective on description : An experimental approach
(Avec Enrique Bernárdez, M. Guarddon, M. Mioduszewska et M. Romano)
- 2000 Participation au *II Congreso de la Asociación Española de Lingüística Cognitiva*
Université complutense de Madrid, 17 – 20, Mai
Communication : Origen y evolución de los agentes ergativos : Una perspectiva conceptual
- 1998 Participation au *The Verb in Cognitive Grammar : Morphological, Semantic and Syntactic Aspects*
Université d'Oslo, 13 – 15, Juin
Communication : Polysynthesis and Instrumental encoding across American Indian languages
- 1998 Participation au *II Congreso de la Asociación Española de Lingüística Cognitiva*
Université d'Alicante, 4 – 6, Mai
Communication : Prefijos instrumentales en lenguas indígenas de la América del Norte
- 1997 Participation au *XXVI Simposio de la Sociedad Española de Lingüística*
Conseil supérieur pour la recherche scientifique, 16 – 20, Décembre
Communication : Del sufijo *-azo* o de cómo nos pegamos en castellano
- 1997 Participation au *I Reunión del Grupo de Trabajo Español de Lingüística Cognitiva*. Université de Logroño, 21 – 22, Février
Communication : La construcción de golpeo en castellano
- 1997 Participation au 5th *International Cognitive Linguistic Conference*
Université libre d'Amsterdam, 14 – 19, Juillet
Communication : Understanding general patterns for grammatical coding : causal and agential perspectives
- 1996 Participation au *XXV Simposio de la Sociedad Española de Lingüística*
Université de Zaragoza, 18 – 20, Décembre
Communication : Instrumentos y Agentes

REPRODUCTION DES TRAVAUX

1. Travaux la syntaxe otomie

- TRAVAIL 1. Palancar, Enrique L. 2016. On the tightrope between infinitives and action nouns : The case of Otomi nominalizations. Claudine Chamoreau et Zarina Estrada (eds.), *Finiteness and nominalization*, pp. 141 – 170. Amsterdam : John Benjamins. 33
- TRAVAIL 2. Palancar, Enrique L. 2013. The evolution of number in Otomi : The many faces of the dual. *Studies in Language* 37(1), 94 – 143. 63
- TRAVAIL 3. Palancar, Enrique L. 2012. Clausal juxtaposition and subordination : New light from Otomi. Volker Gast et Holger Diesel (eds.), *Clause Combining in Cross-linguistic perspective*, pp. 37 – 76. Berlin : Mouton de Gruyter. 113
- TRAVAIL 4. Palancar, Enrique L. 2010. Semántica depictiva sin predicación secundaria. Judith Aissen et Roberto Zavala Maldonado (eds.), *La Predicación Secundaria en Lenguas de Mesoamérica*, pp. 251 – 279. Mexico : CIESAS. 135
- TRAVAIL 5. Palancar, Enrique L. 2008. Juxtaposed adjunct clauses in Otomi : Expressing both depictive and adverbial semantics. *International Journal of American Linguistics* 74(3), 365 – 92. 151
- TRAVAIL 6. Palancar, Enrique L. 2004. Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70(3) : 251 – 278. 179

On the tightrope between infinitives and action nouns

The case of Otomi nominalizations

Enrique L. Palancar
SeDyL (UMR-8202), CNRS
University of Surrey

In this paper, I study the distribution of two nominalization structures in two closely related Amerindian languages of Mexico (Eastern Otomi and Northern Otomi). The structures involve intransitive nominalizations depicting a customary activity performed by humans ('hunting', 'sowing', etc.) as they appear in a complementation frame. I address the question of whether these nominalization structures in Otomi are morphologically non-finite forms or are just nouns. I show how a cognate structure in two close languages can be found in the two opposing ends of a nominalization scale: in Eastern Otomi, the structure is inflectional and should be analyzed as an intransitive infinitive, while its equivalent in Northern Otomi should be seen as an action deverbal noun.

Keywords: action noun; deverbal noun; infinitive; complement; Otomi

1. Introduction: Establishing a common ground

In this paper, I study the distribution and properties of two nominalization structures in two closely related Otomi languages of Mexico: Eastern Otomi (from the linguistic variety of Huehuetla, in the state of Hidalgo, also known as Highlands Otomi) and Northern Otomi (especially from the variety of San Ildefonso Tultepec, in the state of Querétaro). The structures in question appear in a special complementation construction that involves a nominalized form of a verb as a non-finite predicate. Such a complementation construction is special in that it requires the nominalization to be intransitive and depict a customary activity performed typically by humans, such as 'hunting', 'sowing', or 'having a bath', so that it is embedded in a complementation frame such as *I'm going off hunting, you finished sowing, the child started having a bath*, and the like. An example of such a construction is given

in (1) from Mezquital Otomi, a variety of Northern Otomi that is the most widely spoken in the family.¹

- (1) *mānde* *bi=mfa'yo* *ra Xuwa*
 MEZ yesterday 3.CPL=herd.sheep SG John
nu *mí=ma* [*ra 'beni*] *rá* *nānā*
 when 3.IMPF=ss/go SG doing.laundry SG.3POSS mother
 'John herded sheep yesterday as his mother was leaving to do the laundry.'
 (Hernández Cruz et al. 2004)

I address the question of whether nominalization structures in Otomi such as *'beni* 'doing the laundry' in a frame like (1) are morphologically non-finite forms of verbs or are just nouns. In this respect, I propose that the equivalent structure to (1) in Eastern Otomi is of a verbal non-finite form, which is inflectional and could be analyzed as an intransitive infinitive, whereas the actual structure in Northern Otomi in (1) should perhaps be seen as instantiating an action deverbal noun. In order to argue for this in the rest of the paper, in this section I first make a brief excursion into the concept of nominalization that I use here to approach the Otomi data. This excursion is intended to sketch a differentiation between infinitives and action nouns, which although well known in the literature is very rarely spelled out.

The term "nominalization" is used in more than one sense. From a morphological perspective concerned with word-formation, it means in essence "turning something into a noun" (Comrie & Thompson 2007: 334), so that nouns such as CONSTRUCTION_N, CONSTRUCTING_N, or CONSTRUCT_N /'kɒnstɹəkt/ are both possible and existing nominalizations of the verbal lexeme CONSTRUCT_V 'build, construct'. In this sense, nominalization can apply to verbs in the same way as to any other word class, providing the result is a noun: REDNESS is also a nominalization of RED, LENGTH of LONG, and so on. From a syntactic perspective, nominalization is used to refer to "constructions that have properties of noun phrases, but are headed by an element that is to some extent verbal" (Muysken 1999: 248). The syntactic perspective is concerned with the study of the syntactic properties of nominalized forms of verbs in specific grammatical contexts; one of these properties is finiteness.

Finiteness as a concept has in turn been approached from different perspectives. From a syntactic perspective, finiteness is a property of the clause (Givón this volume, Nikolaeva 2007, etc.), while from a morphological perspective it is a morphological property of verbal forms. From the latter perspective, for example, non-finite forms have reduced inflectional properties; in this respect one of the most common non-finite

1. Orthography: *f* /p^h/; *j* /k^h/; ' /ɾ/; *ñ* /ɲ/; *tx* /tʃ/; *r* /ɾ/; *s* /ʃs/ or /ʃs^h/ (in Eastern Otomi); *x* /ʃ/; *y* /j/; the vowels are a /ɔ/ [v]; e /ɛ/; o /ɔ/; i /i/; the " indicates a nasal vowel.

forms of verbs in European languages are infinitives. This has consequences for lesser-known languages, as treating and glossing a given verbal form as an infinitive creates a number of expectations about the behavior of such a form. These expectations are naturally based on a consensual knowledge of how infinitives typically behave in the European languages.

Following this common knowledge, I take an infinitive to be an inflected form of a verb that has been recategorized as a noun to occur in non-finite syntactic contexts. For an infinitival form to be inflectional implies (i) that the form in question is part of the inflectional paradigm of the verbal lexeme (i.e., it is a form of the lexeme made up by productive rules, and is thus accessible for all speakers), and (ii) that it is required for the lexeme by the syntax of specific grammatical contexts, those that typically revolve around the syntax of clausal complementation and clausal adjunction – such as the infinitive in Spanish as illustrated in (2a–b), where it realizes the lexeme SALIR ‘go out’ just like the finite form in (2c), but in a different syntactic context.

- (2) a. *quier-o* [*sali-r*]
 want-1SG.PRS.IND go.out-INF
 ‘I want to go out.’
- b. *lo=vi* [*a-l* *sali-r*]
 3SG.M.OBJ=see.1SG.PST.IND at-DEF.M.SG go.out-INF
 ‘I saw him when I was going out.’
- c. *sal-i*
 go.out-1SG.PST.IND
 ‘I went out.’

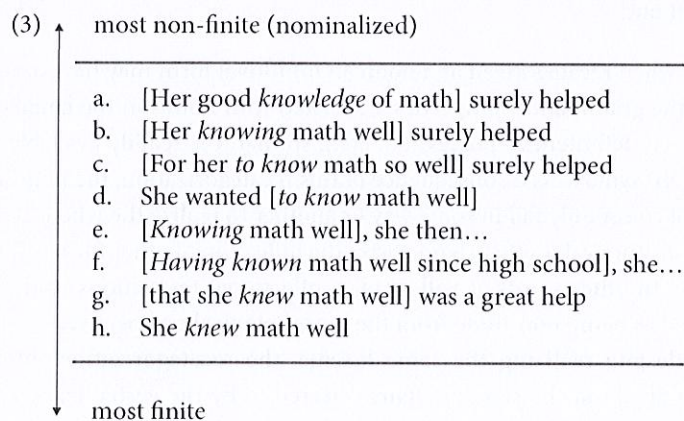
On the other hand, when recategorized as a noun an infinitival form may have access to (all or some of) the grammatical properties associated with nouns in the language (gender, case, number, definiteness, possession, etc.), so that it is readily available to take part in typical NP syntax. As a consequence of this recategorization, the nominal form of the verb will commonly fail in some way or another to realize the whole array of morphosyntactic feature values that characterize the inflection of other forms of the same verbal lexeme. In other words, it will display inflectional restrictions, and can thus be characterized as being non-finite from the morphological perspective.

Nevertheless, despite realizing the same lexeme, the recategorization into a noun allows us to talk about the state of affairs referred to by the verbal lexeme in a holistic way, for example *el reír de los niñ-o-s* [DEF.M.SG laugh-INF of DEF.M.PL child-M-PL] ‘the laughing of the children’. Once the verbal lexeme is used as a noun, its semantics may evolve independently by acquiring a new meaning that may in turn evolve into a new lexeme, as in CANTAR_v ‘sing’ → cantar_{INF} ‘to sing’ > CANTAR_N ‘lyric or epic song’ (e.g. *el cantar de los cantares* ‘the song of songs’); DEBER_v ‘must’ → deber_{INF} ‘must’ > DEBER_N ‘duty’ (e.g. *el deber del ciudadano* ‘the citizen’s duty’); or

SABER_V ‘know’ → *saber*_{INF} ‘to know’ > SABER_N ‘knowledge, learning’ (e.g. *una persona de gran saber* ‘a person of great learning’).

Independent nouns in Spanish from infinitival forms such as CANTAR_N, DEBER_N and SABER_N are few and hard to find. This is because the language has other more common word-formation ways to produce similar deverbal nouns. In this light, a deverbal noun, in contrast to an infinitive, semantically profiles some aspects of the state of affairs referred to by the verbal lexeme, and it does this as an independent lexical unit. It may refer to the agent (e.g. *cantante* ‘singer’), the instrument (e.g. *secador* ‘dryer’), the location (e.g. *consultorio* ‘medical office’), and so on, or to aspects of the action itself, profiling its outcome product or rendering the action as a whole. The so-called “action nouns” are the type of deverbal nouns that typically serve this purpose cross-linguistically; they commonly develop a wide range of meanings (e.g. *escritura* ‘writing, spelling, art of writing’, *enseñanza* ‘teaching, education’). Deverbal nouns are further subject to restrictions on the productivity of certain derivational rules in a given language, and are prone to have lexical gaps.

The morphological properties of the nominalization in question, whether inflectional (e.g. the infinitival type) or derivational (e.g. deverbal action nouns) have a bearing on the syntactic behavior of these elements. According to the finiteness scale in (3) from Givón (this volume), the more inflectional a certain morphological nominalization is, the more finite properties it will display (e.g. the *-ing* forms in English in (3b)), while if it is subject to derivational restrictions, it will be found on the non-finite end of the scale (e.g. the *-tion* forms or others in English found in (3a)).



In this paper, I study a case of nominalization in Otomi to be found in the grammatical space of Givón’s finiteness scale between (a) and (e). The continuum in (3) is designed to accommodate the range of possibilities existing in English. This range is wide because English has three distinct forms in competition (deverbal nouns such as *knowledge* and two other morphologically non-finite forms *knowing* and (*to*) *know*).

In the Otomi languages, verbs can only have one nominalized form to be used equivalently. This significantly reduces the topological space of possibilities available for the finiteness scale in Otomi. However, I show in the paper that for some Otomi languages like Eastern Otomi, the same form can be a deverbal noun or an infinitive depending on the grammatical context, whereas in Northern Otomi their equivalent forms are deverbal nouns in all contexts.

The paper is structured as follows. In the next section, I give a brief introduction to Otomi as a linguistic family and present some basic grammatical notions that are useful for understanding what follows, namely the properties of finite verbs and some basics about NP structure. In Section 3, I present the construction that involves a nominalized complement. This construction, found in all the Otomi languages, is the specific grammatical context from which Eastern Otomi has developed an infinitive. In Section 4, I elaborate on the differences of the two constructions of Eastern and Northern Otomi, and show that in Northern Otomi the nominalizations are deverbal nouns rather than syntactic nominalizations. The paper concludes in Section 5 with a summary of the argument.

2. Otomi

Otomi is a small, shallow language family of Mexico that belongs to the Oto-Pamean branch of Oto-Manguean. The number of languages in the family is still an open question. The National Institute for the Indigenous Languages of Mexico (INALI) acknowledges nine different linguistic varieties in the CLIN (2008).² INALI's linguistic varieties are treated as genuine languages for official – that is, potential educational, administrative, judicial, and information – purposes. However, the dialectal decisions behind the CLIN are not based on a rigorous dialectological survey and they place much emphasis on sociolinguistic criteria. From a purely linguistic point of view, based on the data of Lastra (2001), on the available descriptive materials, and on my own field notes, I argue that the Otomi family consists of six languages spoken in four main dialectal

2. A linguistic variety is defined in CLIN (2008:37): “La variante lingüística... se define como una forma de habla que: (a) presenta diferencias estructurales y léxicas en comparación con otras variantes de la misma agrupación lingüística; y (b) implica para sus usuarios una determinada identidad sociolingüística, que se diferencia de la identidad sociolingüística de los usuarios de otras variantes.” [“A variety of speech that (i) has structural and lexical differences in comparison with other varieties within the same linguistic group, and (ii) has a distinct sociolinguistic mark of identity for its users, different from the sociolinguistic identity borne by speakers of other varieties” (my translation).]

areas. From this position, I consider the existing varieties spoken nowadays in different villages to be dialects of one of these six languages:

- Northern Otomi (comprising the varieties of Mezquital, San Ildefonso Chantepec, San Ildefonso Tultepec, Toliman, Santiago Mexquititlán, Guanajuato, Acambay, and also most probably the variety of Ixhuatlán de Madero in the east)
- Eastern Otomi (comprising the varieties of Huehuetla, Tenango, Tulancingo, Texcatepec, Santa Ana Hueytlalpan, and San Pablito)
- Western Otomi (comprising the varieties of Temoaya, San Felipe Santiago, San Andrés Cuexcontitlán, and most probably the Otomi spoken in Huixquilucan)
- three other Southern languages: Acapulco-Atlapulco Otomi and Tilapa Otomi to the southwest and Ixtenco Otomi to the southeast.

I study here the distribution of a similar nominalization construction in both Northern and Eastern Otomi. The two languages serve as good examples of the two opposing ends of a continuum of inflectional conservatism in the family. In this respect, Northern Otomi and Western Otomi display a large degree of simplification of the morphological complexities in the inflectional system of Colonial Otomi, as described in the sixteenth-century grammar written by Pedro de Cárceres (1580[1907]).³ For example, of the three conjugation classes that we can reconstruct for the old system (Palancar 2011), only one survives in these two languages. In contrast, both Eastern Otomi and all Southern languages have managed to preserve a great deal of the old inflectional contrasts. For example, the three conjugations have survived in all of them, although with some degree of leveling. Because of this, we might treat Eastern Otomi as inflectionally conservative. Nevertheless, for the particular phenomenon I study here, the rise of the infinitive in this language is an innovative feature, as discussed in Section 4.1.

In Otomi languages, verbs are commonly inflected for TAM by means of complex sets of markers that precede the stem. The number of tenses varies across languages, with conservative languages having more distinctions than innovative ones. The markers in question often cross-reference the subject too. Example (4) illustrates the paradigm of the present realis in the Northern Otomi variety of San Ildefonso Tultepec (SIT).

- (4) a. *dí=xóf-o*
 SIT 1.INCPL.R=harvest.maize-F
 ‘I harvest.’

3. Colonial Otomi is the Northern dialect of Old Otomi, which was the common language spoken until the beginning of the 18th century, just before the dialectal spread that gave rise to the family.

- b. *gí=xóf-o*
2.INCPL.R=harvest.maize-F
'You (SG) harvest.'
- c. *xóf-o*
[3.INCPL.R]harvest.maize-F
'He/she/they harvest.'

The inflectional markers in (4) behave like affixes for morphological purposes, but they are not typical affixes in two other senses. Phonologically, despite being unstressed by default, they can become stressed in dubitative speech, as shown in (5) where the suspension points '...' represent a prosodic pause. Syntactically, the markers are hosted on preceding functional words, as shown in (6). For these two reasons, I treat these markers as simple clitics rather than just prefixes (Textual examples are indicated as 'Text').

- (5) *ja bi... ja bi=h<y>ant=ú*
SIT LOC.FOC 3.CPL LOC.FOC 3.CPL=<SS>see[3OBJ]=3PL
'It was there he saw them.' (Txt)

- (6) *hín=df né=gá=ú*
SIT NEG=3.PRS.R want=1SG=3PL
'I don't want them.' (Txt)

On the other hand, most verbs in Otomi consist of a morphological root and a stem formative (marked with +) (e.g. HUE 'yá+i 'carry water', ts'ú+gi 'get thin', ú+ts'i 'catch, fish'). The finite forms of these verbs may surface in two phonological shapes (in the sense used by Zwicky 1990), namely a "free" and a "bound" shape. The occurrence of these shapes responds to a number of prosodic and syntactic conditions.

For example, a free shape (F) is required at a clausal boundary, as shown in (7). In contrast, the verb occurs in a bound (B) shape when it occurs intraclausally, that is, when the finite verbal form is followed by an NP or other constituents within its own clause, as illustrated in (8). A verbal form in the bound shape is phonologically dependent and behaves like a clitic hosted on the first free word of the next syntactic constituent. In (8), the host of the verb is the definite determiner *no* and carries the main stress of the complex phonological word /bi.,mũ.ja.'no.r/ (the verb stem receives a secondary stress).

- (7) [*pá bi=müts'-i*]
SIT PURP 3.CPL=pile.up[3OBJ]-F
'To pile it up.' (Txt)

- (8) [*pa bi=müx-a=[nõ=r hái]*]_{NP}
SIT PURP 3.CPL=pile.up.AS-B=DEF.SG=SG soil
'To pile up the soil.' (Txt)

The determiner cluster *no=r* in Example (8) serves in turn to illustrate that nouns co-occur with number determiners (singular vs. plural) as well as definiteness determiners (definite vs. indefinite).⁴ More examples are given in (9).

- (9) a. [*nu=ʼa* *ra Teo*]_{NP} *di=ho^hk-a=[ndʼa*
 HUE BASE=DEF.SG SG Teo 3.INCPL.IRR=make-B=INDE.SG
 ra ʼyü]_{NP}
 SG path
 ‘Teo will be making a path.’ (E&V 2007:4)
- b. *ndí=thötʼ-a=thó* *kó* [*ʼna=r* *du^htu*]_{NP}
 SIT 1.IMPF=tie[3OBJ]-B=DEL with INDE.SG=SG cloth
 ‘I used to tie it up with a cloth.’ (Txt)

In the next section, I introduce a very specific constructional frame with the typical syntax of complementation in which one finds many of the verbal nominalizations in Otomi. I will refer to such a frame as the “nominalized complement construction.” I first introduce the frame with some emphasis on the similarities between Eastern and Northern Otomi. In later sections I elaborate further on the interesting ways in which they differ.

3. The nominalized complement construction

All Otomi languages have a complement construction in which a complement-taking verb selects a verbal nominalization of the action noun type as its complement. I call this the nominalized complement construction (henceforth NCC). Examples of the NCC appear in (10) and (11). Example (10) is from Eastern Otomi, while the examples in (11) come from two different varieties of Northern Otomi (San Ildefonso Tultepec Otomi and Mezquital Otomi, respectively). In these examples, matrix verbs and NP complements with the nominalization are in boldface.

- (10) *di=tho^hk-a=[ra* *nxadi]* *ma* *tsʼuntʼu*
 HUE 3.INCPL.IRR=keep[3OBJ]-B=SG learning 1POSS boy
 ‘My son will keep on studying.’ (E&V 2007:352)
- (11) a. *pa* *da=ma[=r* *sü]*
 SIT PURP 3.INCPL.IRR=ss/go=SG firewood.making
 ‘So that he’d go to make some firewood.’ (Txt)
- b. *xtä=gwa[=rä* *ndoni]* *yä za...*
 MEZ 3.PRF.IRR=ss/stop.AS=SG blossoming PL plant
 ‘When the flowers stop blossoming...’ (Salinas Pedraza 1983:194)

4. The number determiners have been erroneously characterized in the literature as definite articles.

Table 1. Customary action nouns from SIT Otomi occurring in the NCC

Grooming/body function:	<i>njode</i> 'dressing'; <i>moge</i> 'undressing'; <i>mpit'i</i> 'pissing'; etc.
Food/drink consumption:	<i>ts'ithe</i> 'having a drink'; <i>n̄hūni</i> 'having a meal'; etc.
Food processing:	<i>juni</i> 'corn grinding'; <i>t'aphi</i> 'mead extracting'; <i>ntho'tsè</i> 'cooking'; etc.
Farming/household jobs:	<i>mpohmi</i> 'plowing'; <i>bot'i</i> 'sowing'; <i>sofo</i> 'maize harvesting'; <i>'bamboi</i> 'cattle tending'; <i>njuhwo</i> 'fishing'; <i>'bejwa</i> 'hunting'; <i>t'ēdi</i> 'sewing'; <i>njo^hki</i> 'sweeping'; etc.
Other activities:	<i>nt'opho</i> 'writing'; <i>sadi</i> 'praying'; <i>mā</i> 'selling'; etc.
Work in general:	<i>'befi</i> 'working'; etc.

(13) *nú='mu* *bi=gwá* [=r *t'ēdi*]
 SIT BASE=when 3.CPL=ss/finish.AS.B=SG sewing
 'When she finished sewing' (Txt)

(14) **nú='mu* *bi=gwá* [=r *ntsági*]
 SIT BASE=when 3.CPL=ss/finish.AS.B=SG jumping
 Intended reading: 'When she finished jumping'

To express the intended meaning in (14), language users revert to another complement construction. This other construction consists of a finite subordinated clause that is syntactically juxtaposed to the matrix predicate, as in (15).⁶ This construction, involving a juxtaposed dependent clause, is very common in the languages of this family for the encoding of notions typical of syntactic complementation, as for example in (16a). The same structure is also used to encode purposive semantics, as in (16b).

(15) *nú='mu* *bi=gwád-i* [*bi=n* *tság-i*]
 SIT BASE=when 3.CPL=ss/finish-F 3.CPL=INFL jump-F
 'When she finished jumping'
 (Lit. 'When she finished it, she jumped.')

(16) a. *hing=i=n* *nde* [*da='y>qt'-e*]
 HUE NEG=3.INCPL.R=INFL want 3.INCPL.IRR=<SS>do[3OBJ]-F
 'It was there he sat to sell.' (V&E 2007: 269)
 (Lit. 'He doesn't want it, he'll do it.')

b. *ja* *bi=h-ñ>úh=pu=thó* [*bi=m-pa*] *'něhe*
 SIT LOC.FOC 3.CPL=<SS>sit.AS=there=DEL 3.CPL=ANTIP-sell also
 'It was there he sat to sell.' (Txt)
 (Lit. 'It was there he sat, he sold also.')

6. For a full account of why such juxtaposed clauses instantiate subordination rather coordination, see Palancar (2012).

Although the customary activity required by the NCC is in the default case performed by a human being, we also find examples referring to higher-ranking animals and even other spontaneous events involving plants, such as the blossoming of flowers, as in Example (11b) above, or inanimate participants such as the fermenting of an alcoholic drink. What appears to be crucial for the construction is that the subject performs the action depicted in the nominalization as a customary activity, whatever its type.

3.2 The complement nominalization is an action noun

Nominalizations in Otomi languages are produced by conversion from verbal stems. In this respect, verbs can have up to four different stems for inflectional purposes. These stems may then be later used for derivational purposes through conversion. Examples are given in (17). The primary stem is the lexical representation of the lexeme (in the sense used by Aronoff 1994). Both the secondary and the impersonal stems derive from the primary stem. A fourth stem, which I refer to as the “stative stem,” is in turn derived from the impersonal stem by means of a nasal prefix.

(17)	HUE	PŪTI	HONI	'OT'I
		'peck'	'look for'	'write'
	Primary stem	pŭt-	hon-	'ot'-
	Secondary stem	mbŭt-	hyon-	'yot'-
	Impersonal stem	'bŭt-	thon-	t'ot'-
	Stative stem	m-'bŭt-	n-thon-	n-t'ot'-

Not all verbs have these four stems. For example, derived verbs by a middle or antipassive prefix can only have a primary stem, so nominalizations from those verbs naturally come from this stem only. But if a verb has more than one stem, it is often the case that the morphological nominalizations of these verbs come from stems other than the primary one. The secondary stem is mainly involved in the production of agent nouns, while the impersonal and the stative stems yield other types of noun: action nouns, instrument nouns, result nouns, and so on. In (18), I give some examples of actual nominalizations of the verb *pu* 'grind sugar cane' in Eastern Otomi.

(18)	HUE	v.	n.	
		PŪ		
		'grind cane'		
	Primary stem	pŭ	> —	
	Secondary stem	mbŭ	> MBŪ	'sugar cane grinder' agent n.
	Impersonal stem	'bŭ	→ 'bŭ	'sugar cane grinding' action n.
			> 'bŭ ₁	'sugar mill' instrument n.
	Stative stem	m-'bŭ	> M'BŪ ₁	'ground sugar cane' product n.
			M'BŪ ₂	'sugar cane grinding season' time n.

In general, the resulting noun is an independent lexical item from the verbal lexeme, except for the action noun in Eastern Otomi. The action noun in this language could be taken as a (syntactic) nominalization of the lexeme for the context of the NCC; in this regard it still realizes the same verbal lexeme. The subtle differentiation between inflection and derivation in this context is represented by the use of the arrow \rightarrow rather than the greater-than sign $>$. A different situation applies in Northern Otomi, in which action nouns behave like deverbal nouns.

However, conversion rules appear to be unproductive in the Otomi languages. This means that all deverbal nouns existing in the lexicon are the outcome of past operations effected when the rules were still productive. Today, Otomi speakers often rely on loanwords from Spanish for new concepts.

Most nominalizations of the action noun type are conversions of the impersonal or the stative stems. As these are the types of nominalization we find in the NCC in the two languages, it is worth having a look at the rules involved in the building of these stems in more detail. In Table 2, I have spelled them out for Eastern Otomi, the situation being very similar for Northern Otomi. The sign $>$ in Table 2 now indicates morphophonological derivation.

An impersonal stem is available for the inflection of all transitive verbs, so the building rules in Table 2 are productive for inflectional purposes. The different affixes used to form the impersonal stem are conditioned by the phonology of the primary stem. If the stem has a glottal in onset, a *t*- prefix is used; if it has an aspirated consonant, a nasal prefix is used (with some lexical exceptions). For all other cases, the rules either involve a glottal affix (such as a stop or a fricative), which is infixes with onset stops, or they involve consonant mutation of the onset resulting into a hardening outcome. This mutation is the historical reflex of the affixation of the same glottal prefix that we still observe in other verbs.

The action nominalizations occurring in the NCC are commonly from active intransitive verbs. As the default for these verbs is to use an impersonal stem, this suggests that such a stem is also available to them, at least for derivational purposes. A few examples of the default rule are given in (19). Northern Otomi behaves in this respect like Eastern Otomi, but for this language the process is derivational and has lexical gaps, and a sign $>$ should be employed instead.⁷

7. A note about tone in nominalizations: At times, the tone of a deverbal noun differs from the one we observe in the primary stem of the verbal lexeme. This commonly corresponds to tonal changes that also apply to the inflectional stems, but the phenomenon is more complex and needs further study. As for derivation, there do not seem to be productive tone-changing rules when there is a change in word class, but when the tone of a given deverbal noun differs from the inflectional stem from which it emerged, this could be taken as a sign of lexical independence from the original lexeme. Echegoyen and Voigtlander (2007) at times give

Table 2. Building rules for the impersonal and stative stems in Eastern Otomi

v. tr.		Primary st.		Imper. st.		Stative st.
HONI	'look for'	honi	> t-	t-honi	>	n-t ^h oni
HWÁHNI	'choose'	h ^w á ^h ni	>	t-há ^h ni	>	n-t ^h á ^h ni
'ÚDI	'show'	?údi	>	t-ʔúdi	>	n-t'údi
'WĚTI	'sew'	'wē ^h ti	>	t-'ē ^h ti	>	n-t'ē ^h ti
MÁ'T'I	'guard'	máʔt'i	> h-	h-máʔt'i	>	m- ^h máʔt'i
MBA'T'I	'shout at'	m ^b aʔt'i	>	h-m ^b aʔt'i	>	m- ^h m ^b aʔt'i
NŪ	'see'	nū	>	h-nū	>	n- ^h nū
NDŌNI	'affect'	n ^d ōni	>	h-n ^d ōni	>	n- ^h n ^d ōni
YŌ'T'E	'knock over'	jəʔt'e	>	h-jəʔt'e	>	n- ^h jəʔt'e
PĒ	'steal'	pē	>	p ^h ē	>	m-p ^h ē
KÓDE	'wear (below waist)'	kóde	>	k ^h óde	>	n-k ^h óde
TŪTS'I	'haul'	túts'i	> ʔ-	t ^h úts'i	>	n-t'úts'i
TSA	'eat/bite'	tsa	>	ts ^h a	>	n-ts ^h a
BĒNT'I	'grasp'	bent'i	> HARDENING	p ^h ent'i	>	m-p ^h ent'i
XÁ'TS'I	'rinse'	járts'i	>	ts ^h árts'i	>	n-ts ^h árts'i
ZÓFO	'speak to'	zóp ^h o	>	ts ^h óp ^h o	>	n-ts ^h óp ^h o
PĒ'MI	'slap'	pé'mi	>	'bé'mi	>	m-'bé'mi
FĒI	'whip'	p ^h ēi	>	'bēi	>	m-'bēi
TÓHNI	'push'	tó ^h ni	>	'dó ^h ni	>	n-'dó ^h ni
FĪDI	'use as a mattress'	p ^h īdi	> N-	m-p ^h īdi		
THŪTI	'burn/sear'	t ^h ūti	>	n-t ^h ūti		
JŪT'I	'close'	k ^h ūt'i	>	n-k ^h ūt'i		

(19) v.intr. nom. from impersonal stem

HUE

ZÓNI	→	ts'oni	'weeping'
PŪ	→	'bu	'grinding cane'
HI	→	thi	'bathing'
KŪNI	→	jūni	'grinding'

information for deverbal nouns in the verbal entries; sometimes the tone of these nouns is different, but they do not say whether the same tone change applies in the nominalization that occurs in the NCC. For this reason, and because my understanding of tone is poor in Otomi on the whole, I have not indicated tone in any nominalization.

- (23) *dá=du'm-i* [*dá=wä'-a=n'da*
 HUE 1.CPL=start-F 1.CPL=clean.cornfield.AS-B=INDE.SG
ma hwähi
 1POSS cornfield
 'I started to clean one cornfield of mine.' (E&V 2007:84)
 (Lit. 'I started it, I cleaned one cornfield of mine.')

In their vast dictionary of Eastern Otomi, Echegoyen and Voigtlander (2007) include 3,000 sentential examples to illustrate the usage of the different entries. However, there is only one Example (24) whose translation can mislead us into believing we are witnessing a transitive nominalization. The structure in question is not an instance of the NCC, but of a dative external possessor construction.

- (24) *ga=hon-b-a=[rá* *thai*]NP
 HUE 1.DEF.IRR=search.DTR-3DAT-B=SG.3POSS purchasing
*[ra 'yomm-a-n'i]*NP
 SG get.dry-B-chilli.pepper
 'I'll get money to buy dry chili peppers.' (E&V 2007:495)
 (Lit. 'I'll get its purchasing to the dry-chilli pepper.')

Nominalizations in Eastern Otomi can also work as independent action nouns outside the NCC, and their subject can be encoded as a possessor. In Otomi languages, possession is marked on the head. The possessed noun is marked as possessed by means of a high tone that (commonly) lands on the number determiner, as shown in (25). The NP encoding the possessor is simply juxtaposed. Compare for this purpose (25) with (26), where the possessed noun is the deverbal noun of the verb *PEFI* 'work'.

- (25) *ra mbidi, di=ja* [*rá t'uhni*^{POSSD} [*a*
 HUE SG fright [3.INCPL.R]INFL=exist SG.3POSS child DEF.SG
ra xitsu]^{POSSR}]
 SG woman
 'A fright is what the child of the woman has.' (E&V 2007:14)
 (Lit. '...her child the woman')
- (26) [*nu=rá* *'befi*^{POSSD} [*yö xadi-t'uhni*]^{POSSR}]
 HUE DEF=SG.3POSS work PL learn-child
di=n-xah=yö *t'ohni*
 [3.INCPL.R]INFL=MIDDLE-learn.AS.B=PL letter
 'The job of students is to learn the letters.' (E&V 2007:33)

In the NCC, however, the subject of the nominalization is always controlled by the matrix clause, as shown in (27). In other words, one of the arguments of the matrix predicate has to be co-referential with the subject of the nominalization (Stiebels 2007). As a consequence of this, the controlled argument is never expressed in the

complement and is thus never encoded as a possessor. Example (27b) shows the neutral word order VOS in Eastern Otomi.

- (27) a. *bi_i=ben-i* *bi_i=mba* [*ra nxadi* ____i]
 3.CPL=think-F 3.CPL=ss/go SG learning
 ‘He thought of going to school.’ (V&E 1979 [1985]: 305)
 (Lit. ‘He thought of it, he went studying.’)
- b. *bi_i=jwah-te* [*na ts’ithe* ____i] *ra fani_i*
 HUE 3.CPL=finish.AS-ANTIP INF water.drinking SG mule
 ‘The mule finished drinking water.’ (V&E 1979 [1985]: 235)
 (Lit. ‘It finished water-drinking the mule.’)

The intransitivity of nominalizations in Otomi is a property found in other unrelated language families in the Mesoamerican linguistic area, with the same function of expressing subject-oriented activities. It is reported, for example, in Mayan languages, such as in K’ichee (Can Pixabaj 2009), Tzutujil (Dayley 1985), Mam (England 1983), Chol (Vázquez 2007, 2013), and Tzeltal (Polian 2013), but also in Classical Náhuatl (Uto-Aztecan, Carochi [1645] 1983) and Ayutla Mixe (Mixe-Zoquean, Romero 2013).

In this section, I have presented a set of properties of the two NCCs of Eastern and Northern Otomi that are common to both of them. I have shown, for example, that the nominalization occurring in the frame is of the action noun type, that it commonly comes from intransitive verbs, and that it depicts actions performed customarily by humans. In the next section, I explore the differences between the two constructions.

4. Exploring the differences between the two NCCs

Despite their structural similarities, the NCC of Eastern Otomi differs greatly in its distribution from the NCC of Northern Otomi. I argue that the Eastern construction allows for a nominalization of the inflectional type, whereas its Northern counterpart has lexical restrictions that suggest the frame is working as a moribund syntactic structure: that is, it is almost idiomatic, and can only hold deverbal action nouns in the nominalization slot.

To show this, I first examine the lexical behavior of both matrix verbs and nominalizations in the NCCs of both languages. I show that there are no lexical restrictions in Eastern Otomi, which is not the case in Northern Otomi. Then I show that the nominalization in the Eastern Otomi construction is preceded by a marker that works as an infinitival marker, whereas in Northern Otomi only the singular determiner is possible. This can be taken as morphosyntactic evidence that the nominalization in Eastern Otomi is inflectional, a characteristic further supported by the common use of the construction in discourse.

4.1 The matrix verbs in the NCC

The NCC in Eastern Otomi allows for about twenty matrix verbs of the ones typically found in complementation structures typologically (Cristofaro 2003). The verbs in question are given in Table 3.

Table 3. Matrix verbs in the NCCs of Eastern and Northern Otomi

Control predicates		HUE	STT	
purposive motion	'go to do sth.'	<i>pa</i>	<i>pa</i>	
desiderative	'want to do sth.'	<i>ndé</i>	–	
liking	'like to do sth.'	<i>hō</i>	–	
experiential knowledge	'know how to do sth.'	<i>pādi</i>	–	
	'learn to do sth.'	<i>xadi</i>	–	
modal	'try to do sth.'	<i>tsa</i>	–	
	'help to do sth.'	<i>fats'i</i>	<i>fats'i</i>	
other	'achieve having done sth.'	<i>hwégi</i>		
Phasal and causative				
phasal	'stop/finish doing sth.'	<i>jwádi</i>	<i>jwádi</i>	
		<i>jwáhte</i>		
	'stop doing sth.'	<i>hegi</i>	–	
	'start doing sth.'	<i>judi</i>	–	
	'begin doing sth.'	<i>dú'mi</i>	–	
	'keep on doing sth.'	<i>thógi</i>	–	
	causative	'let sb. do sth.'	<i>hegi</i>	–
		'take sb. to do sth.'	<i>tsixte</i>	
'put sb. to do sth.'		<i>ži</i>	–	
'tell sb. to do sth.'		<i>xifi</i>	–	
	'order sb. to do sth.'	<i>'bépi</i>	–	

Table 3 shows that the NCC of Northern Otomi only allows three verbs: the verb *pa* 'go' used as a purposive motion matrix verb with the meaning 'go elsewhere to do an action'; the modal verb *fats'i* 'help to do an action' and the phasal *jwadi* 'finish doing an action'. Of the three, the motion verb is by far the most common one in usage, and by itself it accounts for more than ninety percent of the tokens of the construction in Northern Otomi. This language allows a roughly similar number of matrix verbs as Eastern Otomi, but they are used in other complementation structures, mainly the juxtaposed complement construction we saw in Section 3.1. This means that when a

speaker of Eastern Otomi uses the NCC, to render its closest semantic equivalent a speaker of Northern Otomi must use an alternative construction (see Section 4.3. for a discussion).

4.2 Lexical gaps in the nominalization slot of the NCC

As I pointed out in Section 3.1 above, the NCC is used for a customary activity typically performed by a human being or by other higher-ranking animals. In Eastern Otomi, all possible verbs that conform to this semantic profile have a nominalization readily available to be used in the frame of the NCC. The same is apparently not possible for Northern Otomi. Table 4 contains a list of nominalizations that can occur in the NCC in the San Ildefonso Tultepec variety of Northern Otomi. Crucially, the list in Table 4 can be considered exhaustive.

Table 4. Exhaustive list of action nouns in the NCC in Northern Otomi

SIT	V.			action nouns
grooming/body function:	KÓDE	tr.	>	NJODE 'dressing'
	MÓGE	intr.	>	MOGE 'undressing'
	MPÍ'T'I	intr.	>	MPIT'I 'pissing'
food/drink consumption:	'E+NÖ	intr.	>	N'ENÖ 'combing'
	TSĪ+DÉHE	intr.	>	TS'IDEHE 'having a drink'
food processing:	NŪNI	intr.	>	NHŪNI 'having a meal'
	N'ŪSPI	intr.	>	N'ŪSPI 'making a fire'
	KŪNI	intr.	>	JŪNI 'grinding'
	KŪNI+N'Ī	intr.	>	JŪNIN'I 'chili-grinding'
farming/household jobs:	HÓ'TS'E	tr.	>	NTHO'TS'E 'cooking'
	MPOHMI	intr.	>	MPOHMI 'plowing'
	PÓT'I	tr.	>	'BOT'I 'sowing'
	XÓFO	intr.	>	SOFO 'harvesting'
	'AFI	intr.	>	T'AFI 'extracting maguey-mead'
	JO+JŪ	intr.	>	NJOJŪ 'collecting beans'
	HONI+K'ANI	intr.	>	THONGAR'ANI 'searching for greens'
TŪ ^h KI+LÓLÓ	intr.	>	THU ^h KLOLO 'plucking fruits'	

(Continued)

Table 4. (Continued)

SIT	V.		action nouns	
	TÚ ^H KI+RÁXT'Ö	intr.	>	THU ^H KA'RAXT'Ö 'plucking prickly pear fruit'
	TÚ ^H KI+KÓHÖ	intr.	>	THU ^H KAKÖHÖ 'plucking prickly pear leaf'
	FÁDI+BÓI	intr.	>	'BAMBOI 'looking after cattle'
	HO+TS'ÚDI	intr.	>	THO ^H PTS'ÚDI 'killing of the pig'
	HO+BÓI	intr.	>	THOBOI 'killing of cattle'
	HO+KÓDO	intr.	>	THO ^H KDO 'killing of turkey'
	JU+HWÓ	intr.	>	NJUHWÖ 'fishing'
	PÉ+JWA	intr.	>	'BEJWA '(rabbit) hunting'
	XÚ	intr.	>	SU 'gathering firewood'
	'WÉDI	intr.	>	T'ÉDI 'sewing'
	NJO ^H KI	intr.	>	NJO ^H KI 'sweeping'
other activities:	'OFO	intr.	>	NT'OFO 'writing'
	JÜTS'I+JÖDO	intr.	>	NJÜXJÖDO 'fencing'
	XÁDI	tr.	>	SADI 'praying'
	PĚ ^H TI	tr.	>	'BĚ ^H TI 'begging'
	PĚ	tr.	>	MFĚ 'stealing'
	PĀ	tr.	>	M'Ā 'selling'
	Ñ'ÁDI	intr.	>	Ñ'ÁDI 'asking for information'
work in general:	PEFI	tr.	>	'BEFI 'working'

The exhaustiveness of Table 4 has an important consequence for the NCC: it reveals that the lexicon allows for certain action nouns to occur in the structure, but not others, and it suggests that lexical selection in the NCC of this language pertains to the realm of derivation. As the existing corpus material would not render sufficient information, the list was gathered from extensive elicitation sessions with two fluent speakers.

In this respect, variation in the number of items allowed for each individual speaker is expected, but the speakers consistently refused a set of nominalizations in the frame. For example, the verbs in (28) have the right semantic profile to be allowed in the construction (i.e., they depict customary actions) and the cognates of such verbs occur freely in the constructional counterpart in Eastern Otomi. Such verbs have deverbal nouns in the lexicon that in principle (if the nominalizing rules were more transparent semantically) could be used in the NCC with an action meaning. In reality, they cannot because such a meaning is unavailable.

(28)	SIT	v.		deverbal noun	potential action noun meaning
	POHO	intr.	>	FOHO	'excrement' *'defecating'
	'ÓHÖ	intr.	>	T'ÖHÖ	'sleep' *'sleeping'
	'WĪ	intr.	>	T'Ī	'dream' *'dreaming'
	NTSĪM+XUDI	intr.	>	NZIMXUDI	'luncheon' *'lunching'
	PŪ	intr.	>	'BŪ	'cane' *'grinding cane'
	WÖHI	intr.	>	HWÖHI	'cornfield' *'plowing'
	HĒNI+ZÁ	intr.	>	THEENZA	'wood' *'cutting wood'
	ÑÖ	intr.	>	HÑÖ	'word, speech' *'talking'
	NĒI	intr.	>	HNĒI	'dance' *'dancing'
	PĒ+ ³ BÍDA	intr.	>	'BĒ'BIDA	'music, violin' *'playing music'
	BĒNI	intr.	>	MFĒNI	'thought' *'thinking'
	ZÓNI	intr.	>	TS'ONI	'cry' *'crying'
	NTHU'TS'I	intr.	>	NTHU'TS'I	'kiss' *'kissing'
	HÁ ^H TE	intr.	>	NTHA ^H TE	'fraud' *'being fraudulent'

The examples in (28) show that actual nominalizations of verbs, which could in principle occur in the NCC, cannot do so because they lack an action noun meaning. I take this restriction as a piece of evidence that the lexical items occurring in the nominalization slot of the NCC in Table 4 are now independent lexical items from the verbal lexemes from which they were once produced as morphological nominalizations. But when the conversion rules ceased to be productive such nouns lost their semantic connection with their source. In other words, the nouns occurring in the NCC of Northern Otomi are synchronically deverbal action nouns, and no longer represent inflectional non-finite forms of their respective verbal lexemes. Further evidence that this analysis is correct is the fact that some other verbal lexemes whose actions could be equally construed as customary have no nominalizations available at all, as for example the verbs in (29), whereas their cognates can be used in Eastern Otomi.

(29)	SIT	PENI	>	*MBENI ⁸	'doing the laundry'
		KAT'A+HME	>	*JAT'AHME	'making tortillas'
		XÁHA	>	*NTXAHA	'bathing, swimming'
		'YO	>	*Ñ'O	'going for a walk'

4.3 The use of the NCC with respect to other structures

The NCC competes with its finite counterpart, the juxtaposed subordinated clause, for the expression of the semantics of customary activities. Examples (30) and (31) come

8. Variation is expected across the different dialectal areas of Northern Otomi. Mezquital Otomi has the action noun 'beni' 'doing the laundry' as illustrated in (1) above, while in San Ildefonso Tultepec Otomi such a noun does not exist, apparently.

are also examples of the NCC in which the nominal complement also co-occurs with the number determiner *ra*, as in (10) and (27a) above. Such examples are instances of action nouns, as in Northern Otomi.

However, the default marker preceding the nominalization in the NCC of Eastern Otomi is *na*, as shown in (32). Following Voigtlander & Echegoyen (1979 [1985], 2007), I analyze this marker as an infinitive marker, and more concretely as an “intransitive infinitive.”

- (32) a. *'b_u má=mba b_u däthe [na thi]...*
 HUE when 1.IMPF.BELOW=SS/go P.LOC river INF.INTR bathing
 ‘When I used to go down to the river to bathe...’ (E&V 2007: 122)
- b. *bi=za-pi [na zete]*
 3.PST=SS/try-3DAT INF.INTR working.for.others
 ‘He tried working for others.’ (V&E 1979 [1985]: 232)

The infinitival construction is interpreted by clausal syntax as an NP. This can be seen, for example, in both the word order and in how the verb reacts to it morphophonologically. As an illustration, consider Example (33) with VOS order, where the structure *na 'b_e'bida* ‘playing music’ occurs in the position of the object.⁹ Notice as well that the matrix verb occurs in bound shape.

- (33) *nu='b_u bi=mb_uh=[na 'b_e'bida]_{OBJ}[='b_u*
 BASE=when 3.PST=start.AS.B=INF.INTR music.playing=DEF.PL
zi mina]_{SUB}...
 DIM squirrel
 ‘When the little squirrels started playing music...’ (V&E 1979 [1985]:26)

In the light of these data, analyzing *na* as an intransitive infinitive marker is comparable to saying that it has an inflectional function. In other words, I take it to encode the productively nominalized form of the complement verb as a non-finite form. For example, when the verb has more than one morphological nominalization, the form co-occurring with *na* refers unequivocally to the action of the verb; that is, it realizes the verbal lexeme. This is shown in (34). Other nominal forms of the verb with other meanings naturally co-occur with the singular determiner, but never with *na*. In other words, the *na* marker in Eastern Otomi is an infinitival marker similar to English *to* in forms such as *to grind*, *to sleep*. The Otomi non-finite structure ‘*na* + NOM’ would be roughly equivalent to made-up forms in English that would look like **to grinding*

9. I want to thank an anonymous referee for making me aware that this does not account for the position of the infinitival phrase in (32a). For this structure, the infinitival phrase appears to function as an adjunct purpose clause, and should thus perhaps be considered as an instantiation of typical extensions that infinitives have in better-known languages.

or **to sleeping* for the syntactic contexts where just *to grind* or *to sleep* would typically be used.

(34)		INE.INTR		deverbal
	T _{AI}	'buy'	→ na th _{ai}	'buying' > ra T _{AI}
	P _U	'grind cane'	→ na 'bu	'grinding cane' > ra 'B _U
				> ra M'B _U
				'ground sugar cane'
	'ÁHĀ	'sleep'	→ na t'ähā	'sleeping' > ra T'ÁHĀ
	PĒ	'steal'	→ na mfē	'stealing' > ra MFĒ
	PI	'relieve oneself'	→ na 'bi	'relieving oneself' > ra 'BĪ
				> ra MPĪ
				'excrement' > ra MPĪ
				'restroom' > ra NTHĪ
	HI	'bathe'	→ na thi	'bathing' > ra NTHĪ
				> 'bath utensils'

At times, the nominalization in question can be used independently as a proper action noun and thus receive the determiner singular too. When this happens, both forms can be used in the NCC, as in the instances in (35).

(35)		INE.INTR		deverbal
	'ÁTS'I	'fish'	→ na nt'ats'i	'fishing' > ra NT'ATS'I
				> 'net'
	XADI	'learn'	→ na nxadi	'learning' > ra NXADI
				> 'classroom'
	XÚ	'gather firewood'	→ na s _u	'gathering > ra s _u
			firewood'	firewood'

There is another piece of evidence in favor of treating *na* as an infinitival marker. In Eastern Otomi, there is a transitive infinitival counterpart marker *dí*. Transitive infinitives are typically used when the complement clause functions as the subject of unaccusative matrix verbs, as for example in (36).

(36)	a.	xó=n-hëi	[dí	t _{ai}	ra	dethä]
	HUE	PRE.ST=ST-be.difficult	INF.TR	buy	SG	maize
		'It's difficult to buy maize.'				(V&E 1979 [1985]: 234)
	b.	bi=za	[dí	pet'i	ra	hme]
		CPL=ss/be.possible	INF.TR	make.tortillas	SG	tortilla
		'It was possible to make tortillas.'				(V&E 1979 [1985]: 170)

The forms *t_{ai}* in (36a) and *pet'i* (36b) are in turn nominalized forms of the transitive verbs T_{AI} 'buy' and P_{ET}'I 'make tortillas' and are produced by the primary stem of such verbs. One piece of evidence that these are indeed non-finite forms comes from the fact that they cannot appear in bound shape, unlike other finite forms of verbs. This is illustrated in the contrast between (37a) and (37b), from Voigtlander & Echegoyen (1979 [1985]: 233).

- (37) a. *bi=manda* [dɨ́ wɛ́ti rá bɔ́tʰe]
 HUE 3.CPL=command INF.TR sew SG.3POSS skirt
 ‘She had her skirt sewn.’
 (Lit. ‘She commanded to sew her skirt.’)
- b. *bi=manda* [da=wɛ́t=rá bɔ́tʰe]
 3.CPL=command 3.INCPL.IRR=sew.B=SG.3POSS skirt
 ‘She told them to sew her skirt.’
 (Lit. ‘She commanded they’d sew her skirt.’)

As the transitive infinitive is built from the primary stem of a verb, younger speakers of Eastern Otomi occasionally also use the primary stem of a verb for the intransitive infinitive. As a result, speakers often have a free choice between two forms for the same function, as for example in (38).

- (38) HUE
 pɛ́fi ‘work’ → *na pɛ́fi* ~ *na bɛ́fi* ‘to work’
 zóni ‘cry’ → *na zóni* ~ *na tʰóni* ‘to cry’
 peni ‘wash’ → *na peni* ~ *na bɛ́ni* ‘to wash’
 mbõnni ‘go for a walk’ → *na mbonni* ~ *na hmbonni* ‘to go for a walk’

A number of questions come to mind: Where does the *na* marker of Eastern Otomi come from? Has Northern Otomi lost it? Witnessing the decay of the NCC in Northern Otomi, one could be easily tempted to think that the language has naturally lost the infinitival marker and now has only the equivalent construction with deverbal action nouns. However, the history of the forms tells a different story.

The intransitive infinitive marker *na* of Eastern Otomi is the reflex of the determiner *na* of Old Otomi, which was a singular determiner occurring with nominalizations. Pedro de Cárceres (1580[1907]:46) characterizes the nouns co-occurring with *na* as ‘abstracts’, as *na* precedes derived nouns of the type of *nānttāxi* (*na nʰāxi*) ‘whiteness’ or *nāphati* (*na foti*) ‘blackness’, but in reality it also occurs with concrete nouns (possibly morphologically derived from unknown sources) like *nadete* (*na rɛte*) ‘ladder’, or *nadættzæ* (*na rɔtsɛ*) ‘tailbone’. Old Otomi also had the singular determiner *no*, which co-occurred with verbal nominalizations to profile the action as a whole, such as *nothūhū* (*no thūhū*) or *nottogui* (*no tɔgi*) ‘sowing’ (Sp. ‘sembrada’) (Urbano 1605:382v). In contrast to *no*, *na* was used to profile the action of the verb as an activity, as in *nathūhū* (*na thūhū*) or *natɔgui* (*na tɔgi*), which are translated into Spanish as ‘sembradura’. For basic nouns, Old Otomi used the determiners *o* and *an*. The determiner *o* was a deprecative and was lost in all modern languages. The default determiner was *an*, which gave rise to the singular determiners *ra* and *=r* in Eastern and Northern Otomi.¹⁰

10. Ixtenco Otomi also has the marker *ni* with nominalizations. A cognate of such a marker is also found in Tilapa Otomi as *rɨ* which nowadays occurs in complementary distribution

The alveolar nasal in both determiners *na* and *an*, appearing in unstressed syllables, was rhoticized in the modern languages from the eighteenth century onward, possibly after having been weakened to an apical alveolar [d] at a previous stage. In the Southern languages, the modern reflexes of Old Otomi *na* and *an* are *ra* and *ar*, respectively. In the Northern language, however, both reflexes merged into one with different outcomes depending on the variety (San Ildefonso Tultepec Otomi has *ar* or simply =*r*, whereas Mezquital Otomi has *rā*).

In Tilapa Otomi, the most conservative language of the family, the two singular determiners *ra* and *ar*, as reflexes of historical *na* and *an*, have a distribution very similar to Old Otomi. The determiner *ra* lexically co-occurs with most nouns derived from the default conjugation class (e.g. *ra tshógi* ‘key’ from *xÓGI* (intr.) ‘open’). But with nouns of the third conjugation, Tilapa has *dí*, which is a cognate of the transitive infinitive marker in Eastern Otomi (e.g. *dí xadi* ‘prayer’ from *xADI* (intr.) ‘pray’). Otherwise, the default singular determiner is *ar*, which is now used both with basic nouns and with derived ones (e.g. *ar hā* ‘soil’, but also *ar ’i* ‘food, chili’ from *’i* (intr.) ‘be spicy’). As expected, in the NCC of Tilapa Otomi both *ra* and *dí* are used with the verbal nominalization, but not *ar*. This is shown in (39).

- (39) *tʃá=ndé* [*ra* *hněi*] / [*dí* ^h*pefi*]
 I.INCPL.R=want NOM.SG dancing / NOM.SG working
 ‘I want to dance/to work.’

What happened then in Eastern Otomi? In this language, the old *na*, the determiner that occurred with nominalizations in Old Otomi both inside and outside the NCC, escaped the rhotization of [n > d > r] inside the NCC. This is because the old determiner *na* had been reanalyzed by speakers as carrying a specific grammatical function in the construction, which was regarded as distinct from its other determiner uses outside the construction. The old determiner thus became an infinitive marker. In all other cases, just like in Mezquital Otomi, Eastern Otomi merged the old determiner *na* (outside the NCC) and *an* in favor of an outcome such as *ra*, possibly through intermediate stages with alternants *ra* and *ar*, as in Tilapa Otomi. This is why we find *ra* for all other nouns outside the construction, whether basic or derived.

The old function of *na* as a determiner co-occurring with nominalizations outside the NCC has left a few traces in Eastern Otomi in a few nouns derived from stative verbs such as *na n’i* ‘spiciness’ along with *ra n’i* ‘chili’ from *’i* (intr.) ‘be spicy hot’, *na nhō* ‘goodness’ along with *ra nhō* ‘(the) good’ from *hō* (intr.) ‘be good’, or *nxidi* ‘width’ from *xidi* ‘be wide’. Besides these nouns, it also occurs in a handful of idiomatic expressions, such as the ones in (40).

with *ar*, mainly with loanwords and as landing site for the suprasegmental marker of the third person possessor.

- (40) a. *i=tiü na mpa/ntse*
 HUE 3.INCPL.R=die SG.NOM heat/cold
 'He's hot/cold.' (V&E 1979 [1985]:236)
 (Lit. 'He dies (of) heat/cold.')
- b. *i=ʔh=na hyä*
 3.INCPL.R=hear.AS.B=SG.NOM word
 'He understands.' (V&E 1979 [1985]:236)
 (Lit. 'He hears word.')
- c. *di=hax=na ts'oni*
 [3.INCPL.R]INFL=dawn.AS.B=SG.NOM cry
 'He spends the night crying.' (V&E 1979 [1985]:235)
 (Lit. 'He dawns cry.')

The nouns *mpa* 'heat' and *ntse* 'cold' in (40a) occur only in this expression; otherwise, they are used as *pa* and *tse* and take the determiner *ra*. The deverbal nouns *hyä* 'word' in (40b) and *ts'oni* 'cry' in (40c) come from the verbs *yä* (intr.) 'speak' and *zoni* (intr.) 'cry', respectively. Outside these expressions, they must take *ra*.

Finally, there are other examples of the type in (41), which should not be taken to be a trace of the old function of *na* as a determiner co-occurring with nominalizations, like (40). Example (41) is an instance of an infinitival form of a verb being used within the syntax of an NP as subject. However, examples like (41) are problematic because a nominalized form of a verb, recategorized as a noun, should not in principle appear as a finite verb in bound shape when it is an NP. The fact that it does is quite puzzling, and should ideally be addressed in future research.

- (41) [*nu=ʔa ná ntih=ra fani*],
 HUE BASE=DEF.SG INF.INTR running.B=SG mule
ge=ʔ í=dä=ʔ
 COP=3SG 3.INCPL.R.ADV=SS/fall.AS=3SG
 'The running of the mule is the cause of its falling.' (V&E 1979 [1985]:234)
 (Lit. 'The running of the mule, it is that what she fell from.')

5. Conclusions

In this paper, I have studied the many ways in which a very similar grammatical construction involving a verbal nominalization can vary in two very closely related languages of Mexico, namely Eastern and Northern Otomi. The construction at issue is a special complementation construction requiring a nominalized complement that expresses a customary activity, such as 'dancing', 'cleaning', or 'harvesting'. The NCCs of these two languages illustrate possibilities in the grammatical space that exists between a highly nominal structure of the type 'I finished the dance' and another of the non-

finite type 'I finished dancing.' A summary of the properties proposed for the two NCCs is given in Table 5.

Table 5. Comparison of the two NCCs

	Eastern Otomi	Northern Otomi
The nominalization refers to a customary action	Yes	Yes
The nominalization in the NCC is intransitive	Yes	Yes
The nominalization in the NCC is of the action noun type	Yes	Yes
Number of matrix verbs allowed in the NCC (in relation to the ones requiring a juxtaposed clause)	All (20/20)	Very few (3/22)
The nominalization slot in the NCC has lexical gaps	No	Yes
An infinitive marker is to be used in the NCC	Yes	No
The language has other non-finite markers	Yes	No
Frequency of usage in relation to the juxtaposed clause	Frequent	Very infrequent

In the paper, I have shown that the nominalization of Eastern Otomi is a typical case of syntactic nominalization in a non-finite environment. In other words, it is morphologically inflectional in that the nominalized form of a verb is one of the available forms in the paradigm of the verb, as required by the syntax of the NCC. The nominal form of a verb is also available to all speakers. The fact that the nominalization is of the inflectional type, and that the language has developed a unique intransitive infinitival marker to be used in the NCC, makes this complement construction functionally vibrant. However, the nominalization processes in this language have the selection restrictions typical of derivation, in that the only type of verbs required to have a nominalized form are the ones that conform to the semantic profile of the NCC.

In contrast to Eastern Otomi, the nominalization of Northern Otomi should be accounted for as an instance of word-formation. This means that the nominalized form of a verb is not required in its paradigm when the verb is to occur in the syntactic frame of the NCC. As derivational rules in Otomi languages are no longer productive, if the action noun of a given verb exists it can be used in the NCC, by virtue of the lexicon. But as speakers do not share the same lexicon, there is much variation in both the use and the acceptability of the nominalizations in the construction. In reality, a very few action nouns account for most of the tokens used.

As the nominalization in Northern Otomi is derivational in character, and as derivational rules are no longer productive, the functional load of the NCC is very low. The common items found in the construction are probably learned as idiomatic expressions. However, as there still exist a substantial number of lexemes that can occur in the construction, making it feasible for the construction to exist in the first place, the

distribution shows some of the regularities expected of inflection. Heuristically, the NCC is a very useful descriptive tool for the linguist to find all the infrequent action nouns in the lexicon that otherwise do not occur in natural discourse.

Abbreviations

ANTIP	antipassive	LOC.FOC	locative focus
ADV	adverbial inflection	LOC.P	locative preposition
AS	adjusted stem	M	masculine
B	bound shape	MEZ	Mezquital Otomi
BASE	base	NEG	negative
CONT	continuous	NOM	nominalization
COP	copula	OBJ	object
CPL	completive	PRF	perfect
DAT	dative	PL	plural
DEF	definite	POSS	possessive
DIM	diminutive	POSSD	possessed
DTR	ditransitive	POSSR	possessor
E&V	Echegoyen & Voigtlander (see references)	PPRF	plu-perfect
F	free shape	PURP	purposive
HUE	Huehuetla Otomi	R	realis
IMPF	imperfect	SG	singular
INCP	incompletive	SIT	San Ildefonso Tultepec Otomi
IND	indicative	SS	secondary stem
INDF	indefinite	ST	stative
INF	infinitive	TIL	Tilapa Otomi
INFL	inflectional marker	TR	transitive
INTR	intransitive	V&E	Voigtlander & Echegoyen (see references)
IRR	irrealis		

Acknowledgments

This research was supported by the ESRC/AHRC project RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguean languages” and by the PICS 2015–2017 “Mésos-Amérique et la syntaxe de la proposition relative.” The original paper improved considerably as I presented it to different audiences. It benefited from the comments received at the *Escuela Nacional de Antropología e Historia* of Mexico in 2010 and at the *Surrey Linguistics Circle* in June 2011, but it was not until Anna M. Thornton had a look at it that I started to finally understand what I was talking about. Thank you Anna for all your help and patience! A close-to-final twist was given to it at the version presented at the international workshop on *Finitude et Nominalisation* organized by Claudine Chamoreau in Paris on October 12–14 2011. I want to heartily thank Claudine Chamoreau for inviting me to the workshop, where I benefited immensely not only from the discussions and papers on the subject but also from the wonderful collegial

atmosphere. I wrote the final version of the paper in Graz, Austria, during a short stay at the Linguistics Department of the Karl-Franz University from May 14 to 25, 2012. I am deeply grateful to Bernhard Hurch for his hospitality and for inviting me to stay at his home during all this time; it proved to be a perfectly quiet, orderly place to help me put my ideas down in writing, and those days will remain precious in my mind. The reader will be a judge as to the degree of success attained. Finally, my sincere thanks to an anonymous referee for his/her comments and suggestions. All errors and deficiencies remain my responsibility.

References

- Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes* [Linguistic Inquiry Monograph 22]. Cambridge MA: The MIT Press.
- Can Pinxabaj, Telma. 2009. Morphosyntactic Features and Behavior of Verbal Nouns in K'ichee'. MA thesis, University of Texas, Austin.
- de Cárceres, Fray Pedro. 1907[1580?]. *Arte de la lengua otomí*. In *Boletín del Instituto Bibliográfico Mexicano*, Nicolás León (ed.). Mexico: Instituto Bibliográfico Mexicano.
- Carochi, Horacio. [1645]1983. *Arte de la lengua mexicana con la declaracion de los adverbios della*. Edición facsimilar de la publicada por Juan Ruys en la Ciudad de México 1645, con un estudio introductorio de Miguel León-Portilla [Facsimiles de Lingüística y Filología Nahuas: 2]. Mexico: Universidad Nacional Autónoma de México.
- Comrie, Bernard & Thompson, Sandra. 2007. Nominalizations. In *Language Typology and Syntactic Description*, Vol. III: *Grammatical Categories and the Lexicon*, 2nd edn, Timothy Shopen (ed.), 334–381. Cambridge: CUP. doi:10.1017/CBO9780511618437.006
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.
- Dayley, Jon P. 1985. *Tzutujil Grammar* [University of California Publications in Linguistics 107]. Berkeley: University of California Press.
- Echegoyen, Artemisa & Voigtlander, Katherine. 2007. *Diccionario Yuhú: Otomí de la Sierra Madre Oriental (Estados de Hidalgo, Puebla y Veracruz, México)*. Versión Preliminar; Serie de Vocabularios y Diccionarios Indígenas. Mexico City: Instituto Lingüístico de Verano.
- England, Nora. 1983. *A Grammar of Mam, a Mayan Language*. Austin TX: University of Texas Press.
- Hernández Cruz, Luis, Torquemada, Moisés Victoria, & Sinclair Crawford, Donaldo (eds). 2004. *Diccionario del hñähñu (otomí) del Valle de Mezquital, Estado de Hidalgo* [Vocabularios Indígenas 45]. Mexico: Instituto Lingüístico de Verano.
- Muysken, Pieter. 1999. Nominalizations. In *Concise Encyclopedia of Grammatical Categories*, Keith Brown & Jim Miller (eds), 248–252. Oxford: Elsevier.
- Nikolaeva, Irina. 2007. Introduction. In *Finiteness: Empirical and Theoretical Foundations*, Irina Nikolaeva (ed.). Oxford: Oxford University Press.
- Lastra, Yolanda. 2001. *Unidad y diversidad de la lengua: Relatos otomíes*. Mexico: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Palancar, Enrique L. 2011. The conjugations of Colonial Otomi. *Transactions of the Philological Society* 109: 1–19
- Palancar, Enrique L. 2012. Subordination and juxtaposition. In *Clause Combining in Cross-linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*, Volker Gast & Holger Diesel (eds), 37–76. Berlin: Mouton de Gruyter.

- Polian, Gilles. 2013. Infinitivos transitivos: Innovaciones del tseltal en la familia maya. In *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamerica*, Enrique L. Palancar & Roberto Zavala Maldonado (eds), 339–380. Mexico: CIESAS.
- Romero Méndez, Rodrigo. 2013. Cláusulas de complemento en el mixe de Ayutla: Una aproximación a sus tipos morfosintácticos. In *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamerica*, Enrique L. Palancar & Roberto Zavala Maldonado (eds), 217–246. Mexico: CIESAS.
- Salinas Pedraza, Jesús. 1983. *Etnografía del Otomí*. Mexico City: Instituto Nacional Indigenista.
- Stiebels, Barbara. 2007. Towards a typology of complement control. In *Studies in Complement Control*, Barbara Stiebels (ed.), 1–80. Berlin: *ZAS Papers in Linguistics* 47.
- Voigtlander, Katherine & Echegoyen, Artemisa. 1979 [1985]. *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. Mexico: Instituto Lingüístico de Verano.
- Voigtlander, Katherine & Echegoyen, Artemisa. 2007. *Gramática del yuhú: Otomí de la Sierra Madre Oriental*. Mexico: Instituto Lingüístico de Verano. Ms.
- Vázquez Álvarez, Juan Jesús. 2007. Los tres tipos de estructuras de complemento en chol. In *Proceedings of the Conference on Indigenous Languages Latino-América III*. (http://www.ailla.utexas.org/site/cilla3/VazquezJ_CILLA_III.pdf)
- Vázquez Álvarez, Juan Jesús. 2013. Dos tipos de cláusulas no finitas en chol. In *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamerica*, Enrique L. Palancar & Roberto Zavala Maldonado (eds), 77–102. Mexico: CIESAS.
- Zwicky, Arnold M. 1990. Syntactic representations and phonological shapes. In *The Phonology-Syntax Connection*, Sharon Inkelas & D. Zec (eds), 379–397. Stanford CA: CSLI.

The evolution of number in Otomi

The many faces of the dual*

Enrique L. Palancar

SeDyL-CELIA (UMR8202), CNRS /

Surrey Morphology Group, University of Surrey

Otomi (Oto-Pamean, Oto-Manguan) is a small group of languages spoken in Mexico which has emerged as a linguistic family in recent times. In this paper, I study relevant changes in the number system of the Otomi languages. These changes constitute an interesting object of study to gain a better understanding of two main phenomena: aspects of language change involving the loss of the dual number; and the constraints on number agreement restricted to conjoining structures. I show that changes in the number system in the Otomi family have different outcomes depending on the dialectal area. Such outcomes include the loss of dual morphology through a semantic rearrangement where the old dual forms have become exponents of either paucal or plural number. In some of the varieties, the old dual is still preserved, but it is relegated to male speech as a token of partnership and camaraderie.

Keywords: Dual, number systems, split conjunction, Otomi, Oto-Manguan

1. Introduction

In this paper, I study relevant changes in the pronominal number system of Otomi which were triggered by the loss of the dual number. Otomi is a language family spoken in Mexico which belongs to the Oto-Pamean branch of Oto-Manguan. The first attestations of Otomi date back to the late 16th and the early 17th centuries and reveal a linguistic system which was the common ancestor of the modern languages that is known as ‘Old Otomi’ or ‘Colonial Otomi’. Given the large presence of the Otomi people across the valleys of Central Mexico from the Classic period and onwards, Old Otomi, the language they spoke, was most likely already a diversified dialectal continuum at that time rather than a unified language. From the 18th century onwards, this old dialectal continuum broke up into a small linguistic family which has come to consist of at least six different languages. Three

of these languages (Northern, Eastern, and Western Otomi) should in turn be viewed as dialectal continua consisting of various local varieties.

The paper is concerned about number, which is one of the most fundamental grammatical features in natural language, and more specifically about the dual number. Following Corbett (2000), which continues to be the most comprehensive study of number from a cross-linguistic typological perspective, the simplest and most widespread number system has an opposition between singular and plural values, as in the English nouns *dog* vs. *dogs*. In such a system, as Corbett (2000: 20) puts it, “[p]lural here refers to more than one real world entity”.¹ The next common number system has three values: singular, dual, and plural, and the dual here is a value that refers to two distinct real world entities,² as in Vedic (nominative) *tanús* ‘body’ vs. *tanúā* ‘two bodies’ vs. *tanúas* ‘(more than two) bodies’ (extracted from Macdonell 1910: 291).

On the other hand, number may involve nominals as well as pronominals,³ and hence we have nominal number and pronominal number, and languages may have both, either, or none.⁴ Pronominal number involves grammatical person, and it is commonly realized in personal pronouns and possessives (e.g. the first person pronouns of Gothic (nominative) *ik* ‘I’ vs. *wit* ‘we two’ vs. *weis* ‘we’, Agud and Fernández 1988: 139), but also in verb agreement (e.g. Vedic present (indicative, active) forms for the first person *bhárāmi* ‘I bring’ vs. *bhárāvas* ‘we two bring’ vs. *bhárāmasi* ‘we bring’, Forston 2004: 89).

In Old Otomi, the number system in pronominals was based on a singular-dual-plural opposition, but the dual became an unstable value in the breaking up of the old dialectal continuum. The instability is witnessed in that some of the modern languages stray considerably from the original situation, while others have preserved it. What is typologically interesting about the dual in Otomi is not so much its loss and the subsequent reanalysis that took place in the system to make up for it, but the traces left behind of its instability, which are reflected in curious places in the grammar. In Northern Otomi, for example, while dual still exists in the language, it only has a pragmatic function to express a sense of camaraderie between two male speakers. In contrast, in the Southern languages, Tilapa Otomi and Ixtenco Otomi, the old dual forms have become exponents of the plural, but there is slight variation depending on the specific degree of how far this change has advanced.

The overall goal of this paper is to present an overview of such outcomes and to judge them when possible in the light of similar changes documented elsewhere. To do this, the structure of the paper is as follows. First I give a brief overview for the loss of the dual cross-linguistically, illustrating the typical scenarios we find in their diachrony and directing the reader to relevant literature. This introduction is intended to frame the developments found in Otomi in a broader typological

context. Similarly, in order to set up the genetic background for Otomi, I also give an overview of the Oto-Pamean number system at the end of this section. In Section 2, I introduce Otomi as a linguistic family and I present a brief description of the morphology involved in the making of number oppositions. Then in Section 3, I introduce the number system of Old Otomi. This is intended as a landmark to understand the relevance for the different situations found in the modern languages, which I develop extensively in Section 4. Section 5 provides a discussion of the relevance of the different outcomes found in Otomi and Section 6 concludes.

1.1 The loss of duals: An overview

The dual is a resilient number value in many language families cross-linguistically, and from the observation of their behavior in a large database, Plank (2003) points out that the life cycles of duals are often remarkably long. However, in the evolution of number systems, duals may also be lost. For example, a dual value is found in the number system of some of the older Indo-European languages,⁵ but it gradually reduced its functionality or became lost altogether in most of the modern languages (see Corbett 2000:269 for references).⁶ In Greek, for example, the dual is relatively common in the Homeric texts (Monro 1891), but its occurrence becomes increasingly rare throughout the classical period, until the Koine Greek of the Hellenistic period, when it is lost altogether in both its nominal and pronominal domains (Forston 2004). Similarly, in Balto-Slavic, the dual has only survived in Upper and Lower Sorbian (Stone 1993a) and in Slovene (Priestly 1993). Nevertheless, the loss of the dual number is not in itself as interesting a phenomenon as the ways this loss may occur.

To briefly illustrate this point, I outline here two representative outcomes attested in Germanic involving pronominal number which are relevant for the purposes of this paper as they are also found in Otomi, together with further, hitherto unknown outcomes.⁷

In Gothic, the earliest attested Germanic language, we find dual forms in both verbs and pronouns, but only for the first and the second persons in the active voice (Wright 1910).⁸ This is illustrated in (1) and (2) when comparing the inflection of the dual and the plural, respectively (from the Wulfila Project hosted by the University of Antwerp, glossing mine).⁹

- (1) a. ...swaswe **wit** ain **siju**
 GOth even.as we.DU one.NOM.SG be.PRS.1DU
 ‘Even as we (two) are one.’ (John 17:22¹⁰)

- b. þau ainzu ik jah Barnabas ni hab-**os** waldufni du ni waurkjan?
 or only I and Barnabas NEG have-1DU power.ACC to NEG work-INF
 ‘Or I only and Barnabas, have not we power to forbear working?’
 (1 Corinthians 9:6¹¹)

(2) duhve weis jah Fareisai-eis fast-**am** filu,
 GOTH why we.PL and Pharisee-PL.NOM fast-1PL much
 iþ þai sponj-**os** þein-ai ni fast-and?
 but the.PL.M.NOM disciple(M)-PL.NOM 2POSS-NOM.PL NEG fast-3PL
 ‘Why do we and the Pharisees fast oft, but thy disciples fast not?’ (Matthew
 9:14¹²)

The situation found in Gothic, as illustrated in (1–2), is taken to be the situation in Proto-Germanic (Ringe 2006). In the rest of the Germanic languages, a dual value is only found in pronouns. In other words, there are dual forms for the first and second person, but the verb has lost the old dual inflection and now only agrees with dual pronouns in the plural. This is the situation found in Old English (Mitchell 1985), in Old Frisian (Bremmer 2009), in Old Norse (Gordon 1992[1957]), and in Old High German (see for example Howe, 1996: 244ff., for more details).

In contrast to the old languages, in the pronominal number of the modern languages, the dual has been lost. The loss of a dual value is preceded by a stage in which the meaning of the plural has also changed to accommodate references to two people.¹³ Typically, the loss of the dual also coincides with a period where the dual is perceived and used as a facultative value (for facultative number, see Corbett 2000: 42ff.). For example, in Old English we still observe that dual pronouns have a dual meaning referring to a group of two people (i.e. the speaker and hearer; the speaker and somebody else; or the hearer and somebody else), yet the old plural forms are already used in the system with the meaning proper of a simple singular-plural system as they can indicate more than one individual, i.e. two or more, and thus freely occur in situations where dual pronouns are used. This is illustrated in the passages in (3–4), where the reference is of only two individuals, but they are indistinctively referred to by dual and plural pronouns in different clauses. In (3), we have the dual (middle/reflexive) *unc* and the plural *we* for the first person, while in (4) we have the plural (accusative) *ge*, the dual (nominative) *gyt*, and (accusative) *inc* for the second person (Mitchell 1985: §258, glossing and translation mine).

- (3) ic 7 þæt cild ga-ð **unc** to gebidd-*enne*,
 OE I and the child go-PL we.DU.DAT/REFL to pray-INF
 7 **we** syððan cum-að sona eft to eow
 and we.PL.NOM after come-PL soon again to you.PL.DAT
 ‘The child and I will go to pray (Lit. pray ourselves (DU)) and we (PL) shall
 come later back to you (PL) soon.’

- (4) *ða cwæð Petrus: seg-e me,*
 OE then said.3SG Peter say-IMP me.DAT
beceap-od-e¹⁴ ge ðus micel land-es?...
 sell-PST-3SG you.PL.NOM thus much land-GEN.SG
hwi gewearð inc swa,
 why became.3SG you.DU.ACC so
þæt gyt dorst-on fandi-an God-es?
 that you.DU.NOM dared-PST.PL tempt-INF god-GEN.SG
 ‘Then Peter said: “Tell me, did you (PL) sell the land for so much?”... “How is
 it that you (DU) agreed to tempt god?”’

Once the old plural forms can do this, the old dual forms may become superfluous and may be lost, as they were in the Middle English period, where they survived in the conservative dialects of the South only until the beginning of the 13th century (Mossé 1952: 55).

In the development illustrated by the English pronouns, the dual forms are lost along with the dual value they convey. While this remains a common development, there is yet another possibility which involves keeping the old dual forms but reanalyzing them as exponents of the plural. This has also happened in the pronominal system of some Germanic languages, most notably in Icelandic.

In Icelandic, the old dual pronouns for the first and second person (*við* ‘we (DU)’ and *þið* ‘you (DU)’) have become plural forms, while the old plurals (*vér* ‘we (PL)’ and *þér* ‘you (PL)’) have become more restricted and are now only used as honorific plurals (Einarsson 1945: 68, 122 and Guðmundsson 1972 *apud* Corbett 2000: 269).

In a similar fashion, vestiges of the dual remain in Norwegian, which has a hybrid paradigm for the first person plural, where the subject form *vi* stems from the old dual *vit* (cognate of Icelandic *við*), while the object *oss* is the reflex of the original plural form. For the second person, similar mismatches have been reported in the Bavarian German dialectal area. In some varieties of Bavarian, the old second person dual pronoun *eß* has fully replaced the original plural *ihr*, while others have retained it (Kranzmayer 1954: 252ff., *apud* Howe 1993: 121ff.).¹⁵ Outside Germanic, a similar development is found in Slavic Kashubian.¹⁶

The situations in Icelandic, Norwegian, and Bavarian require the old dual forms to first be able to extend their semantic scope to include sporadic reference to more than two individuals, possibly three or four, depending on the situation, and once established as a possible reference for a small group of individuals, they can replace the old plural, which by then may have most likely become obsolete. This is what we find in the old dual of Old Norse which had a paucal extension, i.e. the inclusive plural construction *þit móðir mín* could both mean ‘you (SG) and my mother’ (i.e. dual reference ‘the two of you’) as well as ‘you (PL) and my mother’

(i.e. with plural reference ‘you, him, and my mother’) (Gordon 1992[1957]:313, see also Faarlund 2008). This development is exactly the opposite of what happens when plural forms replace the old dual forms, as in Middle English.

In this paper, I show that in the Otomi family a range of similar developments also occurred involving the dual: some of the varieties retained the old dual, while in others it was considerably weakened and eventually lost altogether. Furthermore, in the languages where we witness the loss of the dual, there are those like English (i.e., which have lost the dual markers along with the dual value) and others like Icelandic (i.e., which have lost the dual value but kept the dual markers having reanalyzed them as plurals).

The Otomi case is interesting for a theory of language change because it provides us with a rich scenario of possibilities, where we can observe: (i) how different typological outcomes can be brought about in a relatively short time, and (ii) how duals may survive in specific registers and grammatical contexts. In this connection, the developments in Otomi can be instructive both for the linguistic typologist as well as for the historical linguist. As an illustration of this, consider for example the case of Northern Otomi, which I will develop further in Section 4.4.

In Northern Otomi, the syntax has a singular-plural opposition as in English; dual forms have disappeared both in the pronominal system and in verb agreement. Nonetheless, we still observe sporadic uses of dual forms in verbs, but their usage is conditioned either by the pragmatics or by the specific requirements of the syntax of conjoining. In other words, the old dual value persists, but associated with a special register of adult males where it is perceived as a token of camaraderie. Similarly, the old dual forms have been retained in a conjoining construction dedicated to the expression of comitative semantics, where the idea of ‘I play with you’ is construed with the dual as “I play-DU you”. In such a construction, dual forms are obligatory, but more interestingly they have been reanalysed as exponents of a paucal value, i.e., a structure such as “I play-DU the men” is used to mean ‘I play with the men’ if the men involved are perceived as a handful of individuals, otherwise the plural is used. This resulted in a rather mixed number system which is discussed in Section 5.

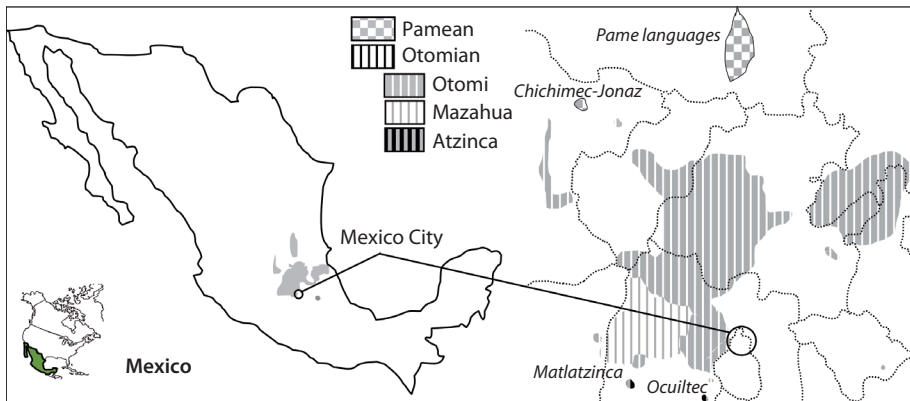
In the remainder of the paper, I present all these outcomes, but an introduction to Otomi in its broader genetic context is given first to understand the relevance of the internal changes. This is the purpose of the next section where I present the basic number system of Oto-Pamean.

1.2 The common number system in Oto-Pamean

Otomi is spoken in Mexico and it belongs to the Oto-Pamean branch of Oto-Manguan.¹⁷ Oto-Pamean itself is made up of two distant subgroups: Otomian

Table 1. The Oto-Pamean languages

Phylum	Branch	Group	Subgroup	Language level	
Oto-Manguean	Oto-Pamean	Otomian	Atzinca:	Matlatzinca	
				Ocuiltec	
			Otomi-Mazahua:	Mazahua	
				Otomi languages	
				Chichimeco-Jonaz	
			Pamean	Chichimec:	Northern Pame
				Pame:	Central Pame
					Southern Pame [†]
				Chinantecan	
				Popolocan-Zapotecan	
		Amuzgoan-Mixtecan			
		Tlapanecan-Manguean			



Map 1. Location of the Oto-Pamean branch of Oto-Manguean

and Pamean, hence the name. This is indicated in Table 1. The languages are geographically located in Map 1 for convenience, which is adapted from Soustelle (1937).

The nouns in Oto-Pamean languages inflect according to a singular-dual-plural number value opposition. This happens in all Oto-Pamean languages except in those of the Otomi-Mazahua subgroup, where the nominal number has been reduced to a simpler singular-plural opposition. This is illustrated in the Otomian group in (5) comparing Matlatzinca (Bartholomew 1966:40–41) and Tilapa Otomi. Number values are realized in these languages by means of number determiners preceding the noun. These determiners appear in boldface in (5).¹⁸

(5) Matlatzinca	SG	in šuʔyowi	‘coyote’
	PL	te šuʔyowi	‘two coyotes’
	DU	ne šuʔyowi	‘three or more coyotes’
Tilapa Otomi	SG	ar mintx’o	‘coyote’
	PL	yi mintx’o	‘coyotes’

The nominal number system is most elaborate in Pamean. This is shown in Table 2 in Chichimec with the inflection of the noun *kinhú* ‘(their) field’, where nominal number intersects with pronominal number of possessor (Angulo 1932: 163, acute indicates high tone; low tone is not represented).

In inalienable nouns in Chichimec like *kinhú* ‘field’ in Table 2, possession is realized by means of tone and stem alternations: the inflection of *kinhú* is given in shading. The singular form for the third person plural possessive is also used in a context where the entity referred to is not associated with any particular possessed entity, so that a form such as *kinhú* can both mean ‘field’ and ‘their field’. Nominal number suffixes (dual *-s* and plural *-r*) occupy the first position after the stem, followed by the number markers of possessor (dual *-(e)s* and plural *-(i)n*). In some of the forms, the occurrence of the dual marker *-s* can ambiguously realize a dual of the possessor (with a singular noun) or a dual of a noun (with a singular possessor), so that a form such as *kunús* can both mean ‘our (yours and mine) field’ or ‘my two fields’. The suffixes *-m^p* and *-hú* encode dual and plural exclusive for a possessor, respectively. The only odd forms in the paradigm are the dual forms *kunúvós* ‘our (his and mine) two fields’ and *kunhún* ‘our (theirs and mine)

Table 2. Inflection of the noun *kinhú* ‘field’ in Chichimec

		Number of possessor				
		SG		DU		PL
SG	1st	kunú	EXCL	kunú-m ^p	EXCL	kun-hú
			INCL	kunú-s	INCL	kunú-n
	2nd	kínu		kínu-s		kínu-n
	3rd	kinú		kinú-s		kinhú
DU	1st	kunú-s	EXCL	kunú-vó-s	EXCL	kun-hú-n
			INCL	kunú-s-és	INCL	kunú-s-in ⁴⁴
	2nd	kínu-s		kínu-s-es		kínu-s-in ⁴⁴
	3rd	kinú-s		kinú-s-és		kinhú-s
PL	1st	kunú-r	EXCL	kunú-r-úm ^p	EXCL	kunú-r-hú
			INCL	kunú-r-és	INCL	kunú-r-in
	2nd	kínu-r		kínu-r-es		kinú-r-in
	3rd	kinú-r		kinú-r-és		kinhú-r

Table 3. Personal pronouns in Chichimec

	SG		DU		PL
1st	i=kág	EXCL	i=kágu=m ^P	EXCL	i=kág=hũ
		INCL	i=kágo=s	INCL	i=kágu=n
2nd	i=xyék ^w		i=xyék'o=s		i=xyék'u=n
3rd	î=no		î=no=s		î=gó

Table 4. Anterior past tense of the verb *nú* 'see' in Chichimec

	SG		DU		PL
1st	tu=nú	EXCL	tu=nú=m ^P	EXCL	tu=nú=hũ
		INCL	tu=nú=s	INCL	tu=nú=n
2nd	ki=nú		ki=nú=s		ki=nú=n
3rd	u=nú		u=nú=s		u=nhú

two fields' where the suffixes of nominal number appear after the marker for the number of the possessor.¹⁹

Similarly, pronominal number in Chichimec is also realized in personal pronouns, as illustrated in Table 3 (Angulo 1932: 155),²⁰ and in verb agreement, as shown in Table 4 with the inflection of the anterior past tense of the verb *nú* 'see' (Angulo 1932: 163–164). The different proclitics in Table 4 are cumulative markers of tense/aspect/mood values and the person of the subject. Notice that the dual and the plural forms are based on the singular by adding the same number markers that encode number for the possessor in Table 2. The only exception to this pattern is the third person plural, which requires stem alternant, i.e. there are no such forms as **î-no-n* for 'they (three or more)' or **u-nú-n* for 'they (three or more) see', but *î-gó* and *u-nhú*, respectively.

Whereas a pronominal number system like the one in Chichimec, based on a singular-dual-plural opposition, remains the default system for Oto-Pamean languages, in some of the Otomi languages of the Otomian group this original number system has undergone a deep restructuring towards a simpler singular-plural number system. This happened for example in Northern Otomi and in the Southern languages, Tilapa Otomi and Ixtenco Otomi. Before studying these outcomes in further detail, in the next section, I first make a brief introduction to Otomi as a language family and to the ways number marking is realized.

2. An introduction to Otomi

In this section, I give a brief introduction to Otomi, aiming to present basic information about how pronominal number is encoded. This is essential to understand the number systems that are found in the different languages.

While the internal linguistic diversity of Otomi is uncontroversial (summarized in Lastra 2001), there is no established consensus as to which of the different linguistic subsystems identified so far as Otomi should be treated as independent languages and which as dialects of those languages. This is why Otomi has traditionally been talked about as forming a dialectal continuum rather than a family and why authors in Mexico avoid using terms such as “language” and “dialect” and instead use “variety” as a convenient cover term. Against the inertia of this tradition, the current official stance by the National Institute of Indigenous Languages (INALI) in CLIN (2008) is to treat Otomi innovatively as forming a “linguistic group” with nine different “linguistic varieties”. INALI’s linguistic varieties are treated as genuine languages for official purposes, that is, for educational, administrative, judicial, social, and informative purposes. This is also the stance I take in the present paper although with my own understanding of the current dialectal situation.

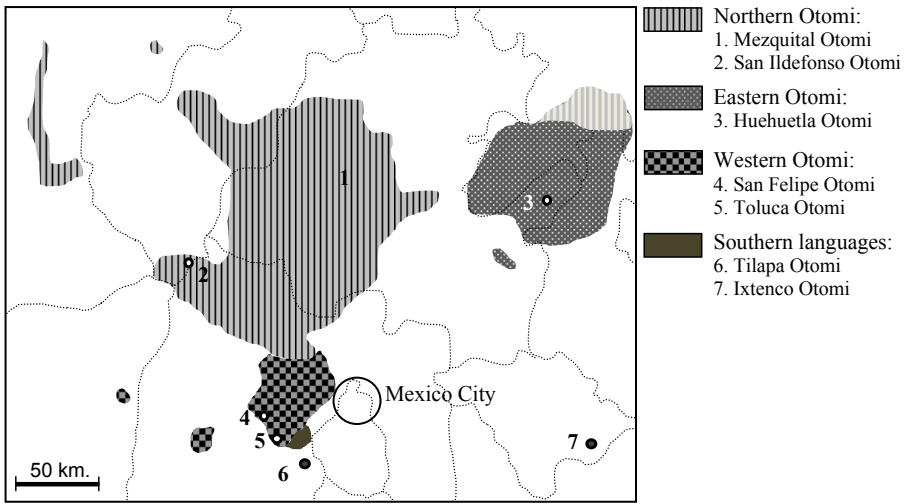
Otomi is interesting for a theory of language change because it has emerged as a linguistic family in recent times from having been a dialect continuum, historically. The first grammar of Otomi, *Arte de la Lengua Othomi*, was written by the Franciscan friar Pedro de Cárceres, who finished it in 1580, although the work was not published until 1907. Cárceres’ grammar is of a relatively early date, especially if we consider that the conquest by the Spaniards of the capital city of the Aztec empire, Tenochtitlan, happened in 1525, and that the first dictionary of Nahuatl by Alonso Molina was published in 1551, and then much enlarged for its second edition of 1571.

The language described in Cárceres is now known as ‘Old Otomi’ or ‘Colonial Otomi’ and it is taken to be the ancestor of all modern Otomi languages, that is, it represents the historical stage when Otomi was still a single language. Other historical documents survive of Old Otomi. The most representative ones are the short grammar titled *Arte Breve de la Lengua Otomi* by Alonso Urbano and a large trilingual dictionary (Spanish-Nahuatl-Otomi), whose authorship Acuña (1990) attributes to an anonymous Otomi amanuensis (both works are found in a manuscript dating from 1605). A third document is an unedited and unpublished anonymous vocabulary dating from 1640, and we also have the Huichapan Codex translated by Lawrence Ecker and nicely edited in Lastra and Bartholomew (2001). Later, two other Otomi grammars were written in the second half of the 18th century: one written by Neve y Molina and printed in 1767, and another anonymous one titled *Luces del Otomi*, later edited by Francisco Buelna in 1893.

Considering that the Viceroyalty of New Spain was well established by 1535, and existed until 1810, and taking the historical documents into consideration, we can roughly date the rise of Otomi as a linguistic family. The crucial thing in this development is that the two existing grammars from the 18th century are already based on an Otomi language which is substantially different from Old Otomi in that, for example, it has a much reduced inflectional system. The language described in these sources is most probably the earlier stage of the Northern language of today, itself a large diasystem including dialects such as Mezquital Otomi, the most widespread Otomi variety with the largest number of speakers. If we estimate an average of 20 years for each generation — most modern Otomi people still marry at a very early age, and there is no apparent reason to believe that this was different in the past — there are only nine generations between Cárceres' grammar and Neve y Molina's, and only six if one starts counting from the anonymous vocabulary of 1640.

All this suggests that the innovative Northern language emerged as a distinct linguistic system in a time span of only one and a half centuries, and it seems reasonable to imagine that a similar break up into different distinct dialectal areas must have also happened elsewhere by that time. One can only conjecture that the new social order imposed by the Spaniards after the conquest was partially responsible for the rapid rate at which this change happened (see for example Lastra 2010). The Otomi as a linguistic community, just like other Mesoamerican Indians elsewhere in New Spain, were confined geographically to their birthplaces, and it is sensible to conclude that this isolation brought about a lack of contact with other communities, which in turn maximized linguistic differentiation. However, the reason why the inflectional system of the Northern language became more innovative than in other areas remains an intriguing question that remains open to date.

For the purposes of the present paper, based on the data in Soustelle (1937), Lastra (2001, 2010), Andrews *et al.* (1983), and my own field observations, the different Otomi varieties spoken today could be seen as grouped in four geographical areas representing six different languages. Three such languages are in turn dialectal continua by themselves and include various linguistic varieties as dialects. If Mexico City were taken as a geographical reference point, we could conveniently identify these four areas by using the four cardinal points. In Map 2, adapted from Soustelle (1937), I indicate the four dialectal areas together with the linguistic varieties from which the representative examples involving pronominal number come, as given in the paper.



Map 2. Location of the Otomi languages and varieties exemplified

2.1 Number marking in Otomi

Otomi languages are head-marking (i.e. there is no case marking on nouns). They have accusative alignment in monotransitive verbs, and an agentive/patientive split in intransitive verbs. Otomi verbs inflect for tense, aspect, and mood (TAM) by means of a set of inflectional formatives.²¹ Many such formatives also cross-reference person of the subject in a cumulative fashion, and for this reason, they are treated here as “P&TAM” formatives (P stands for person of subject).²² Examples are given in (6) for the first and second person in San Ildefonso Tultepec (SIT) Otomi, a variety of Northern Otomi.

- (6) a. $d\acute{a}=n\check{u}\text{'-}i$
 SIT 1.PST.R=see-2OBJ-F
 ‘I saw you.’
- b. $g\acute{a}=n\check{u}\text{'-}g-i$
 2.PST.R=see-1OBJ-F
 ‘You (SG) saw me.’

The examples in (6) illustrate the use of two P&TAM formatives for the inflection of the transitive verb *nú* ‘see’ and they realize grammatical person, past tense, and realis mood. It may be further noted in (6) that objects are cross-referenced by means of suffixes. Phonologically, all P&TAM formatives are clitics in that they are (most commonly) hosted on the verb stem or on a previous functional word.

The P&TAM formatives in (6) make reference to a singular subject by default when occurring alone. A plural or a dual subject is encoded by means of an

additional number enclitic. This is illustrated in (7) with the plural marker *hú* for the first and second persons. Notice that a third person object is encoded by means of the bare stem or a zero morpheme.

- (7) a. *dá=nǔ=hú*
 SIT 1.PST.R=see[3OBJ]=PL
 ‘We (INCL) saw him/her/it/them.’
 b. *gá=nǔ=hú*
 2.PST.R=see[3OBJ]=PL
 ‘You (PL) saw him/her/it/them.’

The first person plural also inflects for clusivity. Example (8) shows the dedicated plural marker for the first person exclusive. This makes the plural marker *hú* in forms such as (7a) express an inherent inclusive sense when it is associated with the first person.

- (8) *dá=nǔ=hé*
 SIT 1.PST.R=see[3OBJ]=PL.EXCL
 ‘We (EXCL) saw him/her/it/them.’

An example with the dual exclusive is given in (9) from Toluca (TOL) Otomi, a variety of Western Otomi (slightly adapted from Lastra 1992: 328). However, because the dual is a complex number value in Otomi, at this point in the exposition I will illustrate the remaining marking possibilities of pronominal number with examples in the plural only.

- (9) *dománninibe y_Λ pháni...*
 TOL *dó=má='winí='bé* *y_i pháni*
 1.PST.R=PURP.MOT=feed[3OBJ]=DU.EXCL PL horse
 ‘We went to feed the horses...’ (Txt)

The plural of an object is also encoded with the same number enclitics we have in (7–8). This is shown in (10). Notice that for this situation, a doubling of object marking is required at least for SIT Otomi.

- (10) a. *dá=nǔ-'-a='í=hú*
 SIT 1.PST.R=see-2OBJ-B=2OBJ=PL
 ‘I saw you (PL).’
 b. *bi=nǔ-g-a=gí=hú/hé*
 3.PST.R=see-1OBJ-B=1OBJ=PL/PL.EXCL
 ‘He/she/they saw us (INCL/EXCL).’

Additionally, the verb can only receive one number marker, so when both subject and object are plural, the plural of the subject is encoded by a free pronoun,

while the verb indexes the number of the object. This is shown in (11) (Palancar 2009: 261).

- (11) nú='i=hú gá=zěngwa=gí=hé
 SIT INTRO=2SG=PL 2.PST.R=visit=1OBJ=PL.EXCL
 'You (PL) visited us.'

Sometimes there can be ambiguity regarding the argument the number enclitic references. In (12), from Toluca Otomi, the clitic *hú* could either refer to the subject or to the object (slightly adapted from Lastra 1992: 34).

- (12) *bi-ñáš-ki-hí*
 TOL bi='ñáš-ki=hú
 3.PST.R=SS/shear.AS-1OBJ=PL
 'They cut my hair.'²³ (Lit. 'They sheared me.')
- 'He cut our (INCL) hair.' (Lit. 'He sheared us (INCL).')

As for personal pronouns, in Otomi, like in Chichimec in Table 2 above, the same number markers are also used for the building of plural forms. There are no independent personal pronouns; pronominal forms are always enclitics hosted on a set of particles such as, for example, the presentational or definite particle *nú* in Table 5. There are various alternative forms for the first and second person in SIT Otomi (Palancar 2009: 342).

Table 5. Personal pronouns in SIT Otomi (Northern Otomi)

	SG	PL	
		INCL	EXCL
1st	nú=gá	nú=khú	nú=khé
	nú=gí	nú=gí=khú	nú=gí=khé
		nú=gí=hú	nú=gí=hé
2nd	nú='í	nú='í=hú	
	nú='í=gé		
	nú=gé		

In Table 5, the forms of the plural are formed on the singular by means of similar number markers, except with the markers =*khú* and =*khé*, which have morphologized as cumulative exponents of first person and plural, inclusive and exclusive, respectively.²⁴

Similarly, as in Chichimec, number markers in Otomi are also added to nouns to realize number values of a possessor. This is illustrated in (13), again with a plural possessor.

- (17) a. nũ 'see'
 HO gë 'roar'
 ho 'kill'
 b. hwé+^hk- 'love'
 pú+^ht- 'peck'
 bɛn+t' 'grasp'
 yä+^hp- 'accuse'

In Huehuetla Otomi, each plural morpheme has two allomorphs; indicated in (18).

- (18) {hú} h_ú, ú 'PL'
 {hé} hé, é 'PL.EXCL'

The allomorphs *h_ú* and *hé* are the default realization of {hú} and {hé}, because they have a wider distribution (i.e., apart from verbs, they are also used in pronouns and in nouns to mark possessor). With verbs they are selected by lexemes with monosyllabic stems as in (19a), as well as with those that have the stem formatives in (19b) and (19c). In the examples, I illustrate it with instances of {hú}, but the same would apply to {hé}. In (19c), the stem formative undergoes a slight phonological adjustment when hosting the number enclitic.

- (19) a. nũ 'see' > nũ=h_ú
 HO b. hwé+^hk- 'love' > hwé+^hk=h_ú
 pú+^ht- 'peck' > pú+^ht=h_ú
 yä+^hp- 'accuse' > yä+^hp=h_ú
 c. bɛn+t' 'grasp' > bɛn+t=h_ú
 ɛ+ng- 'push' > ɛ+nk=h_ú

The allomorphs *h_ú* and *hé* are also used with stems carrying a nasal formative, but here all nasals are reduced to the clusters /mb/ and /nd/ (where the nasal formative is followed by a homorganic excrescent stop /b/ and /d/). The underlying sequence undergoes subsequent metathesis, where <...> stands for the specific location of the consonantal onset of *h_ú* after the metathesis.

- (20) ho+n- 'seek' > ho+mb=h_ú > hom<h>b_ú
 HO tsǎ+mm- 'retain' > tsǎ+mb=h_ú > tsǎm<h>b_ú
 ǎ+hm- 'put someone to sleep' > ǎ+mb=h_ú > ǎm<h>b_ú
 tó+'m- 'wait for' > tó+mb=h_ú > tó<h>b_ú
 'a+nn- 'ask someone' > 'a+nd=h_ú > 'an<h>d_ú
 tó+hn- 'push' > tó+nd=h_ú > tón<h>d_ú

On the other hand, the allomorphs *ú/é* occur before stem formatives with sibilants:

(21)	kǎ+x-	‘bend’	>	kǎ+x=ú
HO	tsĩ+ts’-	‘retain’	>	tsĩ+x=ú
	ě+’ts’-	‘propose’	>	ě+s=ú

But *ú/é* are also used with a special set of verbs, like the ones in (22), which require a special stem to generate the form of the plural. Notice that the resulting form does not undergo subsequent metathesis as in (20) above, although in principle it could.²⁷

(22)	tǎh-	/	tǎmb-	‘win’	>	tǎmb=ú
HO	pǎ+d-	/	pǎmb-	‘know’	>	pǎmb=ú
	pé+g-	/	pémb-	‘put on the ground’	>	pémb=ú
	zó+f-	/	zómb-	‘talk to someone’	>	zómb=ú
	fě+i	/	fěmb-	‘lash’	>	fěmb=ú

In this section, I have introduced the basics about the morphosyntactic and morphological distribution of number markers in Otomi. I have purposely left examples with dual markers out of the discussion because they represent a more complex case, as will soon become clear in the following sections, where I introduce the different number systems we find, one by one, starting with Old Otomi.

3. The basic number system in Old Otomi

In the previous section, I have presented the basics of pronominal number encoding in Otomi; we have seen that number marking encodes number of subject in pronominal forms and in verb agreement as well as number of possessor with nouns. In this section, I introduce the historical system found in Old Otomi. I do this to provide a perspective that serves to understand and evaluate the deviations that have occurred in the daughter languages.

While Otomi-Mazahua had already lost the dual value in nominal number from Oto-Pamean (this was seen, for example, in 9 above), it preserved the dual in pronominal number. The pattern survived into Old Otomi, as illustrated in Table 6 by the subparadigm of the present indicative of the verb *xohnābate* /ʃ̄ŋ̄apāɬe/ ‘teach’ as extracted from Cárceres (1907[1580]: 65, 69).

The marker *wi* /wí/ realizes the dual, while *hæ* /hí/ encodes the plural. Like in the modern languages, when these markers are used with a first person subject, they also convey an inclusive value. This is by virtue of the existence of dedicated morphemes to realize an exclusive value, like dual exclusive *be* /bé/ and plural exclusive *he* /hé/.

Table 6. Present indicative of the verb *xohnābate* ‘teach’ in Old Otomi

	SG		DU		PL
1st	tānāxohnābāte	EXCL	tānāxohnābātebe	EXCL	tānāxohnābātehe
		INCL	tānāxohnābātewi	INCL	tānāxohnābātehoē
2nd	cānāxohnābāte		cānāxohnābātewi		cānāxohnābātehoē
3rd	nāxohnābāte		nāxohnābāte(wi)		nāxohnābāte(hoē)

While number marking was obligatory for the first and second persons, it was already optional for the third person. This facultative character was indicated in Table 6 by means of the parentheses and represents a phenomenon that did not go unnoticed by Cárceres, as he himself points out (1907[1580]:65, free translation mine) in the following:

“[L]a 3ª persona del plural es semejante a la del singular, en todos uerbos y por tanto exadjuntis, se a de conocer quando es plural o singular, aunque ay esta diferencia y es que quando todos los que dan supuesto al uerbo se expresa con conjuncion copulatiua que al uerbo preceda entonces se rrequiere boz del plural.”

‘The 3rd person plural is similar to the singular for all verbs. Because of this, the context dictates when it is plural and when singular. But there is one exception, and this happens when the arguments of the verb are expressed with a conjunction preceding the verb. In such a case, plural marking is required.’

In this way, plural marking is not used in situations like in example (23) because the plural subject does not appear in any conjoined NP. Here and elsewhere I indicate the absence of a number marker by an underscore. The example in Old Otomi (OO) is taken from Cárceres (1907[1580]: 44), but as in other instances, the conversion into modern Otomi orthography, the glossing, and the translation are mine.

- (23) *nūhnāpicha nūye apostol nūbuæmibuæy nūguayximāyāmāhoy*
 OO *nūyahīntācha, tīchaya, nūkochāyāy*
 nūh=na Ø pi=k^ha__ nū=y=e apostol
 INTRO=DEM.SG REL 3.PST.R=do INTRO=DEM.PL=PL apostle
 nū=bu mí=’bu-i__ nū=gwa ^hxi-mayamahay
 INTRO=when IMPF=[3.PRS]be-F INTRO=here HON-earth
 nū=’ya hin=ta=kha ti=kha=ya nū=k²=o khāyāi
 INTRO=now NEG=3.IRR=exist 3.IRR=do[3OBJ]=ENCL INTRO=DEM.SG=HON person
 ‘That which the apostles did when they were here on earth, now the people cannot do it.’

In contrast, number agreement with a third person dual or plural subject is used when a conjoined NP subject precedes the verb. Conjoined NPs are indicated

by brackets, while the coordinating conjunction is underlined to facilitate recognition. Plural number is shown in (24); dual in (25) (Cárceres 1907[1580]:65).²⁸

(24) *nũko San Pedro xĩ mēko San Pablo xĩ mēko San Andres, pnixohmābātehø*
 OO [nũ=k'o San_Pedro xĩ='ne=k'o San_Pablo
 INTRO=DEM.SG=HON Saint_Peter also=and=DEM.SG=HON Saint_Paul
xĩ='ne=k'o San_Andres] p<n>i=xāh nabate=**hu**
 also=and=DEM.SG=HON Saint_Andrew <INTR>3.PST.R=teach=PL
 'Saint Peter and Saint Paul and Saint Andrew preached.'

(25) *nũko Santa Maria xĩ nēkanimācābāṭṭi pinbuø biban Belen*
 OO [nũ=k'o Santa_Maria xĩ='ne=k'a ní makā-bātsi]
 INTRO=DEM.SG=HON Saint_Mary also=and=DEM.SG.III DEM.SG.3POSS ?-child
 pi-n='būbi=**b**=an Belen
 3.PST.R-INTR=live.AS=DU=LOC.P Bethlehem
 'Saint Mary and her child lived in Bethlehem.'

The pronominal system of Old Otomi is given in Table 7, where all pronominal forms are based on the presentational particle *nũ* (*nṹ*) (Cárceres 1907[1580]:50).²⁹

Table 7. Personal pronouns in Old Otomi

	SG		DU		PL			
1st	<i>nũguĩ</i>	nũ=gi	EXCL	<i>nũgābe</i>	nũ=ga='bé	EXCL	<i>nũgāhe</i>	nũ=ga=hé
	<i>nũcāgo</i>	nũ=ka=gá	INCL	<i>nũgāui</i>	nũ=ga=wí	INCL	<i>nũgāhø</i>	nũ=ga=hú
2nd	<i>nũkĩ</i>	nũ=k'í		<i>nũkāvi</i>	nũ=k'a=wí		<i>nũkāhø</i>	nũ=k'a=hú
	<i>nũkāgue</i>	nũ=k'a=gé						
3rd	<i>nũnø</i>	nũ=nũ					<i>nũyø</i>	nũ=yú

In Table 7, the forms for the third person are built on two demonstratives, singular *nø* (*nṹ*) 'that' and *yø* (*yú*) 'those'. The chart illustrates the lack of regular number for the third person. Interestingly, Western Otomi, which as I show has preserved to date the historical situation of Old Otomi, also has dual and plural pronominal forms for the third person, while these are not based on a demonstrative. This situation suggests that the same happened in Old Otomi, but was not recorded by Cárceres.

4. The number system in the modern Otomi languages

The historical number system of Old Otomi has undergone a considerable restructuring in most of the modern languages. In this section, I introduce the different

outcomes one by one. We can start with Western Otomi, which is the most conservative in this regard, having preserved the historical situation almost intact.

4.1 The number system in Western Otomi

Western Otomi is a dialectal continuum of very different varieties, which include Toluca Otomi, San Felipe Otomi, Temoaya Otomi, San Pedro Otomi, and Jiquipilco Otomi. In the pronominal system of Toluca Otomi, the variety that I take here to be representative of the language, there is regular dual and plural marking for all persons. This is shown in Table 8 (Lastra 1992: 19). Notice that the dual and plural forms for the third person are based on a genuine third person pronoun (*gégé*) and not on a demonstrative pronoun like the attested Old Otomi forms in Table 7.³⁰

Table 8. Personal pronouns in Toluca Otomi (Western Otomi)

	SG		DU		PL	
1st	<i>nugó</i>	nu=gó	EXCL <i>nugó-bé</i>	nu=gó= 'bé	EXCL <i>nugó-hé</i>	nu=gó= 'bé
			INCL <i>nugó-wí</i>	nu=gó= wí	INCL <i>nugó-hí</i>	nu=gó= hú
2nd	<i>nukʔígé</i>	nu=k'í=ge	<i>nukʔígé-wí</i>	nu=k'í=ge= wí	<i>nukʔígé-hí</i>	nu=k'í=ge= hú
3rd	<i>gégé</i>	<i>gégé</i>	<i>nugégé-wí</i>	nu=gégé= wí	<i>nugégé-hí</i>	nu=gégé= hú

Interestingly there is more to the third person dual and plural forms in Table 8 than meets the eye. In reality, such pronominal forms can only be used when they refer to human beings. In other cases, the number enclitics cannot be used. The excellent textual material in Lastra (1992) provides us with plenty of information about this semantic distribution.

In Western Otomi, number agreement with a third person dual or plural is conditioned by the semantics. The conditioning most probably dates back to Old Otomi. In the animacy hierarchy, only personal pronouns and human subjects can trigger number agreement on the verb, while the rest of the nouns in the hierarchy do not. This is illustrated in (26) with two instances: one is an inanimate subject, the other is animate but not human (Lastra 1992: 347, morphemic analysis, glosses, and English translation are mine).

- (26) *y_A cithǎ níhǐ bí ni ʔa híçj este mǎhke šo indici y_A ngú*
 TOL *y_U tsi-thǎ 'níhǐ 'b_U(=***h_U**)=ni 'a hítsǐ,*
 PL DIM-maize also [3.PRS]be.located.AS(=*PL)=there LOC.P loft
 'The corn-cobs are also there on the loft.'

*este_mǎ*ke xo i=ndutsi(=***h_U**) y_i ngú*
 even so 3.PRS=feed.on(=*PL) PL rat
 'But even so, the rats eat them.'

With animals, number agreement is only possible with reciprocal verbs, as shown in (27), but here again it is optional.

- (27) *ši toʔitébíkʔí ma rʌ kʔí izó ʔa hʔí, porke dantʔhí*
 TOL xi=to=’üté-bí=k’ú ma rí k’ú i=’ó(=*hú)
 also=1.PST.R=make-3DAT=3PL other INDEF.SG REL 3.PRS=be.down(=*PL)
 ’a hǎi,
 LOC.P ground
 ‘And I also made those other (nests) that are below on the ground.’
 porke da=n-tü=**hú**/n-tühn-i__
 because 3.PRS.IRR=MIDD-fight.AS=PL/MIDD-fight-F
 ‘Because they (the hens) quarrel (with each other).’ (Lastra 1992: 344)

With humans, agreement is further restricted to subjects with specific referents in the discourse. This is illustrated in (28), from San Felipe (SF) Otomi, where the plural is used when the speaker is taking about a specific group of people (i.e. her relatives), notice the additional plural as possessor; and in (29), where the referent of the noun *khä’ni* ‘indigenous people’ is also concrete (notice for instance the use of the deictic demonstrative); and similarly in (30), where the speaker refers to her parents. Further on in this example, the speaker switches to a generic reference, and consequently stops using plural agreement on the verb.

- (28) *nin, te ci di ʔbihtihí nír nguhi, te ʔi pæhmi*
 SF nin, te tsu di=’buh=ti=**hu** nú=r ngu=**hu**,
 no what little DUR=[3.PRS.IRR]be.AS=DEL=PL DEM.SG.3POSS=SG house=PL
 ‘No, they are hardly ever at their home.’

te i=péh=**mu**
 what 3.PRS.R=work.AS=PL
 ‘They do all kinds of work.’ (Andrews 1993: 83)

- (29) *yi khä’ni mí ʔatihí mbaškhwa cada ocho dia, ští dʔ thi ir marte, geh ni mí ʔatihí mbaškhwa ni*
 SF yu khä’ni mí=’oti=**hu** mbaxkhwa cada ocho dia,
 PL person IMPF=[3.PRS.R]do[3.OBJ]=PL feast every eight day
 ‘These people used to have fiestas once a week.’

x=tí=dət=**hu** ur marte,
 PRF=3.PRS.IRR=SS/find[3OBJ]=PL SG Tuesday
 ‘Whenever a Tuesday came.’ (Lit. ‘(When) they’d find a Tuesday.’)

geh=nü mí=’oti=**hu** mbaxkhwa=nú
 IDENT=DEM.SG IMPF=[3.PRS.R]do[3OBJ]=PL feast=there
 ‘They were then throwing a party.’ (Andrews 1993: 74)

- (30) *asta bizáhí yΛ naštaná, bizáhí kwalkiera yΛ pásí yí šo minuhí. Ta šo dúnthi yΛ khání biduṇá*
 TOL *asta bi=zǎ=hú yí naxtä='ní,*
 even 3.PST.R=SS/eat=PL PL raw.prickly.pear=CIT
 'They (my parents) even ate raw prickly pears.'
bi=zǎ=hú kwalkiera yí páxí yu xo mí=nu=hú.
 3.PST.R=SS/eat=PL any PL grass REL well 3.IMP=see[3OBJ]=PL
 'They ate any herb that they well saw.'
ta xo dúnthi yí khání bi=dü__='ní
 even well many PL person 3.PST.R=SS/die=CIT
 'Many people indeed died, they say.' (Lastra 1992: 360)

When the plural subject is non-specific, that is, when it does not refer to specific persons, the verb does not show number agreement. As I have pointed out, this may already be seen in the last clause in (30), where we find the noun for people used as a generic in the last clause. The point is further illustrated in (31) where the speaker is talking about women in a rather general way. A particularly interesting example is (32), where the speaker portrays her five brothers as an undifferentiated collective. Had she used the plural marker *hú*, she would have implied that all the brothers hit her together as a group, when in reality, regardless of how unpleasant the experience must still have been, she implies that she was hit by one or two, now and then.³¹

- (31) *pe maqhamá bastante misufri yΛ bēhñá*
 TOL *pe mākhamá bastante mí=sufri__ yí 'bēhñí*
 but time.ago rather 3.IMP=suffer PL woman
 'But a time ago women suffered a lot.' (Lastra 1992: 361)
- (32) *...máke mižngí yí mǎzídá...*
 TOL *mǎ^hke mí='ǔn-gí__ yu mā 'ídá*
 although 3.IMP=hit-1OBJ PL 1POSS brother.of.woman
 'Although my brothers hit me.' (Lastra 1992: 332)

On the other hand, as we saw above in (24–25) for Old Otomi, in Western Otomi number agreement is also required with a conjoined NP subject. This is shown in (33).

- (33) *kar David kor Lasaro ci žñāhəwi*
 SF [ka=r David ko=r Lasaro] tsi-'ñohə*(=wí)
 DEF.SG=SG David with=SG Lazaro [3.PRS]DIM-man=DU
 'David and Lazarus are little boys.' (Andrews 1993: 37)

The conjoining structure in (33) represents an instance of the widely attested comitative coordination (Schwartz 1985, Stassen 2000) and thus remains uncontroversial from a cross-linguistic point of view.³² There is, however, another con-junctional construction found in all Otomi languages which is treated by Palancar (2012b) as an instance of ‘split conjunction’. Such a construction is typologically much rarer and more interesting for our present purposes, as it displays a number of intriguing restrictions regarding number agreement. Because of this, it deserves to be studied separately in the next section.

4.2 Number and split conjunction

All Otomi languages have a special conjoining construction dedicated to the expression of comitative and reciprocal semantics involving human beings. The construction in question is syntactically interesting because of the three following features:

- i. The conjoined subject NP is always a split one, that is, the two conjuncts appear in different positions in the clause: one is a topical subject and occurs in the typical position of a subject (whether the language is SVO or VOS); the other conjunct appears in the typical position of the object.
- ii. The construction is restricted to intransitive verbs.
- iii. The verb must show resolved number agreement: dual or plural.

The construction is illustrated in (34) in Western Otomi in the variety of Toluca Otomi. In the examples, I further show that number agreement is obligatory.

- (34) a. ...gégé šo biʒiniwí kʒʌ rʌ prima khwé
 TOL [gégé] xo bi=ʒini*(=wí) [kʒi rí prima_khwé]
 s/he so 3.PST.R=PLAY=DU DEF.SG SG female.cousin
 ‘She was playing with her cousin.’ (Lastra 1992: 328)
- b. domá a múlínú domébé na maámíga
 TOL [pro] dó=má a múlínú
 1.PST.R=SS/go LOC.P mill
 dó=mé=ʒé/*ma___ [ʔna mä ámíga]
 1.PST.R=SS/go.DU=DU.EXCL//SS/go one 1POSS friend
 ‘I went to the mill with a friend.’ (*ibid.* 349)
 (Lit. ‘I went to the mill, I went with a friend.’)

Example (34a) involves two singular NPs, and the verb has resolved agreement in the dual number. Notice that the two conjuncts appear in two different positions. Example (30b) is similar to (34a), but here we have an instance of the

clause combining construction involving the juxtaposed clauses we saw in (16) in Section 2.1, above. This construction requires a coreferential subject, which in this context is the first person. The subject of the second clause *dó mĕ'bé 'na mä ámíga* 'I went with a friend' is obligatorily elided; that is, it cannot have an explicit subject personal pronoun as in (34a), which in this case would have been *nugó* 'I'. Had the speaker wanted to add such a pronoun, she would have had to place it before the first clause *do má a múlínú* 'I went to the mill' attending to the syntactic restrictions of the construction. Example (34b) illustrates that the language is pro-drop (the elision of a pronoun is indicated by *pro*).

This conjoining construction is old as it is also found in Old Otomi and it is even shared by Mazahua (Stewart 1966: 76, 108) which split from Otomi around 500 AD. The only two examples to be found in Cárceres (1907[1580]) appear in (35).

- (35) a. *tāmévike Pedro*
 oo [pro] ta=mĕ=wi[=k'e Pedro]
 3.PRS.IRR=SS/go.DU=DU=DEM.SG Peter
 'He may go with Peter.' (Cárceres 1907[1580]: 94)
- b. *cobānyovike Pedro*
 [pro] ko=ba=n-'yo=wi[=k'e Pedro]
 2.PRS.IRR=TRANSLOC=INFL.INTR-walk=DU=DEM.SG Peter
 'You (SG) will go away with Pedro.' (*ibid.* 116)

As it will become clear in the following sections, this conjoining construction has played a very significant role in the preservation of the old number system in the modern languages, which have elsewhere innovated considerably compared with the original situation. The result is an interesting mixture of number systems occurring within one language, which is rarely seen typologically. I will introduce the different situations one by one, moving anticlockwise around the dialectal compass for convenience, beginning with the system in the East, then the North, and finally the South.

4.3 The number system in Eastern Otomi

In the previous section, we have seen how Western Otomi preserves the number system of old which consists of three values: singular, dual, and plural. Number agreement is used with all persons in conjoined structures, whether of the coordinating type in (33) or of the split type in (34). Number agreement is a morpho-syntactic requirement with the first and second person, but with the third person it has become a morphosemantic feature because it is restricted to subjects that are human and specific.

The first deviation from this original situation is the total loss of number in the third person. This happens in all other Otomi languages, but we will concentrate first on the situation found in Eastern Otomi, which bears the closest structural similarity to that of the West, seen before.

Eastern Otomi, also known as Highlands Otomi, is also a dialectal continuum which includes the varieties spoken in Huehuetla (San Gregorio and San Antonio el Grande), Tenango de Doria, Texcatepec, and other peripheral varieties such as Tulancingo Otomi and San Pablito Otomi. The personal pronouns of Huehuetla Otomi, the variety of Eastern Otomi I take to be representative of the language, are given in Table 9. Here we have a contrast of neutral vs. emphatic forms for the first person. For the second person, the contrast is formal vs. intimate. As for number, the third person pronouns are different in that there is no dual pronoun, and the plural is not regularly built on the singular form, (Voigtlander and Echegoyen 1985[1979]: 271).

Table 9. Personal pronouns in Huehuetla Otomi (Eastern Otomi)

		SG			DU			PL	
1st	NEUTRAL	nú=gí	EXCL	NEUTRAL	nú=g='bé	EXCL	NEUTRAL	nú=k=hé	
	EMPH.	nú=gá		EMPH.	nú=ga='bé		EMPH.	nú=ga=hé	
			INCL	NEUTRAL	nú=g=wí	INCL	NEUTRAL	nú=k=hú	
				EMPH.	nú=ga=wí		EMPH.	nú=ga=hú	
2nd	FORMAL	nú='é		FORMAL	nú='e=wí		FORMAL	nú='e=hú	
	INTIMATE	nú='í		INTIMATE	nú='a=wí		INTIMATE	nú='a=hú	
3rd		nú='ã			nú='ú				

In Eastern Otomi, the number system has a healthy contrast of dual and plural values for the first and second person, but verb agreement with a third person subject is no longer sensitive to number, at least in the default situation. This situation is shown both in (36a) and in the coordinating construction in (36b) with the conjunction *'ne* 'and'. These examples illustrate that the verb cannot receive number agreement with a plural or dual third person subject, regardless of whether the subject referent is human, specific, or whether it appears in a coordinating context. In contrast, the conversation in (37) illustrates that number agreement is obligatory with a first and second person dual or plural (Voigtlander and Echegoyen 1985[1979]).

- (36) a. 'ne ga'tho yo xitsu, ga'tho yo t'uhni bi=n-'unb-i ___/*n-'unb-a=hú
 HO and all PL woman all PL child 3.PST=MIDD-suffer-F/MIDD-suffer-
 B=PL

nange ra tsui Ø bi=n=kha
 with SG fight REL 3.PST=INFL=exist
 'All the women and children suffered because of the war that was'
 (p. 301)

- b. bi=mba__/*mbe=wi [ra Xuwa 'ne ra Zigo]
 3.PST=ss/go//ss/go.DU=DU SG John and SG Francisco
 'John and Francisco went (there).' (p. 56)

- (37) -ha gí=xu*(=wi)?
 HO INT 2.PRS=cut.wood=DU
 'Are you (DU) cutting wood?'
 -dí=xu*(='be) thã?
 1.PRS=cut.wood=DU.EXCL indeed
 'Indeed we (DU) are cutting wood'
 -'darbu da=n=xu*(='hu)=bu
 then 1.PRS.IRR=INFL.INTR=cut.wood=PL=so
 'Then let's cut wood together.' (p. 149)

The new situation is summarized in Table 10 with the subparadigm of the present realis of the verb 'yo 'walk', as extracted from Voigtlander and Echevoyen (1985[1979]:98, 147ff.).

Table 10. Present realis of the verb 'yo 'walk' in Huehuetla Otomi (Eastern Otomi)

	SG		DU		PL
1st	dí='yo	EXCL	dí='yo='bé	EXCL	dí='yo=hé
		INCL	dí='yo=wí	INCL	dí='yo=hú
2nd	gí='yo		gí='yo=wí		gí='yo=hú
3rd	i='yo				

Interestingly, however, number marking for the third person has not been fully lost. It survives in the split conjunction construction, where the use of number agreement is obligatory, as illustrated by the ungrammaticality of its elision in (38), where example (a) is plural and (b) dual. Notice that Eastern Otomi is a VOS language, and that the conjuncts in (38) do not represent a contiguous conjoined NP by juxtaposition, but two different NPs in different syntactic positions in the structure. The topical conjunct appears in the position of the subject, hence the translation (Voigtlander and Echevoyen 1985[1979]).

- (38) a. bi=n=tsi-hme*(=**hu**) [y_u xambate][='u y_u xadi]
 HO 3.PST=INFL/INTR=ingest-tortilla=PL PL teacher=DEF.PL PL student
 'The students ate with the teachers.' (p. 307)
- b. 'b_uh=**mi**/'*b_u-i [rá ndombe] [n_u ra t'uhni]
 [3.PRS]live.AS=DU/live-F SG.3POSS grandmother DEF.SG SG child
 'That child lives with his grandmother.' (p. 307)

All the other languages that have lost number in the third person behave like the Eastern Otomi in this respect. The maintenance of number agreement with a third person in the split conjunction construction serves two purposes. One is to indicate resolved agreement with the two NP conjuncts involved, so that the action is seen being carried out by both entities acting as a group. The other purpose is to serve as a very convenient morphosyntactic flag for the hearer to process that the speaker is in fact using this specific conjoining structure. In other words, when a verb has a third person plural subject and appears with number markers, it is irrevocably an instance of split conjunction. If it has no number agreement, it is not. This is shown by the contrast between (39–40). In example (40) — an instance of split conjunction — only the topical subject is present in the surface structure; the other conjunct expressing the comitative participant is inferred from the context. This is indicated by *pro* (Voigtlander and Echegoyen 2007).

- (39) xóngu yo khä'i i=yä__
 HO many PL people 3.PRS=talk
 'Many people are talking.' (p. 102)
- (40) [xóngu yo khä'i] i=yä=**hu** [pro]
 HO many PL people 3.PRS=talk=PL
 'Many people are talking with him/her/them.' (p. 102)

In both the Western and the Eastern languages, the dual and plural are healthy morphosyntactic values. In the rest of the languages, the dual is a disappearing value, but in this respect, its treatment has been very different in the North and the South. Let us first see what the situation is in Northern Otomi.

4.4 The number system in Northern Otomi

Northern Otomi is a large dialectal continuum of closely related varieties including San Ildefonso Tultepec (SIT) Otomi, Santiago Mexquititlán Otomi, Mezquital Otomi, and other small varieties such as Chantepec Otomi, Tolimán Otomi, Acambay Otomi, and possibly the isolated variety spoken to the east in Ixhuatlán de Madero.

4.4.1 *The dual in the split conjunction construction.*

In Northern Otomi, the old dual markers survive in the split conjunction construction, where their occurrence is obligatory. Example (41) is from the variety of SIT Otomi, which shows an SVO word order.

- (41) [nú='í=gé] gá=n-thé*(=wí)[=r Zü'wē]
 SIT intro=2=2SG 2.PST.R=MIDD-meet=DU=SG devil/worm
 'You came across with the devil.' (Palancar 2009: 545)

In (41), the number marker *wí* on the verb conveys resolved agreement with the two NPs referring to the sum of the two individuals, i.e., the hearer and the devil. In this respect, it is naturally expected that we gloss it as a dual *as*, after all, the marker is cognate with the dual markers found in other languages we have seen.³³ However, in other cases the actual sum of the individuals included in the set referred to by the agreement marker is more than two. This is illustrated in (42), where it makes reference to three individuals, or more significantly in (43), where the mason masters referred to are interpreted as being a handful, although not very many. In these examples, I again indicate the reference to an elided conjunct NP by means of *pro*, and when appropriate I use a subscript letter *i* and *j* to track the anaphora to a particular referent (Palancar 2009: 560).

- (42) nú=ya bötsi, bi=ma ba=tsú=r
 SIT DEF.PL=PL child 3.PST=SS/go 3.PST.CISLOC=reach.up.to.B.3POSS[3OBJ]=SG
 mǎle_j
 grandmother

pa [*pro*_i] bi=m-'uh=wí [*pro*_j]
 PURP 3.PST=SS.MIDDL-live.AS=DU
 'The (two) children went to reach up to their granny to live with her.'

- (43) kó ya [*pro*] m-pǒh=wí [nú=ya mástro]
 SIT because P [3.PRS.R]MIDD-know.AS=DU DEF.PL=PL master
 'Because he gets on well with the mason masters.'

I take examples (42–43) as evidence that the old dual markers in the split conjunction construction of Northern Otomi have been reanalysed as exponents of a paucal number value (the reason why I still gloss it as dual will become clear below in Section 4.3.1). In fact, plural markers are only used in the construction when there is a large group of people. In other words, the plural has become a plural of abundance or a remote plural, like in (44), where the members of the family are interpreted as forming a large collective of individuals (Palancar 2009: 560).

- (44) yó_kreo_ké [*pro*] bi=gó=hú [nú'=yá fámiliá]
 SIT I.think.that 3.PST=SS/remain.AS=PL DEF.PL.POSS=PL.3POSS family

'I think they stayed with their family.'

Outside the split conjunction construction, however, dual forms never appear in the elicitation of inflectional paradigms of verbs, and there are no dual forms of pronouns. This is illustrated in Table 11, which completes the information partially introduced in Table 5 above.³⁴ Here, like in Eastern Otomi in Table 10, the plural form of the third person is a demonstrative and is not based on a regular plural, i.e. there is no such a form as **nǒ'óhú* for an equivalent of 'they'.

Table 11. Personal pronouns in SIT Otomi (Northern Otomi)

		SG		PL	
1st	NEUTRAL	nú=gá ⁴⁵	EXCL	NEUTRAL	nú=khé
		nú=gí			nú=gí=hé
	EMPH.	nú=gí=gá		EMPH.	nú=gí=khé
			INCL	NEUTRAL	nú=khú
				EMPH.	nú=gí=khú
2nd	NEUTRAL	nu=gé			nú='í=hú
		nú='í			
	EMPH.	nú='í=gé			
3rd		nǒ=ǒ			nú='ú

Similarly, the typical paradigm of a verb reflects a singular-plural opposition, which is shown in Table 12, illustrated with the subparadigm of the present realis of the verb 'yo 'walk'.

Table 12. Present realis of the verb 'yo 'walk' in SIT Otomi (Northern Otomi).

		SG		PL
1st	dí='yo		EXCL	dí='yo=hé
			INCL	dí='yo=hú
2nd	gí='yo			gí='yo=hú
3rd	'yo			

In this light, the number system in the Northern language is complex because it is a hybrid one. Outside the split conjunction construction, the language has a basic singular-plural number opposition and only for the first and second person. In contrast, in the split conjunction construction, number marking for the three persons is obligatory, and the historical dual markers work here as paucal markers, while the same plural markers in Table 12 are only rarely used, and when they are,

they convey a plural of abundance. This situation has made some authors analyse the historical dual markers in the split construction as associative morphemes (Ecker 1952, Hess 1968), mainly because the construction is used to convey comitative semantics. In the next section, however, I show that the dual still survives as a number value in Northern Otomi, making their proposed treatment as associative markers impractical or unnecessary.

4.4.2 *Duals conditioned by register*

Languages often keep their treasures from us, only to reveal them once we are lucky enough to strike at where they lie hidden. For instance, none of the main references of Mezquital Otomi, the widespread variety of Northern Otomi, such as Bartholomew (2004) or Hess (1968), an often-cited work used in the *World Atlas of Language Structures* (Dryer and Haspelmath 2011), mention the dual as an existing number value in this language. However, a closer look at natural texts reveals that there is in fact more to the dual in this language than meets the eye.

In Northern Otomi, the old dual has become a number value restricted to a very particular register. The dual can only be used among adult men, and here the speaker uses it to portray actions that he and a friend can carry out together. By doing so, he enhances the comradeship between himself and his friend.³⁵ In this context, the use of the dual is optional and alternates with the plural, which is otherwise obligatory. As the use of the dual is motivated by the pragmatics of the situation at hand, it is no longer a morphosyntactic feature value required by the syntax in these languages.

As an illustration of this male register, consider the textual extract in (45). The text comes from the book *Etnografía Otomí*, written in Mezquital Otomi by Salinas Pedraza, himself a native male speaker. In particular, the extract is from a story about two merchants (A and B) who have to overnight in a forest and decide to sleep in a cave, where they would later come across a monstrous snake that will devour one of them. The piece of text features their conversational exchange at the point when they decide to go and sleep in the cave. Notice that there are eight instances in (45) where a non-singular first person form is used. Here, the dual alternates with the plural half of the time (Salinas Pedraza 1983: 138).

(45)

MEZ	<i>N'a di ge'u bi 'neembä rá n'owi:</i>	One of them told his comrade:
A	<i>-“te gä 'ot'wi, n'o?</i>	DU –“What shall we do, man?
	<i>bi nxüi</i>	It's dark
	<i>ha hindi handwi rä 'nüü;</i>	DU and we don't see the way;
	<i>xä nts'uni,</i>	it's dangerous,
	<i>maa grä tahuwa ha yä t'oho”,</i>	PL we could trip here in the woods,”

	<i>bi ʼneembä toʼo mi nʼowiʼna</i>	he said to him, with whom he was walking, they say.
B	–“hää, nʼo”– <i>bi dädiʼä mänʼaʼna</i> –	–“Yes, man”– the other one answered –
A	–“te gi mää <i>gä honhu</i> <i>habu gä kohwi?</i> <i>xtä hayaatsʼi,</i> <i>gä ʼapäwi rä zi Daada</i> <i>rä mäkä ʼnüü mänʼaki</i> ” – <i>di nʼembiʼna</i>	–“What do you say PL (if) we find DU where we could stay? In the morning, DU weʼll pray to God that he stretches the way again,” – he said to him, they say.
B	–“di pääka <i>habu kha nʼa rä kutsʼi,</i> <i>kha dä za gä ʼäähuni</i> <i>yaa hingä yaabu mädeewatho</i> ”	–“I know where thereʼs a cave, PL where we can sleep itʼs not a long distance to walk.”
A	–“ <i>maahaʼmu</i> ” <i>bi ʼneengä nuʼä mänʼaʼna</i>	PL –“Letʼs go then,” he said to the other one, they say.

In this register, the dual is perceived as a grammatical token to increase the degree of social symmetry between two men. It is not found in the discourse of women, who in similar circumstances would only use the plural, never the dual, and it is never used by a minor when addressing an elder. In this respect, consider for instance the extract in (46), which comes from a story told by the same author, but this time about himself and his own father. Notice that the plural is used throughout the story when the speaker, in his role of son, is recounting the story, but the dual is used in the words spoken by the father when he is addressing the speaker to suggest the two of them should quickly go home together. The use of the dual, in the mouth of the father, conveys a sense of bonding between father and son (Salinas Pedraza 1983: 122).

(46)

MEZ	<i>Ndi nʼooʼmee mä daada</i> <i>xki meengä rä theexʼuada nʼa rä xeni</i>	PL	I was walking with my father coming back from cutting maguey at a place called Naxtʼey.
	<i>rä thuuhu Nääxtʼeei.</i> <i>Ha tsukhee rä nduthe rä thuuhu.</i>	PL	Thirst and hunger had seized us.
	<i>Mi ʼeengägi mä daada:</i>		My father said to me:
Father:	–“ <i>ntiihi, nʼo,</i> <i>gä tsoni gä n̄unui bi ndee</i> ” – <i>ʼenä.</i>	DU	–“run, man! (so that) we get home and eat in the afternoon” – he said
	<i>Ndi ʼrañhee yä kati, ya hñee</i>	PL	We crossed the country, the canyons
	<i>pa gä maahee rä ʼmeñi.</i>	PL	so that weʼd go to work
	<i>ha pa gä penkhee mä ngühee.</i>	PL	so that weʼd return to our home.

<i>Dä tsoohee geetuä ha n'a rä hñe.</i>	PL	We arrived at a creek.
<i>Nde kha maagä tsohee ha rä hñee</i>	PL	Just as we were entering the canyon
<i>Nt'agi,</i>		of the Pantheon,
<i>nguu xä dä nk'axee</i>	PL	when we were looming up there
<i>bi ungä n'a ndä mfoote</i>		it (the rattlesnake) gave a great leap
<i>pa maa xä mut'a mä daada.</i>		to bite my father.

These data show that the dual is a number value that is still accessible to speakers of Northern Otomi, but it is only licensed by the right pragmatics of the situation, whereas otherwise the plural is used. In Section 5, I advance a likely scenario for why the dual has attained this meaning. In contrast, the dual thrives in the split conjunction construction where it has been reanalyzed as a paucal. On the other hand, for agreement purposes, Northern Otomi has also lost number agreement for the third person (dual or plural), except in the split conjunction construction. This is also the situation in Eastern Otomi. In general, the situation we find in these languages varies much from the original historical situation still found in Western Otomi. As I will show next, however, the Southern languages display yet another outcome.

4.5 The number system in the Southern languages

The Southern dialectal area of Otomi comprises three different languages which have developed in relative isolation from the rest of the family. Otomi has been spoken in these enclaves long before the arrival of the Spaniards. The three varieties are Tilapa Otomi, Ixtenco Otomi, and the language spoken in the villages of Atlapulco and Acazulco.

As in Northern Otomi, the languages of the South have rearranged the original number system and have reduced the functional space of the dual. In the South, while the historical dual markers have been retained, they have become exponents of a plural value. The clearest example is illustrated in Ixtenco Otomi in Table 13 with the subparadigm of the past (realis) of the verb *mati* 'call someone' (adapted from Lastra 1997: 33). The same situation is shown in the pronominal forms in Table 14 (adapted from Lastra 1997: 29).

Table 13. Past of the verb *mati* 'call' in Ixtenco Otomi (a Southern language)

	SG		PL
1st	dá=mati	EXCL	dá=mati=ga=mé
		INCL	dá=mati=wí
2nd	gá=mati		gá=mati=wí
3rd	bi=mati		

Table 14. Personal pronouns in Ixtenco Otomi (a Southern language)

	SG			PL	
1st	<i>nugagá</i>	nú=ga=gá	EXCL	<i>nugagá-mé</i>	nú=ga=ga=mé
			INCL	<i>nuga-wí</i>	nú=ga=wí
2nd	<i>nukʔí</i>	nú=k'í		<i>nukʔí-wí</i>	nú=k'í=wí
3rd	<i>nuzá</i>	nú='á		<i>nuʔí</i>	nú='ú

The texts in Lastra (1997) have plenty of instances of inflected forms in the plural, but only a few of them actually indicate a clear plural reference, that is, cases where the subject involves more than two individuals. One of such examples is given in (47) and serves well our purpose to show that the markers are not dual (or paucal). Here the subject is a group of construction workers who are heating up lunch for the rest of their fellow workers and speak of themselves as a group (adapted from Lastra 1997: 187).

(47) *nuya* γ *ñõhi dipaʔtí* γ *hmé zembibi tegiʔitewí* // *bidá* γ *ñõhi zena bi paʔtí* γ
 IXT *hméwá nge di zitegamé mayúthegamé* // *hă zăkhá zena šícu hă zăkhá pĕmí,*
pĕmí

nu=ya γ *ñõh* di=pa'tí γ *hmé,*
 INTRO=now PL male.mate 3.PRS.IRR=heat.up[3OBJ] PL tortilla
 'When the men were heating up the tortillas.'

'embi-bi: -te *gí='ute=wí?*
 [3.PRS.R]say.AS-3DAT what 2.PRS.R=do[3OBJ]=PL
 'She says to them: -What do you (PL) do?'

bi=dá= γ *ñõh*, 'ena, Ø bi=pa'tí
 3.PST.R=SS/answer.AS=PL male.mate SAY REL 3.PST.R=heat.up[3OBJ]
 γ *hmé=wá:*
 PL tortilla=here

'It's being said that the men who were heating up the tortillas here answered.'

-nge *dí='ute=ga=mé* ma *yúthe=ga=mé*
 INTRO 1.PRS.R=do[3OBJ]=1=PL 1POSS water.duct=1=PL
 'We're building our water ducts.'

-hă, zăkhá -'ena *šítsu-* *hă, zăkhá, pĕ=mí, pĕ=mí*
 yes well [3.PRS.R]say woman yes well work=PL work=PL
 '-Yes, well -says the woman- yes, well, work!, work!'

In Ixtenco Otomi, the plural markers of Old Otomi have been lost,³⁶ the only surviving number inflection are the original dual markers that have become

exponents of the plural number. In Tilapa Otomi, the situation is similar, but shows a more mixed outcome.

In Tilapa Otomi, the default number system consists again of a singular-plural opposition. The language has, however, preserved the old plural marker *hú* but uses it as a dedicated plural inclusive marker of first person, rather than as the general plural marker as in the varieties of the other areas. This mixed system is given in Table 15 by the subparadigm of the present continuous realis of the verb *ñũ* 'see', and in the pronominal system in Table 16, where we find the reflexes of the old dual *'mbé/wí* functioning as plural markers.

Table 15. Past of the verb *ñũ* 'see' in Tilapa Otomi (a Southern language)

	SG		PL
1st	txá=ñũ	EXCL	txá=ñũ='mbé
		INCL	txá=ñũ=hú
2nd	grá=ñũ		grá=ñũ=wí
3rd	ra=ñũ		

Table 16. Personal pronouns in Tilapa Otomi (a Southern language)

	SG		PL
1st	ñú=gá	EXCL	ñú=ga='mbé
		INCL	ñú=ga=hú
2nd	ñú=k'é		ñú=k'e=wí
3rd	ñú=á		ñú=ki

The situation in the split conjunction construction is useful again to illustrate a number of points about this particular number system. Examples (48–49) show that the old dual markers now convey a clear plural reference because the number of individuals involved features more than two. It should be noted that the markers in question are not conveyors of a paucal number value as in Northern Otomi in (42–43) above. Like in the examples above, the two conjuncts in the construction appear in brackets; I use *pro* to indicate an elided conjunct.

- (48) [ñú=gá] gu=tu='mbũ='mbé=[yú]
 TIL INTRO=1 1.PRS.IRR=1=live=PL.EX=DEM.PL
 'I will live with them.' (Txt)

- (49) [*pro*] g<w>a=tú-n-thé=wí yu khá'ni
 TIL <CISLOC>2.PRS.AMBU.R=INFL-MIDD-meet=PL DEM.PL man
 'You are meeting those men over there.' (Txt)

Similarly, as in all other Otomi languages, number agreement is obligatory with a third person in the split conjunction construction. The marker used for this is the old general dual marker *wí*, but just as in (48–49), it is also an exponent of a plural value. This is illustrated in (50), where again the number of referents is more than two (and also more than a handful).

- (50) [’ně=ha=k’ú] bi=mbɛ=wí [yi xú^htsi]
 TIL and=ENCL=3PL 3.PST.R=SS/go.PL=PL PL girl
 ‘And they went with the girls.’

But more interestingly, within the split conjunction construction, the old dual marker *wí* still realizes a dual value for the first person inclusive. This is shown in (51), which contrasts with the use of the plural marker *hú* in (52), which renders a plural inclusive.

- (51) ’ně=[k’é] [pro] tú=n-thé=wí
 TIL and=2 1.PST.R=MIDD=meet=DU.INCL
 ‘And I met with you (SG).’
- (52) ’ně=[k’e=wí] [pro] tú=n-thé=hú
 TIL and=2=PL 1.PST.R=MIDD=meet=PL.INCL
 ‘And I met with you (PL).’

After the reanalysis of dual markers as plurals, Tilapa Otomi obtained some mismatches in the otherwise regular assignment of semantic value to exponents. While the old dual marker *’mbé* becomes an exponent of the first person plural exclusive, the old regular dual *wí* now has two functions: to serve as a dual inclusive for the first person; and as a plural for the second and the third persons (but only within the split conjunction construction). Similarly, the old regular plural *hú* has become a dedicated first person plural inclusive.

4.6 Summary of the number system of Otomi

Up to this point, we have seen that the original number system of Old Otomi based on a singular-dual-plural value opposition underwent considerable restructuring in the modern languages except in Western Otomi. A summary of the outcomes of this restructuring is given in Table 17.

As indicated in Table 17, Western Otomi is the only modern language which has preserved the situation of Old Otomi. Northern and Southern Otomi are more innovative. The Southern languages have lost dual number altogether, the old dual markers having been reanalysed as exponents of the plural number. In Northern Otomi, vestiges of the dual survive in male speech and in the split conjunction

Table 17. The number system in the Otomi family

	Old Otomi	West	East	North	South	
Plural	Plural agreement for 1st & 2nd person subjects	Yes	Yes	Yes	Yes	Yes
	Plural agreement for 3rd person subjects (only specific humans)	Yes	Yes	No	No	No
	Plural agreement for 3rd person subjects in the split conjunction construction	Yes	Yes	Yes	Yes (as plural of abundance)	Yes
	Plural pronominal forms	Yes: 1/2/3	Yes: 1/2/3	Yes: 1/2	Yes: 1/2	Yes: 1/2
	Meaning of plural	'more than two'	'more than two'	'more than two'	'more than one'	'more than one'
	The dual markers now realize plural	No	No	No	No	Yes
Dual	Dual agreement for 1st & 2nd person subjects	Yes	Yes	Yes	Yes (only in male speech)	No
	Dual agreement for 3rd person subjects (only specific humans)	Yes	Yes	No	No	No
	Dual agreement for 3rd person subjects in the split conjunction construction	Yes	Yes	Yes	Yes (as paucal)	No
	Dual pronominal forms	Yes: 1/2/3	Yes: 1/2/3	Yes: 1/2	No	No
	Meaning of dual	'two'	'two'	'two'	'two male friends'	—

construction where the old dual now works as a paucal. Eastern Otomi shows a mixed situation, leaning towards the old system. In the next section, I will discuss these outcomes in the light of similar trends found elsewhere to understand their typological relevance.

5. Discussion: The evolution of number in Otomi

In this paper, I have shown two interrelated phenomena in the grammar of number in Otomi: one involving subtleties in the distribution of number agreement in verbs; the other involving the loss of the dual as a morphosyntactic value in the number system of some Otomi languages.

As for the first phenomenon, we have seen that Western Otomi has preserved the historical situation present in Old Otomi based on a singular-dual-plural opposition. This situation is typologically close to the one we can still observe in the Pamean branch of Oto-Pamean. In Western Otomi, however, agreement is only possible when the referent is human and specific, as shown in Section 4.1. In contrast, all other Otomi languages have lost number agreement with a third person subject except in the framework of the special conjunctive construction I call ‘split conjunction’, presented in Section 4.2, which is used to convey comitative semantics. In such a construction, number agreement with a third person subject is obligatory.

On the other hand, while the dual remained a healthy morphosyntactic value in the syntax of both Western and Eastern Otomi, as shown in Section 4.1 and 4.3, respectively, in Northern Otomi and in the Southern languages it became dramatically weakened. There is no evidence that this weakening was contact induced, and since it yielded very different outcomes, we can safely claim that the loss of the dual in the different areas of the family took place as a completely independent process which happened in a relatively short span of time (see Section 2 for a brief discussion).

The loss of the dual in Otomi displays a number of unreported developments, it informs us of intriguing traces left behind in its disappearance, and more importantly, it represents a case where a variety of different possible scenarios happened within very closely related languages.

In Northern Otomi, the dual became a facultative number value. Although facultative duals are well-known in the literature, as discussed in Corbett (2000: 42ff.), they commonly involve nominal number only, like in Slovene (and in a number of modern Arabic dialects, see Blanc 1970). What is typologically significant about the facultative dual of Northern Otomi is its distribution and usage. As for its distribution, it is not associated with nouns, but with pronominal number, and more specifically it is only found in verbal agreement (i.e., the language does not have dual personal pronouns). As for usage, the dual is a facultative value restricted to a specific male register where it is used to enhance partnership and camaraderie between two men.³⁷ This was shown in Section 4.4.

In this connection, the use of the dual in this language, conditioned as it is by the pragmatics, should also be understood as involving a subtle change in the

meaning of the value which has yet to be reported to my knowledge. The dual is not only used here as a result of a computation involving the formula ‘one person plus one person equals two people’, there is more to it: the two individuals involved are perceived as forming a unit of two friends who act together as a team. Additionally, this duality is also gendered, that is, it can only involve two males, and they should be of equal social status, i.e. they stand in a symmetrical relation to each other. One question remains: How did this new sense evolve in the dual of Northern Otomi? I suggest that the preservation of the dual in the split conjunction construction could serve as a possible scenario for this semantic development.

The split conjunction construction is used to convey comitative semantics involving two human subjects (or human-like entities, e.g. deities). In other words, the construction is used to convey common acts of companionship such as ‘I ate with you’ or ‘the man came along with the priest’ framed in a conjunctive frame such as “I ate-DU you” or “the man came-DU the priest”, where the two conjuncts appear in different positions in the clause, hence the split. This construction cannot be used to encode states of affairs involving inanimates such as ‘the knife lies with the fork’ and not even events involving animals such as ‘the man went away with the dog.’³⁸ This is the type of semantics the construction bears in all Otomi languages, and I propose it had a bearing on the meaning of the dual.

The dual is the default value in this conjoining construction; the plural is rarely used because acts of accompaniment prototypically involve just two individuals. One likely scenario to account for how the dual became a sign of male camaraderie is to propose (i) that the semantics of companionship, proper of split conjunction, percolated to the uses of the dual outside the construction when the dual was still a healthy value; but (ii) that this development only occurred in the speech of male speakers, who would have used it to establish symmetrical male bonding with other men of their own age within the speech community. In contrast, as women rarely establish symmetrical bonds with other women, they did not use it. For example, tradition has it that newlywed Otomi women must move away from their own homes to live with their in-laws, and this situation compels them to establish new social bonds with the women of their new family. Such bonds are commonly non-symmetrical.³⁹

In this sociolinguistic context, speakers may have strongly perceived that the dual has come to mean something else and much more subtle than just two people. Consequently, this perception gave rise to its loss in usage by the linguistic community at large in broader situational contexts, while it was progressively relegated to a specific type of male speech, motivated by specific circumstances. In this way, as the meaning and usage of the dual had changed, the old plural morphology took over its morphosyntactic function.

But as the dual was becoming a token of male register in Northern Otomi, within the split conjunction construction itself, the old dual forms survived. However, they underwent another semantic change by virtue of which they were no longer perceived as exponents of dual semantics per se, but as exponents of a paucal number value. This change in meaning may have been motivated again by the common situation that an accompaniment situation commonly involves two individuals, and as the dual was losing its morphosyntactic function outside the realm of the construction, it also opened up the possibility that its semantic reference could become more flexible to also incorporate instances of three or more people, as long as the sum of the individuals involved was only a handful. This new change forced the semantics of the old plural markers to be reinterpreted as exponents of a plural of abundance within the same conjunctive frame, resulting in a hybrid number system for the same language: while the default case has a singular-plural opposition, it is singular-dual-plural for the specific register of males, and paucal-plural for the conjoining construction.

In this scenario, the dual became a paucal, which appears to be a rather rare development according to Corbett (2000: 25); the notorious sources for paucals being trials (e.g. the Malaitan languages spoken in the Solomon Islands, Simons 1986: 33). The only example pointed out in Corbett's sample is a small number of Arabic dialects, presented in Blanc (1970), and this only involves nominal number. Another instance, also involving nominal number, is the odd case inflection of nouns modified by numerals 2 to 4 in Russian. Traditionally interpreted as a genitive singular, Corbett (2012) prefers to treat this inflection as the 'adnumerative case'. The adnumerative, shown in (53a), contrasts with the use of the genitive plural with numerals 5 and above, as in (53b).

- (53) a. *dva* *žurnal-a*
 RUSS two.M.NOM/ACC journal-ADN
 'two magazines'
- b. *pjat'* *žurnal*⁴⁰
 five journal(GEN.PL)
 'five magazines'

What is crucial in this scenario is that the form realizing the adnumerative in (53a) is the vestige of the old dual, which went lost in Slavonic except in Sorbian and Slovene. One possible interpretation for this change involves dual number first being reanalysed semantically as a paucal, covering references from two to four entities in the world, just as in Northern Otomi. In Russian the old dual was confined to the numeral modification construction; in Northern Otomi to the specific frame of the split conjunction construction.⁴¹

As the dual was being weakened in Northern Otomi, in the isolated languages of the South, the dual was reanalysed as a plural, but here, the change involved the subsequent loss of the old plural markers. This development is presented in Section 4.5 and it lies in connection with the reanalysis of the pronominal system of Germanic languages like Icelandic and Norwegian, as we saw in Section 1.1. An almost total loss of plural markers is evidenced in Ixtenco Otomi,⁴² while a partial one is observed in Tilapa Otomi, as the dual and the plural markers were preserved with their original meaning only for the first person inclusive. It is more difficult to envision how this process may have happened, but whatever the process might have been, prior to it, the change must have also involved a semantic change from dual to paucal, as we saw above, and then a later reinterpretation as a general plural.⁴³

6. Conclusion

Otomi started to diversify as a linguistic family, probably in the beginning of the 18th century, from having been a dialectal continuum in the colonial times of New Spain. One of the grammatical domains in which the Otomi languages differ most remarkably is in the many ways they exploit pronominal number in both verbs and pronouns, especially in their treatment of the dual number. For example, in Northern Otomi, while the dual vanished from the pronouns, it survived in pronominal cross-referencing on the verb in two contexts: (i) as circumscribed to a male register where it is a token of male camaraderie; and (ii) within the constructional frame of split conjunction where the old dual has been reanalysed as a paucal. In contrast, in Southern languages such as Tilapa Otomi and Ixtenco Otomi, the old dual markers have become exponents of a plural value. This is surprising, because when the dual number is lost as a semantic value, its morphological exponence is typically also lost.

In general, the distribution and use of the pronominal number systems of the Mesoamerican languages of Mexico and Guatemala remain understudied to date. The dual is not a very common number value in the languages of the area, although it is well attested in the Oto-Pamean languages of Oto-Manguéan (Suárez 1983:81–82). While Otomi has been known to have had dual number for a long time, I hope to have been able to show in this paper that the situation is not that simple when looked at in further detail. This means that stating that a given language has a dual is no longer as informative as one would want that label to be. One also needs to state what type of dual it is. In Mesoamerica, just as in Oto-Pamean, the dual has also been reported in other languages where it occurs sporadically, such as in Huave (Stairs and de Stairs, 1983) and in Lachixío Zapotec (Persons *et al.*, 2009), a Zapotecan language of the Soltecan group whose dual pronoun (only

for the first person) appears to be a newly discovered innovation not reported for other Zapotecan languages (see Operstein 2003). As more languages in this complex linguistic area, and other areas of the Americas, receive better and more complete descriptions, we will be able to uncover subtleties regarding number systems which may shed more light onto how such systems can evolve and decay.

Notes

* Acknowledgments: This paper was written under the auspices of the ESRC/AHRC project RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguan languages”. I heartily thank Greville G. Corbett and Anna M. Thornton for all their invaluable comments on an earlier draft of this paper. I am most grateful to the two anonymous referees for their comments and suggestions, and to the journal editor, Balthasar Bickel, for providing ways to improve its structure. Finally, I also want to thank Penny Everson for proofreading the English in the text. All errors and deficiencies remain my responsibility.

1. The language in question establishes whether the division with the singular starts at ‘more than one’ (i.e., English *one and a half days*, where the plural is used just as in *two days*) or ‘two or more’ (i.e., Spanish *un día y medio* (**uno y medio días*), but *dos días*).

2. In Classical Greek nouns, however, the dual is already restricted to things that may come in pairs, such as *hýppos* ‘horse’ whose dual form *hýppō* (nominative/accusative/vocative) refers to ‘the two horses of a chariot’, not just two horses in general (Monro 1891: 161).

3. Another possibility is verbal number, the study of which lies outside the scope of this paper.

4. About the dual number, there is the expectation framed under universal 564 in the Universals Archive accessible at <http://typo.uni-konstanz.de/archive/nav/browse.php?number=1&PHPSESSID=rhbji3mj02r3anuh5op4e3v49rjp2qcc>, which states that ‘if the dual extends to nouns, then it also extends to pronouns’ (see Plank and Schellinger 2000). There are numerous counterexamples to this universal (e.g. Celtic languages have vestiges of nominal number, but no trace of pronominal number), all of them discussed in Plank (1989). An additional counterexample is Otomi.

5. This is most notably the case in Indo-Iranian (mainly in Vedic), in Greek, in Germanic, in Balto-Slavic, and in Tocharian (in the dialect known as Tocharian B). However, it is not found in Anatolian, the earliest attested branch of the phylum, nor in Italic and Armenian, while in Celtic it only appears in nominal number. Whereas the absence of dual in such groups could in principle be interpreted as a historical loss of this value, the consensus is, nonetheless, that the dual number is a later development in the phylum so that it should not be reconstructed for the proto-language.

6. Another well-documented case of duals being lost is Arabic, where the modern dialects have also lost the old dual in the pronominal system (Kaye 1987).

7. The interested reader is directed to Corbett (2000) for an extensive exploration of the typological distribution of number systems with a dual value, to Jensen (1952) and Plank (1989) for dual diachrony, and to Plank (2003) and Plank & Schellinger (2000) for well-informed reviews of the implicational relations duals hold with other features typologically.

8. As for pronouns, in other old Indo-European languages with a singular-dual-plural opposition, the dual was also already relegated to the first and second persons only (e.g. Vedic nominative singular *ahám* 'I', dual (*ā*)*vám* 'we two', plural *vayám* 'we', and nominative singular *tvám* 'you (sg)', dual *yūvám* 'you two', plural *yuyám* 'you (pl)'; or Attic Greek nominative singular *egó* 'I', dual *nó(i)* 'we two', plural *hēmeís* 'we', and nominative singular *sú* 'you (sg)', dual *sphó* 'you two', plural *hūmeís* 'you (pl)'; Adrados *et al.* 1998: 31ff.). Anaphoric demonstratives, which are reanalyzed as third person pronouns in the different families, have only singular-plural opposition in all Indo-European languages.

9. Abbreviations: ACC accusative; ADN adnumerative; AMBU ambulative; AS adjusted stem; B bound shape; CISLOC cislocative; CIT citative; DAT dative; DEF definite; DEM demonstrative; DET determiner; DIM diminutive; DIR directional; DIST distal; DU dual; DUR durative; ENC phrase final enclitic; EXCL exclusive; F free shape; FEM feminine; GOTH Gothic; HO Huehuetla Otomi; HON honorific; IDENT identificational particle; IMM immediate; IMP imperative; IMPF imperfect; INCL inclusive; INF infinitive; INFL index of inflectional class; INTR intransitive; INTRO introductory particle (definite, topical, etc.); IRR irrealis; IXT Ixtenco Otomi; LOC.P locative preposition; M masculine; MEZ Mezquital Otomi; MIDD middle; NEG negative; NOM nominative; OBJ object; OE Old English; OO Old Otomi; P anterior particle; PL plural; POSS possessive; PRF perfect; PRS present; PURP.MOT purposive motion; PST past; REFL reflexive; REL relative; R realis; SF San Felipe Otomi; SIT San Idefonso Tultepec Otomi; SG singular; SS secondary stem; TIL Tilapa Otomi; TOL Toluca Otomi; TRANSLOC translocative.

10. <http://www.wulfila.be/gothic/browse/text/?book=2&chapter=17>.

11. <http://www.wulfila.be/gothic/browse/text/?book=6&chapter=9>.

12. <http://www.wulfila.be/gothic/browse/text/?book=1&chapter=9>.

13. Although perhaps a banal point, one should remember that the existence of a dual value in a given number system has a bearing on the meaning of the plural in that system, as the plural will then be used for three or more real world entities. This means that the plural of English is not the same as the plural in Sanskrit (a general point made by Saussure 1916/1971: 161, also in Corbett 2000: 5).

14. The verb *beceapian* 'sell' appears here in singular. Verbs in Old English may not agree in number with their subjects when they precede them (see Mitchell 1985: 635–639 for numerous examples).

15. In addition, Simon (2005) claims that in honorific usage, the pronoun *Sie* in Bavarian indicates an inclusive second person (i.e., a plurality of hearers), while *eß*, the default second person plural pronoun, would in contrast be used with an exclusive value (i.e., the hearer plus somebody else not present). I remain sceptical whether this is the right interpretation of the data. I want to thank the first anonymous referee for pointing this reference out to me.

16. Stone (1993b) reports that the old dual forms for the first and second person in Kashubian (i.e. *ma* 'we' and *wa* 'you (PL)') have acquired a plural meaning and alternate with the historical plural forms *mē* 'we' and *wē* 'you (PL)', although it seems they still preserved a dual meaning in the northern variety up to the first half of the 20th century (Breza and Treder 1981: 125 *apud* Stone 1993b: 773).

17. In a way, Oto-Pamean is to Oto-Manguean what the Anatolic branch is to Indo-European, that is, it is the earliest branch to split from the proto-language and is thus structurally very distant from the other branches of the phylum.

18. For the orthography of Otomi I use here the consensus used by most authors. As for consonants: *C* represents an ejective, and ^h*C* pre-aspirated; *f* /p^h/; *ʔ* /ʔ/; *ñ* /ɲ/; *c* /t͡ʃ/; *r* /r/; *x* /ʃ/; and *y* /j/. As for vowels, *a* represents /ɔ/ [v]; *e* /ɛ/; *o* /ɔ/; *u* /i/, and *i* /ɛ/, while the dieresis indicates a nasal vowel (e.g. ä /ã/ [ã], etc.).

19. Additionally, we find here *-n* instead of the expected *-s*.

20. Angulo's orthography has been slightly modified. The circumflex represents a middle tone; low tone is not represented; a superscript vowel indicates a whispered vowel; and an upperscript *p* indicates a salient plosive release.

21. I use the term 'inflectional formative' in the sense of Bickel and Nichols (2007), who use it as a cover term for any marker of inflectional information, regardless of its morphological type, that is, regardless of whether it is an affix, a clitic, or a functional word.

22. While lexical roots have a three-tone contrast (high, low, and ascending), functional formatives only have a high vs. low tone contrast.

23. Here the plural pronoun has specific reference (e.g. 'Peter, Mary and their friends' or 'those men standing over there'). In other words, the reading is not ambiguously impersonal like in English. For the impersonal reading Otomi has a different construction.

24. These in turn emerged from a combination similar to the form *nú=gí=hú* which underwent phonological reduction of the pronominal clitic, i.e., **nú=g(i)=hú* > **nú=k=hú* > *nú=khú*.

25. The verb in (11) appears in a special morphological shape in SIT Otomi called 'bound shape'. This shape is used when the verb is followed by an NP or some other constituent of its own clause; see Palancar (2004, 2011) for more details.

26. Out of context, example (11) can also mean 'I went to store our luggage,' but the plural subject reading is preferred.

27. The fact that these verbs cannot have metathesized outcomes (e.g. **támhbú*, **pámhbú*, etc.) like the ones in (16) indicates that an alternant such as *ú* is not a merely phonetically conditioned allomorph.

28. Cárceres further points out that the verb does not inflect for the number of the subject when the conjunct bearing the coordinating conjunction follows the verb. This is shown in (i). This is a structure that can, nonetheless, be well accounted for as an instance of gapping.

(i) *nūko San Pedro pñioxhmābāte xīnēko San Pablo*

oo nū=k'=o San_Pedro p<n>i=xāhñabate__ xi='ne=[k'=o
INTRO=DEM.SG.III=HON Saint_Peter <INTR>3.PST.R=teach also=and=DEM.SG.III=HON
San_Pablo]
Saint_Paul
'Saint Peter preached and also Saint Paul.'

29. In Table 7, I have not included predicative forms such as *cogueque* (*ko=geh=ke*) ‘it’s me’, which are introduced by means of the focus particle *co* (*ko*) and the identificational particle *gue* (*ge/geh-*), and require special pronominal forms (i.e., *ke* ‘I’).

30. The form *gégé* of Toluca Otomi is already attested in the predicative form for the third person in Cárceres, i.e., *coguegue* (*ko=ge=ge*), and stems from a lexicalization of the identificational particle *gue* (*ge/geh-*) plus the genuine third person form (i.e., *ge* ‘he/she’).

31. There is only one example in Lastra’s (1992) extensive corpus that shows number agreement with an apparent non-specific plural subject. This is given in (ii). Whether the speaker had some specific people in mind is not known.

- (ii) *pũ xo dañũtu hũ damábihúyĩ wũ damákáhúya wũ*
 TOL *pũ xo da=ñũtu=hũ da=má=bĩ=hũ=yĩ=wũ*
 maybe so 3.PRS.IRR=SS/get.in.AS=PL 3.PRS.IRR=PURP.MOT=SS/steal[3OBJ]=3PL=then
 ‘It may be that they got in to steal it (the food) then,’
da=má=ká=hũ=yĩ=wũ
 3.PRS.IRR=PURP.MOT=bring[3OBJ]=3PL=then
 ‘and then take it.’ (p. 323)

32. In Old Otomi, the coordinating conjunction *xi* (*ma*) ‘*ne*’ and’ (Lit. ‘also (other) and’) was replaced by the loanword *ko(n)* from Spanish, although not entirely, depending on the variety.

33. In Section 4.4.2, I further show that the dual is a number value characteristic of specific male register.

34. Similarly, there is no dual of possessor. The only context where the dual is still used associated with a noun is a nominal predication construction where a proper noun is used to predicate the identity of namesakes between two referents, as illustrated in (iii), where the proper name *Tasya* ‘Anasthasia’ functions as a nominal predicate with the meaning ‘be called Anasthasia’. The use of the dual here is related to the use of number of possessor.

- (iii) *ma=ga=hong-a=ma Tasya=’bé*
 SIT IMM=1.PRS.IRR=look.for[3OBJ]-B=1POSS Anasthasia=DU.EXCL
 ‘I’m going to look for my namesake Anasthasia.’

35. Salinas Pedraza (1983: 116–117) gives some illustrative ethnographic examples of typical Otomi male camaraderie interaction, which involve teasing and sexual wordplay.

36. Lastra (1997: 36) indicates that in Ixtenco Otomi, the plural markers of other varieties are still understood and they are perceived by speakers as plurals of abundance, but in reality they are not used.

37. We only have examples available in the literature with the first person, but it could be expected that the dual can also be used for the second person, although less frequently, or to the addition of different connotations.

38. Unless the situation described involves anthropomorphized animals as if they were humans, such as in folk tales of the type ‘the turtle walked along with the coyote’.

39. For anthropological information about the daily lives of Otomi women, see Molinar Palma (1997) or Van De Fliert (1988).
40. For the inflectional class to which *žurnal* 'magazine' belongs, there is syncretism between the nominative singular and the genitive plural.
41. Another possibility is that the dual was no longer used with nouns but survived when the noun was modified by the numeral '2'. Having lost its original dual meaning, the old dual forms may have been extended to neighboring numerals '3' and '4', but devoid of meaning.
42. See note 38.
43. In this connection, Corbett (2000:268) proposes that a change involving paucal to plural was a likely scenario for the number marking of Mokilese, a Micronesian language (Harrison 1976), where the original trial developed into a paucal before becoming a plural.
44. Angulo (1932) gives the form *kinusín* for this cell, but the high tone is probably a typo, as the second person possessive is systematically low in the second syllable for this class of nouns.
45. The pairs *nú=gá* and *nú=gí* for the first person singular, *nú=gé* and *nú=ʔ* for the second person and *nú=khé/khú* and *nú=gí=hé/hú* for the plural, are in free variation. The only difference lies in their usage, i.e., the first form is much more frequent in use than the second one.

References

- Acuña, René. 1990. Introducción. In René Acuña (ed.), *Arte breve de la lengua otomí y vocabulario trilingüe: Español-náhuatl-otomí*, i–lxx. Mexico City: Universidad Autónoma Nacional de México.
- Adrados, Francisco R.; Bernabé, Alberto and Mendoza, Julia. 1998. *Manual de Lingüística Indoeuropea, Volumen III: Morfología: pronombres, adverbios, partículas y numerales. Sintaxis: diferenciación dialectal*. Madrid: Ediciones Clásicas.
- Agud, Ana and Fernández Álvarez, Ma Pilar. 1988. *Manual de Lengua Gótica*. Salamanca, Universidad de Salamanca.
- Andrews, Henrietta. 1993. *The function of verb prefixes in Southwestern Otomí* (Summer Institute of Linguistics and University of Texas at Arlington Publications in Linguistics 115). Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Andrews, Henrietta, Doris Bartholomew, Richard Blight, Artemisa Echevoyen, Nancy Lanier, Donald Sinclair, Katherine Voigtlander, Steven Eglan, Paul Mellema & David Persons. 1983 [1978]. Otomi. In Steven Eglan (ed.) in collaboration with Doris Bartholomew and Saúl Cruz Ramos, *La inteligibilidad interdialectal en México: resultados de algunos sondeos*, 46–48. Mexico City: Instituto Lingüístico de Verano.
- Angulo, Jaime de. 1932. The Chichimeco language (Central Mexico). *International Journal of American Linguistics* 7. 152–194.
- Bartholomew, Doris. 1966. Observations about Matlatzinca Grammar. Ms.
- Bartholomew, Doris. 2004. Notas sobre la gramática. In Luis Hernández Cruz, Moisés Victoria Torquemada & Donald Sinclair Crawford (eds.), *Diccionario del hñāhñu (otomí) del Valle*

- del Mezquital, Estado de Hidalgo* (Vocabularios Indígenas 45), 473–495. Mexico City: Instituto Lingüístico de Verano.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume III: Grammatical categories and the Lexicon*, 2nd edn, 169–240. Cambridge: Cambridge University Press.
- Blanc, Haim 1970. Dual and pseudo-dual in the Arabic dialects. *Language* 46. 42–57.
- Bremmer, Rolf H., Jr. 2009. *An introduction to old Frisian: History, grammar, reader, glossary*. Amsterdam: John Benjamins.
- Breza, Edward & Jerzy Treder. 1981. *Gramatyka kaszubska*. Gdańsk: Zrzeszenie Kaszubsko-Pomorskie.
- Buelna, Francisco (ed.). 1893. *Luces del Otomí, ó Gramática del idioma que hablan los indios otomíes en la República Mexicana*, completed by 1767. Mexico City: Imprenta del Gobierno Federal.
- Cárceles, Fray Pedro de. 1907 [1580]. *Arte de la Lengua Othomi*, edited by Nicolás León. *Boletín del Instituto Bibliográfico Mexicano* 6. 38–155.
- CLIN: Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. 2008. *Diario de la Nación* (14 January 2008), Mexico DF. Accessible at <http://www.inali.gob.mx/pdf/CLIN_completo.pdf>
- Comrie, Bernard and Corbett, Greville G. (eds.). 1993. *The Slavonic Languages*, London: Routledge.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2011. *The world atlas of language structures online*. Munich: Max Planck Digital Library. Accessible at <<http://wals.info>>
- Ecker, Lawrence. 1952. Compendio de gramática otomí: Introducción a un diccionario otomí-español. *Anales del Instituto Nacional de Antropología e Historia* IV-32. 121–174.
- Einarsson, Stefán. 1945. *Icelandic: Grammar texts, glossary*. Baltimore, London: John Hopkins University Press.
- Faarlund, Jan Terje. 2008. *The syntax of old Norse: With a survey of the inflectional morphology and a complete bibliography*. Oxford: Oxford University Press.
- Forston, Benjamin W. 2004. *Indo-European language and culture: An introduction*. Oxford: Blackwell.
- Gordon, E.V. 1992[1957]. *An introduction to old Norse*, edn revised by A. R. Taylor. Oxford: Clarendon Press.
- Guðmundsson, Helgi. 1972. *The pronominal dual in Icelandic* (University of Iceland Publications in Linguistics 2). Reykjavik: Institute of Nordic Linguistics.
- Harrison, Sheldon P. with the assistance of Salich Y. Albert. 1976. *Mokilese reference grammar*. Honolulu: University Press of Hawaii.
- Hess, H. Harwood. 1968. *The syntactic structure of Mezquital Otomi*. (Janua Linguarum. Series Practica 43). The Hague: Mouton.
- Howe, Stephen. 1996. *The personal pronouns in the Germanic languages: A study of personal pronoun morphology and change in the Germanic languages from the first records to the present day*, (Studia Linguistica Germanica). Berlin: Walter de Gruyter.
- Jensen, Hans. 1952. Die sprachliche Kategorie des Numerus. *Wissenschaftliche Zeitschrift der Universität Rostock, Reihe Gesellschafts- und Sprachwissenschaften* 1(2). 1–21.
- Kaye, Alan S. 1987. Arabic. In Bernard Comrie (ed.), *The world's major languages*, 664–685. London, Sydney: Croom Helm.

- Kranzmayer, Eberhard. 1954. Der pluralische Gebrauch des alten Duals 'eš' und 'enk' im Bairischen: Ein Beispiel für Homonymenflucht. In *Festschrift für Dietrich Kralik: Dargebracht von Freunden, Kollegen und Schülern*, 249–259. Horn: Berger.
- Lastra, Yolanda. 1992. *El Otomí de Toluca*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Lastra, Yolanda. 1997. *El Otomí de Ixtenco*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Lastra, Yolanda. 2001. *Unidad y Diversidad de la Lengua: Relatos otomíes*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Lastra, Yolanda. 2010. Diversidad lingüística: variación dialectal actual. In Rebeca Barriga & Pedro Martín Butragueño (eds.), *Historia Sociolingüística de México*, vol. II, 841–880. Mexico City: El Colegio de México.
- Lastra, Yolanda & Doris Bartholomew (eds.). 2001. *Códice de Huichapan, paleografía y traducción*. Lawrence Ecker. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Macdonell, Arthur Anthony. 1910. *Vedic grammar*. Strassburg: Karl J. Trübner.
- Mitchell, Bruce. 1985. *Old English syntax. Volume I: Concord, the parts of speech, and the sentence*. Oxford: Clarendon Press.
- Molinar Palma, Patricia. 1997. Vida cotidiana y salud en mujeres otomíes del estado de México. Paper read at the Latino American Studies Association, Guadalajara, Mexico (Apr. 17–19). Accessible at <<http://lasa.international.pitt.edu/LASA97/molinarpalma.pdf>>
- Monro, David B. 1891. *A grammar of Homeric Greek*. Oxford: Clarendon Press.
- Mossé, Fernand. 1952. *A handbook of middle English*. Baltimore/London: John Hopkins University Press.
- Neve y Molina, Luis. 1767. *Reglas de ortografía, diccionario y arte del idioma othomí: Breve instrucción para los principiantes*. Mexico City: Imprenta de la Bibliotheca Mexicana.
- Operstein, Natalie. 2003. Personal pronouns in Zapotec and Zapotecan. *International Journal of American Linguistics* 69(2). 154–185.
- Palancar, Enrique L. 2004. Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70(3). 251–278.
- Palancar, Enrique L. 2011. Shape conditions in Otomi, invited paper read at the 3rd Morphology Meeting of the Southeast, University of Surrey, Guildford, UK (8 Apr.)
- Palancar, Enrique L. 2012a. Clausal juxtaposition and subordination: New light from Otomi. In Volker Gast & Holger Diesel (eds.), *Clause combining in cross-linguistic perspective*, 37–76. Berlin: Mouton de Gruyter.
- Palancar, Enrique L. 2012b. A typology of split conjunction. *Linguistic Typology* 16. 265–320.
- Palancar, Enrique L. 2009. *Gramática y Textos del Hñõñhõ: Otomí de San Ildelfonso Tultepec, Querétaro. (Volumen 1: Gramática)*. Mexico City: Plaza y Valdés.
- Persons, David D., Cheryl A. Black & Jan A. Persons. 2009. *Gramática del Zapoteco de Lachixío*. Mexico City: Instituto Lingüístico de Verano.
- Plank, Frans. 1989. On Humboldt on the dual. In Roberta L. Corrigan et al. (eds.), *Linguistic categorization*, 293–333. Amsterdam: John Benjamins.
- Plank, Frans. 2003. The selective elaboration of nominal or pronominal inflection. In Frans Plank (ed.), *Noun phrase structure in the languages of Europe*, 253–287. Berlin: Mouton De Gruyter.
- Plank, Frans & Wolfgang Schellinger. 2000. Dual laws in (no) time. *Sprachtypologie und Universalienforschung* 53. 46–52.

- Priestly, Tom M. S. 1993. Slovene, in Bernard Comrie and Greville G. Corbett (eds.), pp. 388–451.
- Ringe, Donald. 2006. *A history of English: Volume I: From Proto-Indo-European to Proto-Germanic* (A Linguistic History of English, Volume 1). Oxford: Oxford University Press.
- Salinas Pedraza, Jesús. 1983. *Etnografía del Otomí*. Mexico City: Instituto Nacional Indigenista.
- Saussure, Ferdinand de. 1971 [1916]. *Cours de linguistique generale* (publie par Charles Bally et Albert Sechehaye avec la collaboration de Albert Riedlinger), 3rd edn. Paris: Payot.
- Schwartz, Linda. 1985. Plural pronouns, coordination, and inclusion. In Nancy Stenson (ed.), *Papers from the Tenth Minnesota Regional Conference on Language and Linguistics*, 152–184. Minneapolis: Department of Linguistics, University of Minnesota.
- Simon, Horst J. 2005. Only you? Philological investigations into the alleged inclusive-exclusive distinction in the second person plural. In Elena Filimonova (ed.), *Clusivity: Typology and case studies of the inclusive-exclusive distinction* (Typological Studies in Language 63), 113–150. Amsterdam: John Benjamins.
- Simons, Linda 1986. The pronouns of To'abaita (Solomon Islands). In Ursula Wiesemann (ed.), *Pronominal systems* (Continuum 5), 21–35. Tübingen: Narr.
- Soustelle, Jacques. 1937. *La Famille Otomi-Pame du Mexique Central*. Paris: Travaux et Mémoires de l'Institut d'Enthnologie.
- Stairs, Glem A. & Emily F. de Stairs. 1983. *Huave de San Mateo del Mar* (Archivo de Lenguas Indígenas de México). Mexico City: El Colegio de México.
- Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4. 1–54.
- Stewart, Donald. 1966. *Gramática del mazahua*. México: Instituto Lingüístico de Verano. MS.
- Stone, Gerald. 1993a. Sorbian (Upper and Lower). In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 593–685. London, New York: Routledge.
- Stone, Gerald. 1993b. Cassubian. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 759–794. London, New York: Routledge.
- Suárez, Jorge. 1983. *The Mesoamerican Indian languages*. Cambridge: Cambridge University Press.
- Urbano, Alonso. 1605[1990]. *Arte breve de la lengua otomí y vocabulario trilingüe: Español-náhuatl-otomí*. Facsimile of Manuscripts Américains de la Bibliothèque Nationale, Paris, No. 8, XVII siècle, edition by René Acuña. Mexico City: Universidad Autónoma Nacional de México.
- Van De Fliert, Lydia. 1988. *Ar Hñahnu Hongar Nzaki: El otomí en busca de la vida*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- Voigtlander, Katherine & Artemisa Echegoyen. 1985. *Luces contemporáneas del otomí*. Mexico City: Instituto Lingüístico de Verano. [First printed in 1979 without the authors' names].
- Voigtlander, Katherine & Artemisa Echegoyen. 2007. Gramática del yuhú (otomí de la Sierra Madre Oriental), adapted from Katherine Voigtlander & Artemisa Echegoyen (eds.), *Luces contemporáneas del otomí*, 1979. Ms., Instituto Lingüístico de Verano.
- Wright, Joseph. 1910. *Grammar of the Gothic language*. Oxford: Clarendon Press.

Author's address

Enrique L. Palancar
 SeDyL-CELIA, CNRS
 7 rue Guy Môquet
 94800 Villejuif, France

Or:

Clausal juxtaposition and subordination: New light from Otomi

Enrique L. Palancar

1. Introduction

In this paper, I argue that in some languages clausal juxtaposition can be used not only as a mechanism of semantic or pragmatic subordination, but also as a device signaling syntactic subordination. This implies that the presence of clausal juxtaposition in the discourse of a given language does not necessarily imply the absence of the typical morphosyntactic hallmarks that characterize dependent clauses formally in other languages, especially if the relevant languages use juxtaposition as a typical clause combining device to encode semantically dependent events. In other words, I argue that it is important to keep apart the presence or absence of (i) syntactic dependency and (ii) markers of subordination (cf. Lehmann 1988). I illustrate this claim with data from a specific grammatical construction in Northern Otomi (Oto-Pamean, Oto-Manguéan) from the variety of San Ildefonso Tultepec, which involves a juxtaposed dependent clause with a number of (morpho)syntactic properties commonly associated with subordination, in particular "cross-clausal dependencies" as described by Gast and Diessel (this volume), and the prosodic realization of the clauses concerned in a single intonation contour.

1.1. Clausal juxtaposition: Parataxis and hypotaxis

Clausal juxtaposition is traditionally taken to be a type of clause-combining strategy used to encode paratactic relations among clauses. In this function, clausal juxtaposition is rather common cross-linguistically. An English example is given in (1), where the two clauses involved are conveyed to be textually related in some way (Quirk et al. 1985: 1472).

(1) *He ate too much for dinner. He was ill the next day.*

In examples like (1), it is essential to keep two distinct phenomena apart which are often erroneously equated: (i) a type of a clause linkage mechanism

and (ii) a type of a linking relation. The linkage mechanism used in (1) is that of “clausal juxtaposition”. This term implies that two (or more) clauses are chained by just placing one after the other without using an overt conjunction. The type of linking relation that the two clauses in (1) display is that of “parataxis”, traditionally defined as “the linking of elements of equal status” (Halliday 1985: 198). In (1) we have two matrix clauses of equal status.

Juxtaposing two clauses as a linking device naturally implies that there is no formal means indicating what semantic relation holds between them, but for examples like (1), Quirk et al. (1985) point out that the mere fact of using juxtaposition “raises the expectation that the second utterance [*He was ill the next day*] followed the first [*He ate too much for dinner*] as an iconic representation of being sequential in time or consequential in reasoning” (Quirk et al.: 1472).¹ In other words, in English examples like (1) clausal juxtaposition serves as a linguistic icon of notions such as sequence and/or consequence, which are the quintessential conceptual notions associated with sentential conjunction.

As an alternative to juxtaposition, the use of coordinating conjunctions (*and, but, or, etc.*) is yet another way in which sentential conjunction can be encoded, as for example in (2) (adapted from Quirk et al. 1985: 1472).

- (2) [*He ate too much for dinner*] *and* [*he was ill the next day*].

Coordinative constructions with conjunctions like *and* can also express sequence or consequence, but are regarded by many as imposing asymmetry on the conjuncts – beginning with Ross (1967), Schmerling (1975), Haiman (1985), Lakoff (1986), Deane (1992), and Johannessen (1998) – for example in the sense that the order of conjuncts cannot be reversed without affecting the meaning of the entire sentence. As both structures in (1) and (2) express semantic coordination, the coordinating construction in (2) is often regarded as “syndetic coordination”, while the clausal juxtaposition linking strategy in (1) is commonly treated as “asyndetic coordination”.

Haiman (1985) provides various examples of languages which make productive use of both asyndetic and syndetic coordinating structures. In all such cases, the asyndetic construction conveys a closer conceptual distance between the events involved, presumably as a result of iconic motivation. This may be illustrated with the contrast between (3a) and (3b) from

¹ Readings like these emerge when juxtaposition is used as a linking device encoding parataxis.

Feʼfe? Bamileke, a Bantu language (Hyman 1971, *apud* Haiman 1985: 113, translation and brackets mine).²

- (3) a. à kà gén nīē [ɲjwēn lwàʔ]
he PST go market buy yams
‘He went to the market and bought yams.’
b. à kà gén nīē [nī ɲjwēn lwàʔ]
he PST go market and buy yams
‘He went to the market and bought yams.’

According to Hyman (1971: 43 fn), example (3a) “definitely implies that yams were bought at the market”, whereas in (3b), the use of the conjunction disassociates the conjuncts, so that the second event is interpreted as having taken place at a later time and place.

While semantic coordination and parataxis could in principle be taken to be similar concepts at the semantic level, the literature also commonly treats the concepts of parataxis and clausal juxtaposition as being synonymous. This may be seen, for example, in Bowerin (2008: 171) and in Cosme Noonan (1985, 2006), Thompson & Longacre (1985), Longacre (1985), and Thompson et al. (2007).

In this contribution, I intend to show that clausal juxtaposition, despite its common paratactic function, can also be used to encode hypotactic relations, even in cases of highly integrated subordinate structures. I illustrate this claim in Otomi, a Mesoamerican, Oto-Manguean language family spoken in Central Mexico, in which such a strategy is extensively employed, and for this purpose I will use data from Northern Otomi, more specifically from the variety of San Ildelfonso Tultepec.

1.2. Distinguishing parataxis and hypotaxis

The common contention that juxtaposition equals parataxis can be illustrated with the chapter on subordination written by Aarts (2006) for *The Encyclopedia of Language and Linguistics* (second edition). Aarts (2006: 252) claims

² An analogous reading is found in the serial verb constructions of a number of languages, for example in Ono (Trans-New Guinea), where according to Wacke (1931, *apud* Haiman 1985: 112), juxtaposed verbs tend to render simultaneous events while coordinated VPs render sequential readings.

that “[s]ubordination is a matter of degree”, and referring to structures like in (1), he explicitly states that “clauses that are merely juxtaposed are paratactically linked”. The particular example under discussion is given in (4), about which the author rightly says: “Although there is an obvious semantic (causal) link between the first and second sentences in this example and co-reference is established through the use of the pronoun *he*, which may (but need not) refer back to Tim, grammatically neither sentence is subordinate to the other.”

(4) *Tim fell. He broke his arm.*

Note that what is at issue in (4) is not that this example is an instance of subordination – which it is not – but that the use of clausal juxtaposition should not be taken to be the same thing as parataxis. Not even in languages like English are such categories always clearly associated. For example, Culicover & Jackendoff (1997) argue in favour of an analysis that regards syntactic coordinating structures like those in (5) as instances of semantic subordination, despite the typically paratactic function of such constructions, as in examples like (2) above.

- (5) a. *One more can of beer and I'm leaving.*
 b. *You drink another can of beer and I'm leaving.*

Similarly, the construction in (6) – an instance of clausal juxtaposition – is also regarded as involving semantic subordination by Hoeksema and Napoli (1993) – again, in contrast to typical paratactic examples like (1).

(6) *I fainted, the sun was so hot.*

Prosody is one the most crucial features leading to a subordinate interpretation of examples like (6). Based on data from an experimental study carried out on English native speakers, Hoeksema and Napoli (1993) claim that the clauses in (6) are uttered within a single intonation contour. As we will see in Section 2.2, prosody is an important criterion that can be used to identify subordination when clausal juxtaposition is involved. The structure in (6) contrasts with the one given in (7), where the second clause is conventionally regarded as being syntactically subordinated.

(7) *The sun was so hot (that) I fainted.*

In this respect, Hoeksema and Napoli (1993) argue that the structure in (6) is located at an intermediate level of syntactic dependency, somewhere

between the structure in (1), which conveys semantic coordination of the asyndetic type, and the structure in (6), which is clearly subordinative.³ This problem has been identified by Givón (1990: 826):

[T]he absolute distinction between subordinate (“dependent”) and coordinate (“independent”) clauses works only within rigidly prescribed typological bounds; and even there only with rigid pre-selection of the facts. Both its sharp distinction and its utility dissipates rapidly when one considers a broader typological sample, and analyses a wider range of functional and grammatical facts. In particular, a careful parallel analysis of the structural facts of clause integration and the functional facts of event integration reveals that in many languages, the sharp binary distinction between subordination and coordination is a gross over-simplification of the facts.

In examples (5) and (6), clause-combining strategies which are typically associated with parataxis can be regarded as encoding semantic subordination, even in English. Semantic subordination, as conceived of by Culicover & Jackendoff (1997), refers to the semantic dependence between two clauses which is not (necessarily) realized formally in the typical syntactic terms that characterize other, more conventional hypotactic structures. In view of such cases, and in order to accommodate a wider range of structures which do not always show the same formal properties, Cristofaro (2003: 2) takes a functional typological approach to subordination as a grammatical phenomenon, providing the following definition: “(A) particular way to construe the cognitive relation between two events, such that one of them (which will be called the dependent event) lacks autonomous profile, and is construed in the perspective of the other event (which will be called the main event).”

This definition implies a notion of subordination that is independent of the way in which clause linkage is realized across languages. Such a definition allows Cristofaro to include languages which make extensive use of clausal juxtaposition to encode dependent events, such as Mandarin Chinese. In this new light, clausal juxtaposition is regarded as a possible strategy to encode dependent clauses. The claim may be illustrated with example (8) from Mandarin Chinese (Li & Thompson 1973: 98).

- (8) *nǐ guì-xiàlái [qǐn Zhāng-sān]*
 you kneel.down beg Zhang-san
 ‘You knelt down in order to beg Zhang-san.’

³ Hoeksema and Napoli (1993) use the term “parataxis” to refer to this intermediate level. In my opinion, the choice of the term is rather unfortunate in this context.

come from Northern Otomi, more specifically from the variety spoken in San Ildefonso Tultepec.⁵

- (9) a. *ně da = hñ̃>ōis'-i* [da = mǎ = pyá]
and 3.PRES.IRR=<SS>take[3OBJ]-F 3.PRES.IRR=SS/sell[3OBJ]=now
'And they took them (the tortillas) to sell them.' (Txt)
- b. *dá = jwá = hé* [dá = né = hé]
1.PST-finish.AS=PL.EXCL 1.PST=dance.AS=PL.EXCL
'We finished dancing.' (Txt)

The juxtaposed dependent clauses in (9) are syntactically reduced and are functionally (more or less) equivalent to typical infinitival clauses in Indo-European languages. The phenomenon under study in this paper appears to be a typological feature of Oto-Manguean, as similar structures have been reported in other groups, such as Trique (Longacre 1966) and Chatino (Pride 1965).

The present paper is structured as follows. In the next section, I introduce the grammatical construction instantiated in (9) in more detail, focusing on its prosody. In Section 3, I present a detailed analysis of the various grammatical properties that characterize the dependent juxtaposed clause type under consideration as a "deranked" structure, namely: it exhibits tense/aspect/mood (TAM) restrictions, it features a rather reduced syntactic structure, and it shows inherent control. In order to illustrate the various properties more clearly, I compare this clause type with other dependent clauses in Otomi which are headed by conjunctions, such as other types of purpose clauses and complement clauses. In Section 4, a number of additional morphosyntactic characteristics are presented which involve the relation between the main and the dependent clause, and which suggest that the construction at hand shows a high degree of syntactic integration. This can be seen in the marking of plural subjects and in the morphological adjustment of the verb in the main clause. In Section 6, I summarize the proposal and revisit the relation between clausal juxtaposition and the parataxis/hypotaxis distinction.

5 The source of a textual example is indicated as (Txt). All textual examples come from a large database collected in Mexico under the auspices of the SEP-CONACyT-47175 research project.

Example (8) instantiates a dependent purpose clause in Chinese, despite involving clausal juxtaposition, because "this is the way Mandarin expresses the conceptual situation associated with purpose clauses in other languages" (Cristofaro 2003: 2).

Nevertheless, the relation between clausal juxtaposition and subordination is still taken to work at a semantic or functional level only, but not at a syntactic level. Consider for this purpose Cristofaro's (2005: 506) own position, also in Cristofaro (2003: 98): "[B]egging Zhang-san is the goal of the person's kneeling, but the relevant clause does not show any of the morpho-syntactic hallmarks that differentiate purpose clauses from main clauses in other languages." It is the purpose of this paper to show that clausal juxtaposition can equally be used as a mechanism of syntactic (rather than just semantic) subordination. To do so, in the next section I introduce the phenomenon in question in Otomi.

1.3. Clausal juxtaposition in Otomi

Otomi is a Mesoamerican language family of the Oto-Manguean stock which comprises about six to seven different languages. All Otomi languages have juxtaposed dependent clauses which may express purpose as in (9a) or may serve as clausal complements of specific matrix verbs, as in (9b).⁴ The examples I used in this paper to illustrate the phenomenon

4 The standard Otomi orthography I use here deviates from the IPA as follows: <x> /ʎ/; <ɸ> /pʰ/; <ɰ> /kʰ/; <y> /j/; <x> /ʃ/; <tx> /tʃ/; <n̄> /n/; <ɸ̄> /f/; <a> /a/; <ɔ̄> /ɔ/; <ɛ̄> /ɛ/; <ɔ̄> /ɔ/; <ɰ̄> /ɰ/; Umlaut indicates a nasal vowel (<ā> /ā/; <ǖ> /ū/; etc.). Otomi languages have three tones: ascending ǎ; high ȳ; and low, which is not represented. Abbreviations: ◊ non-concatenative morpheme; = clitic; _ lexicalized morpheme; AS adjusted stem; ACC accusative; ADVP adverbial phrase; ANTIP antipassive; B bound shape; C complementizer; DAT dative; DEF definite; DEL delimitative; DIM diminutive; DU dual; EMPH emphatic; EXCL exclusive; F free shape; FOC focus; IDEN identifier; IMM immediate; IMPF imperfect; INCL inclusive; INT intensive; intr. intransitive; IRR irrealis; IS impersonal stem; lab. labile verb; LOC.ACT locative actualizer; LOC.P locative preposition; MIDD middle; NEG negation; OBJ object; P (aspectual) particle; PL plural; POSS possessive; PPERF plu-perfect; PERF perfect; PL plural; PRES present; PROG progressive; PST past; PURP purposive; QUOT quotative; R realis; REL relative; SG singular; SS secondary stem; ST stative; TAM tense/aspect/mood; tr. transitive.

2. The purposive juxtaposed construction in Otomi

2.1. Syndetic and asyndetic purpose clauses

Otomi is a Mesoamerican language family spoken in Central Mexico which pertains to the Oto-Pamean subgroup of Oto-Manguean. The analysis I present here is based on first-hand textual data from the Otomi language spoken in the village of San Ildefonso Tultepec (SIT), in the State of Querétaro, but the phenomenon under study is a family-trait, and earlier descriptions of the phenomenon have been provided for other Otomi languages, e.g. for Mezquital Otomi in Lanier (1968) and for Mexico State Otomi in Bartholomew (1973). Details of distribution within the family remain poorly understood, however. Typologically, Otomi languages are nominative-accusative head-marking languages with complex synthetic morphology and extensive use of clitics.

A purpose construction like in English *I went downtown to see the movies* is defined by Cristofaro (2005: 506) as “[...] one encoding a particular relation between events. This relation is such that one of the linked events (the one coded by the main clause, or the main event) is performed with the goal of obtaining the realization of another one (the one coded by the purpose clause, or the dependent event)” (cf. also Gast and Schäfer this volume for a definition of purpose clauses). In SIT Otomi, purpose semantics may be encoded by means of two types of subordinating constructions, which in many situations are freely interchangeable (see Section 3 for a number of restrictions):

- One construction which involves a main clause and a dependent finite clause expressing purpose introduced by the conjunction *pá* – a loan from Spanish *para* – as shown in (10).

(10) *hǒ, dá = hǔd-i [pá dá = k’ó’i’-i]*
 yes 1.PST=sit-F PURP 1.PST=look.at[3OBJ]-F
 ‘Yes, I sat down to look at it.’ (Txt)

- Another construction which consists of a main clause and a juxtaposed dependent clause that follows it. I will refer to this construction as the “juxtaposed construction”. In this construction, the dependent clause is taken to be “finite” in a traditional sense, as it is marked for TAM and for subject agreement, but as it will be shown in Section 3.3, the juxtaposed clause is also structurally deranked, and should be considered as less finite than the dependent clause in (10). An example is given in (11) from the same text as (10).

(11) *dá = ’bat’-i [dá = hand-a = nǒ = r fǒ’to]*
 1.PST=turn.around-F 1.PST=look.at[3OBJ]-B=DEF.SG=SG bulb
 ‘I turned around to look at the bulb.’ (lit. ‘I turned I looked at the bulb.’)
 [18/10-11]

When considered in isolation, the finite juxtaposed clause in (10a) could be taken at a surface level at least to be just like the finite clause in (10b), the only difference being that there is no subordinator. In this sense, I will also refer to this type of clause as the “asyndetic” dependent clause, whereas the first one will be called “syndetic”. I show in the next sections, however, that these two examples instantiate two very different subordinate constructions.

2.2. Prosodic properties of purpose adjunct clauses

One fundamental difference between syndetic and asyndetic purpose clauses concerns their prosodic structures. The dependent clause in (11) above is embedded in the overall phonological structure of the main clause. This motivates the prosodic realization of the dependent clause within the intonational contour of the main clause (Nespor & Vogel 1986). Consequently, no pause – indicated by the symbol “#” – can occur at the juncture between the clauses. This may be seen in the contrast between (12a) and (12b); the latter example not allowing the purposive interpretation of (12a).

(12) a. *ntǒnses ya (dá =)’bǎ = hé = ’pu [dá = ntǒndǎ = hé] #*
 then P 1.PST=stand.AS=PL.EXCL=there 1.PST=have.a.look=PL.EXCL
 ‘We then stood up there to have a look.’ (Txt)

b. *ntǒnses ya (dá =)’bǎ = hé = ’pu # [dá = ntǒndǎ = hé] #*
 then P 1.PST=stand.AS=PL.EXCL=there 1.PST=have.a.look=PL.EXCL
 *Intended reading: ‘We then stood up there to have a look.’
 Meaning conveyed: ‘We then stood up there (and) had a look.’

While the pause is not possible at clausal juncture, it may freely occur in the middle of either clause, as shown in (13), where the speaker is mentally searching for phrasal and lexical options within the dependent clause.

(13) *mú = gá n-dí = pa [n-dí = ’óx-i ... #*
 DEF=ISG IMPF-1.PRES.R=go IMPF-1.PRES.R=overnight-F
ásta ká ngǔ mú ma ... #
 up.to P.LOC.3POSS house DEF 1POSS

nó ma kómpá 'né má n'á ntxúúsi] #
 DEF.SG 1POSS friend and other one girl
 'I used to go to spend the night ... at the house of my ... of a friend of mine and another girl's.' (Txt)

In contrast to the juxtaposed construction in (11), (12a) and (13), the syndetic purpose clause headed by *pá* in (10) forms an independent phonological phrase, and a pause may freely occur at the juncture between the clauses – in particular, after the main clause, as shown in (14). This is also true of all other syndetic dependent clauses in Otomi, as in example (15), a complement clause headed by the complementizer *ge*.

(14) *má n'á_gi bá = pēng-i (#)*
 other one_time 3.HITHER.PST=return-F

[pá bi = ñūng-a = má n'á_gi] #
 PURP 3.PST=have.a.meal.AS-B=other one_time
 'He returned again to eat again.' (Txt)

(15) *kábu ya ga = xi-p = ká ar txítá (#)*
 after.all P 1.PRES.IRR=say-3DAT=ISG.EMPH SG deity
[ge hín = dá tññ- 'i] #
 C NEG=1.PST find.AS-2OBJ-F
 'After all, I will tell God that I didn't find you.' (Txt)

The fact that juxtaposed dependent clauses fall under a single intonation contour together with the main clause indicates that the juxtaposed construction is a complex clause. Using prosodic criteria to identify complex clauses is common (for example, Givón 2009: Chapter 4), and may ultimately be motivated by an iconic principle involving conceptual distance and formal marking, as proposed in Haiman (1985).

In contrast, if clausal juxtaposition is used to encode paratactic relations, a pause may freely occur at the juncture between the clauses. This may be seen in (16), which illustrates a chain of clauses standing in a paratactic relation to each other, thus expressing a sequence of events.

(16) *[bi = n-kūh = wí] #*
 3.PST=MIDD-chase.AS=DU
 'It (the snake) chased her.'

[bi = mñh-i] #
 3.PST=catch[3OBJ]-F
 'caught her,'
[bi = xó't'-e] #
 3.PST=coil.around[3OBJ]-F
 '(and) coiled around her.' (Txt)

Accordingly, a juxtaposed clause in Otomi discourse may either be subordinated to another clause – as in (11) – or it may be paratactically linked to a previous clause – as in (17). Using juxtaposition necessarily implies that the type of relation that a given clause bears to a preceding clause is not explicitly marked, prompting the hearer to rely on other cues, a crucial one being intonation. For example, in the textual string in (17), we have two different instances of juxtaposed clauses:

- (17) a. *[bi = gú^hk-a = nó = r káhá*
 3.PST=SS/extract[3OBJ]-B=DEF.SG=SG box
há mí = kú = r bojó] #
 where 3.IMPF.R=for.a.mass.to.be.located=SG money
 'They got the box out where the money was.'
- b. *[bí = med-e] #*
 3.PST=SS/tell[3OBJ]-F
 'He counted it (the money).'
- c. [*'né bi = hē^hk-a = máđē = thó*
 and 3.PST=divide[3OBJ]-B=half=DEL
 'And he divided it up in order to give it to him.' {2ln; 61–64}
 d. *bi = 'úmm-b-i] #*
 3.PST=give.to.3.AS-3DAT-F
 'And he divided it up in order to give it to him.' {2ln; 61–64}

The juxtaposed clause in (17b) is paratactically linked to (17a). The clause in (17d) is an asyndetic dependent clause expressing purpose and is embedded within (17c). If the speaker had wanted to present the event involving the counting of the money in (17b) as a goal event, he would have included the clause *bi mede* 'he counted it' within the same intonational phrase as (17a). However, as (17a) is phonologically heavy, opting for a syndetic purpose clause would be a better option, like in (18).

- (18) a. *bi = gũ^hk-a = nŃ = r* *káhá*
 3.PST=SS/get.out[3OBJ]-B=DEF.SG=SG box
ha mí = kũ = r *bojŃ (#)*
 where 3.IMPF.R=for.a.mass.to.be.located=SG money
 b. [*pá bi = med-e*] #
 PURP 3.PST=SS/tell[3OBJ]-F

'They got the box out where the money was to count it.'

Likewise, the juxtaposed clause *bi úmbi* 'he gave it to him' in (17d) is construed as a purpose clause by the hearer because of its prosodic embedding. A prosodic structure as in (19) would have rendered a paratactic alignment.

- (19) [*né bi = hé^hk-a = máǎ = thŃ*] # [*bi = úm-b-i*] #
 and 3.PST=divide[3OBJ]-B=half=DEL 3.PST=give.to.3.AS-3DAT-F
 'And he divided it up (and) he gave it to him.'

These facts seem to imply that two sentences considered in isolation could in principle be taken to be ambiguous at a surface level if prosody were not taken into account. Such a point of view, nevertheless, may risk oversimplifying the facts by taking for granted that the internal structure of these apparently ambiguous sentences is the same. In the rest of the paper, I aim to show that the dependent juxtaposed clause in Otomi in (17d) is a deranked structure which is functionally equivalent to other, more typical, deranked structures like infinitives or gerundives in other languages. This clause is different from infinitives and gerundives in that it is marked for TAM and subject agreement, and in that it emerged from the truncation of a full-fledged clause and has never had the external syntax of an NP. While the case is made with data from Otomi, a similar syntactic behaviour is expected in other languages where clausal juxtaposition is used to encode semantic dependency.

Note that, while a purposive interpretation is the most common reading rendered by the juxtaposed adjunct clause in (17d), a simultaneous manner interpretation is at times possible, as for example in (20).⁶

- (20) a. *mí = húh = pu* [*mí = kǎxt' -i*]
 3.IMPF.R=sit.AS=there 3.IMPF.R=message[3OBJ]-F
 'He was sitting there giving it (his finger) a message.' (Txt)

⁶ I want to thank Eva Schultze-Berndt for making me aware of this phenomenon.

- b. *i ya nŃ = r* *ndŃ*
 and P DEF.SG=SG man
sígi bi = 'bá = nu [*bí = tŃ 'm-i*] = 'na
 AUX 3.PST=stand.AS=there 3.THITHER=wait[3OBJ]-F=QUOT
 'And they say that the man kept on standing there waiting for her.' (Txt)

In examples like (20), the preferred reading is one of simultaneity of two events. Such a reading is obtained when the main clause expresses an aspectually imperfective event. The imperfective aspect conveys a stable state of affairs which takes place without a significant change in the dynamism. In such cases, the clause does not favour an interpretation of the event as being performed with the goal of realizing the dependent event. There are also cases with an ambiguous reading between purpose and manner, as for example in (21).

- (21) *dá = yŃ = hé = pu* *ká tai*
 1.PST=walk=PL.EXCL=there LOC.P market
 [*dá = nix-Ńnda = hé*
 1.PSD=have.a.look[3OBJ]=PL.EXCL
té má Ø gǎ 'thŃ Ø = ja = 'pu
 something other REL all 3.PRES.R=exist=there
 'We walked around the market to have/having a look at all that was there.' (Txt)

The relative freedom in the interpretation of the type of adverbial semantics expressed by the dependent clause is a consequence of clausal juxtaposition as a linking mechanism, given that there is no overt marker to favour one interpretation over another. In this respect, the semantic relation between the main and the dependent clause very much relies on the pragmatic interpretation between the two events involved based on conversational implicatures.⁷

⁷ On the other hand, with syndetic constructions, having an overt marker to signal the introduction of a dependent clause does not always entail obtaining a single interpretation because conjunctions are often polysemous.

3. Morphosyntactic properties of the juxtaposed construction

The construction illustrated in Section 2 shows a high degree of syntactic integration between the main and the dependent juxtaposed clause. This can be seen not only in prosodic terms, but also in other properties which will be presented in this section.

3.1. Marking of TAM

In Otomi languages, TAM is encoded by paradigms of verbal clitics which also cross-reference the person of the subject. Table 1 surveys common tenses and moods in SIT Otomi.⁸

Table 1. Common TAM markers in SIT Otomi.

Mood	Tense/Aspect	1 st p.	2 nd p.	3 rd p. ⁸
Realis	Present	<i>dí</i>	<i>gí</i>	(∅)
	Imperfect	<i>ndí</i>	<i>ngí</i>	<i>mí</i>
	Past	<i>dá</i>	<i>gá</i>	<i>bi</i>
	Hither Past			<i>bá</i>
	Perfect	<i>xá</i>	<i>xá</i>	<i>xi</i>
	Plu-perfect			<i>xki</i>
Irrealis	Present	<i>ga</i>	<i>gi</i>	<i>da</i>
	Immediative	<i>ma = ga</i>	<i>ma = gi</i>	<i>ma = da</i>
	Imperfect	<i>nga</i>	<i>ngi</i>	<i>nda</i>
	Perfect	<i>xka</i>	<i>xki</i>	<i>xta</i>

Examples of the present realis are given in (22) for all persons:

- (22) a. *dí = tsú = gá*
 1.PRES.R=get.scared=1EMPH
 'I am scared.' (Txt)
- b. *ja = gí = 'búh = kwa*
 LOC.ACT=2.PRES.R=live.AS=here
 'You're indeed living here.' (Txt)

⁸ Intransitive verbs that encode their subject argument morphologically as an object inflect for TAM values by means of third person markers.

- c. *zóng-a = má* *n'á gi*
 [3.PRES.R]weep.AS-B=other one_time
 'She's weeping again.' (Txt)

The dependent clause in the juxtaposed construction has TAM encoding restrictions (cf. also Gast and Schäfer this volume for a discussion of the relationship between mood or facticity and purposive clauses). It may appear in the irrealis mood, as shown in (23), or otherwise it copies the TAM-values of the main clause, as illustrated in (24) with present and perfect forms of the realis.

- (23) *bi = ma* *nó = r* *'béhñö*
 3.PST=SS/go DEF.SG=SG woman
 [*ma = dá = hó = r* *déhe = 'na*]
 IMM=3.PRES.IRR=bring[3OBJ]=SG water=QUOT
 'The woman went to bring some water.' (Txt)

- (24) a. PRES=[PRES]
pe tó 'be = dí *pa* [*dí = zéngwa*]
 but still=1.PRES.R go 1.PRES.R=visit[3OBJ]
 'But I'm still going there to visit her.' (Txt)
- b. PERF=[PERF]
ya nó = r *'béhñö* *ya xi = náng-i* *n i' 'ho*
 P DEF.SG.3POSS=SG woman P 3.PERF.R=get.up-F early
 [*xi = n-ká'n-i*]
 3.PERF.R=ANTIP-grind-F
 'His wife has already got up early to grind corn.' (Txt)

The choice of TAM-value is semantically motivated when the dependent clause expresses purpose. When the irrealis is used, as in (23), the purpose event is presented as a potential event that has not yet happened, from the perspective of the temporal reference-point established in the main event. In contrast, when the TAM is copied, as in the examples in (24), the goal event in the purpose clause is regarded as factual, that is, it is conceived of as an attained goal—as having actually occurred after the main event. Consider example (25). Here, the act of sitting was performed prior to the act of seeing a given woman, and with the objective of obtaining the realization of that seeing event. The use of the past tense in the juxtaposed

dependent clause moreover indicates that the man indeed saw the woman in question.⁹

- (25) pst=[pst]
niŋse bi = h·ñ·ñih = pu [bi = h·y·and-i]
 then 3.PST=⟨SS⟩sit.AS=there 3.PST=⟨SS⟩see[3OBJ]-F
 'He then sat down to see her (and he saw her).' (Txt)

If no irrealis tense is used – as in (23) – the copying of the TAM-features of the main clause is obligatory. This can be seen in (26), when compared to the ungrammaticality of the examples in (27).

- (26) impf=[impf]
mí = pa [mí = k'ön-i] = 'na
 3.IMPF.R=go 3.IMPF.R=have.a.look.out-F=QUOT
 'It's said that they were going there to have a quick look (and he had a look).' (Txt)

- (27) a. *IMPF≠[PST]
 **mí = pa [bi = k'ön-i] = 'na*
 3.IMPF.R=go 3.PST=have.a.look.out-F=QUOT
 Intended reading: like (26)
 b. *PST≠[IMPF]
 **bi = ma [mí = k'ön-i] = 'na*
 3.PST=SS/go 3.IMPF.R=have.a.look.out-F=QUOT
 Intended reading: like (26)

Similarly, purpose clauses headed by the conjunction *pá* can also appear in the irrealis mood if the goal event is seen as potential, as for example in (28), or in any of the tenses of the realis mood if it is regarded as factual, like in (29).

⁹ By using the irrealis the speaker does not commit him/herself to the factuality of the goal event, i.e. in the interpretation of (22), there is no semantic entailment that the woman did not bring the water, i.e. she could or could not have brought it.

- (28) *mí = hōx-a = r xóni kó déhe*
 3.IMPF.R=take.A[3OBJ]-B=SG jug with water
 [*pá ga = xut = hé nú = ya 'bada-dōni*]
 PURP 1.PRES.IRR=wash.up[3OBJ]=PL.EXCL DEF=PL vase-flower
 'She was taking a jug with water so that we could wash up the vases.'
 (Txt)
- (29) *bi = 'y'am-b-a = bí = r sēi*
 3.PST=SS+extract.liquid.A-3DAT-B=3DAT=SG agave.sap
 [*pá bí = zī*]
 PURP 3.PST=SS/ingest[3OBJ]
 'They took some *pulque* from her in order to drink it (and they drank it).'¹⁰ (Txt)

However, in contrast to the juxtaposed dependent clause, the TAM-features of the purpose clause headed by *pá* are not copied from the main clause, as is the case in any other dependent clause headed by a conjunction. This may be seen in the examples of (30) when compared with the ungrammaticality of (27) above.

- (30) a. IMPF≠[PST]
kó ya hé'mi mí = jō mí = njanu
 with PL paper 3.IMPF.R=put.together[3OBJ] 3.IMPF.R=be.like.that
 [*pá bí = kwá'h-i*]
 PURP 3.PST=put.away[3OBJ]-F
 'They put it together with papers like this to put it away.' (Txt)
- b. PST≠[IMPF]
bi = ma bá = tsý = r
 3.PST=SS/go 3.HITHER.PST=teach.AS.3POSS[3OBJ]=SG
 [*pá mí = m- 'úh = wí*]
 grandmother PURP 3.IMPF.R=MIDD-live.AS=DU
 'They went to reach their grandmother to live with her.' (Txt)

¹⁰ *Pulque* is a Spanish word for a fermented drink made from maguey sap, which is a large kind of agave growing in Central Mexico.

3.1.1. Verbs in the main clause and the TAM-features of the juxtaposed clause

When the asyndetic dependent clause functions as a clausal adjunct, there is no lexical restriction as to which verb may occur in the main clause – provided, of course, that it is semantically compatible with the event described in the second clause. Therefore, we may have verbs in the main clause expressing transfer such as *ʔini* ‘give’ (cf. [31a]), or change of state as *ʔini* ‘spread’ (cf. [31b]).

- (31) a. *bi = ʔim-b-i* [bi = za]
 3.PST=give.to.3.AS-3DAT-F 3.PST=SS/bite[3OBJ]
 ‘He gave it to him to eat.’ (Ttxt)
- b. *da = ʔin-i* [da = ʔy>ót'-i]
 3.PRES.IRR=spread[3OBJ]-F PRES.IRR=<SS>get.dry[3OBJ]-F
 ‘One spreads it (the corn) for it to dry.’ (Ttxt)

Nevertheless, the most common main predicates in the construction are motion verbs. This is typologically expected, as pointed out by Cristofaro (2005), because the participant moving in the main event is identified with the actor of the goal event, and this identification is favoured by the cognitive presumption that people move with the volition and intention of doing other actions. The most frequent motion verbs entering in the juxtaposition construction of SIT Otomi are given in Table 2, together with a number of common verbs expressing change of posture:

Table 2. Verbs in the main clause with purpose asyndetic clause

Motion:	Generic	<i>pa</i> <i>ʔehē</i> <i>pēngi</i>	‘go’ ‘come’ ‘return’
	Deictic	<i>tsopho</i> <i>tsomi</i>	‘arrive here’ ‘arrive there’
	Boundary-crossing	<i>pōni</i> <i>kaʔiʔi</i>	‘exit’ ‘get in’
Postural:		<i>nāngi</i> <i>hiādi</i> <i>bāi</i> <i>ñēgi</i> etc.	‘get up’ ‘sit’ ‘be standing’ ‘turn around to look’ etc.

3.1.2. Juxtaposed clauses in complement function

As mentioned in Section 1.3, the juxtaposed dependent clause is also used in Otomi as a grammatical means to encode the clausal complements of specific matrix verbs (mainly desiderative, manipulative, modal, phasal, and a few others). In this way, the construction serves as a “complementation strategy” in the sense of Dixon (1995, 2006). Two examples are given in (32) with the matrix verbs *ne* ‘want’ and *ndʔi* ‘start’.

- (32) a. *ya hín = dá ne* [dá = bʔh = pʔ]
 P NEG=1.PST want[3OBJ] 1.PST=live.AS=there
 ‘I didn’t want to live there.’ (Ttxt)
- b. *bi = ndʔi-i* [bi = ñʔm-i]
 3.PST=start-F 3.PST=thunder-F
 ‘It started thundering.’ (Ttxt)

The matrix verbs most often found in this construction are given in Table 3. When functioning as a complement, the juxtaposed clause has the same syntactic properties as when it functions as an adjunct, but it shows further restrictions concerning the selection of TAM-values. These restrictions are indicated in the table. The verbs in Table 3 are also grouped according to their semantic classes. Note that the construction is typologically coherent with the proposal made by Cristofaro (2003), as the matrix verbs selecting a juxtaposed dependent clause are those which are expected to select a deranked or less complex clausal complement from a typological point of view.

Table 3. Matrix verbs requiring a juxtaposed clausal complement.

	Only copy of TAM	IRR possible in imperfect and negative; otherwise a copy	Only IRR
Emotion:	<i>tsʔi</i> (intr.) <i>kóstá</i> (dat)	‘be/get scared to do X’ ‘be difficult for one to do X’	✓ ✓
Cognition:	<i>pōdi</i> (tr.) <i>dōhni</i> (intr.)	‘know how to do X’ ‘be/get lazy at doing X’	✓
Desiderative:	<i>jéhni</i> (intr.) <i>ne</i> (tr.) <i>gústá</i> (intr.)	‘set at doing X’ ‘want to do X’ ‘like to do X’	✓ ✓ ✓

	Only copy of TAM	IRR possible in imperfect and negative; otherwise a copy	Only IRR
Manipulation:	<i>já</i> (dtr.) 'make/force sb. to do X'	✓	Only IRR
	<i>'bɛfɪ</i> (tr.) 'command sb. to do X'	✓?	IRR
	<i>'o'ʰe</i> (tr.) 'obey sb. to do X'	✓	
Causative:	<i>hɛgɪ</i> (tr.) 'allow/let sb. do X'	✓	
Modal:	<i>fás'i</i> (tr.) 'help sb. do X'	✓	
	<i>tsá</i> (tr.) 'try to do X'	✓	
	<i>tsɔ</i> (intr.) 'be possible to do X'		✓
Phasal:	<i>jáhni</i> (intr.) 'be fully engaged in doing X'	✓	
	<i>sigt</i> (intr.) 'continue doing X'	✓	
	<i>thége</i> (intr.) 'affect completely by doing X'	✓	
	<i>tége</i> (tr.) 'finish/consume sth. by doing X'	✓	
	<i>fíádi</i> (intr.) 'begin doing X'	✓	
	<i>ndúí</i> (intr.) 'start doing X'	✓	
	<i>dí'ni</i> (tr.) 'start doing X'	✓	
	<i>hɛgɪ</i> (tr.) 'stop doing X'	✓	
	<i>jwáadi</i> (lab.) 'finish doing action X'	✓	
	<i>jwá'ti</i> (lab.) 'finish carrying out event X'	✓	

3.2. Negation

Asyndetic dependent clauses cannot be morphosyntactically negated. When the main clause is negated, the scope of negation comprises the dependent clause. This can be seen in (33) with a purpose clause.¹¹

¹¹ In example (32), the fact that the goal event does not take place does not necessarily preclude the use of the past in the dependent clause. When the past is used, the goal event is presented as a presupposed expected goal. Using a tense of the irrealis mood, as in (i), entails either that the subject is not a drinker or that the speaker discards the idea that the act of drinking was the real reason why he didn't go.

(i) *hím=bi ma [da n=tsi-the]*
 NEG=3.PST SS/go 3.PRES.IRR SS=ingest-water-AS
 'He didn't go to have a drink.'

- (33) *hím=bi ma [bá=tsi-the=yá]=na,*
 NEG=3.PST SS/go 3.HITHER.PST=ingest-water.AS=P=QUOT
bi=ma=yá tá Níai
 3.PST=SS/go=P up.to Amealco
 'They say he didn't go to have a drink, he went to Amealco.' (Txt)

In example (33), the actor of the main event performs the action of going – although the main clause is negated – but not with the objective of having a drink, but of visiting a nearby town called *Amealco*. The marking of negation in (33), the procliticized morpheme *hím-*, is syntactically associated with the main clause but has scope over the event in the dependent clause. Placing the negative marker in the juxtaposed clause renders the sentence ungrammatical, as shown in (34).

- (34) **bi=ma [hím=bá tsi-the]*
 3.PST=SS/go NEG=3.HITHER.PST ingest-water.AS
 Intended reading: 'He went away in order not to have a drink.'

In the intended reading of (34), the main event is performed in order to avoid the happening of the goal event. For such cases, speakers of SIT Otomi use a negated purpose clause headed by *pá*, as in (35):

- (35) a. *ga=tí'k=hé*
 1.PRES.IRR=cut.off.plant[3OBJ]=PL.EXCL
[pá hím=da yá]
 PURP NEG=PRES.IRR rot[3OBJ]
 'We cut it off so that it doesn't rot.' (Ttxt)
- b. *bi=zé'm=ya do*
 3.PST=SS/cover.up[3OBJ]=PL stone
[pá hím=bi dím-b-a=bí ya mbóhɔ]
 PURP NEG=3.PST SS/find-3DAT-B=3DAT PL mestizo
 'They cover it (the money) up with stones so that the mestizos wouldn't find it on them.' (Ttxt)

When the dependent clause functions as a complement clause, the tight structural relation between the main predicate and its clausal complement allows the main predicate and the dependent event to be negated as a whole, as for example in (36).

- (36) a. *ya hín = dí ne [ga 'yó = thó]*
 P NEG=1.PRES.R want[3OBJ] 1.PRES.IRR walk=DEL
 'I don't want to walk.' (lit. 'I don't want it I will walk.') (Txt)
- b. *ya hím = bi hɔyeg-i [bi = ma]*
 P NEG=3.PST <SS>let[3OBJ]-F 3.PST=SS/go
 'She didn't let him go.' (lit. 'She didn't let him he went.') (Txt)

3.3. Structural complexity and finiteness

SIT Otomi is a language with flexible SVO order. A matrix clause has up to six syntactic positions available to the left of the VP before the subject NP (for more details see Palancar 2009). These positions include: TAM markers, a number of topicalized constituents and indefinite pronouns, the marking of negation, a number of topicalized constituents like adverbials and other argumental Ss. Example (37a) illustrates the use of the intensive clitic *xa* and the adverb *ya*. Example (37b) shows a topicalized adverbial, and (37c) a topicalized NP.

- (37) a. *pórke nó = r sèi ya xa = xi = zá t' - i*
 because DEF.SG=SG agave.sap P INT=PERF=SS/thicken[3OBJ]-F
řá-ñ-ho
 [PRES]INT-ST-be.good[3OBJ]
 'Because the *pulque* has already thickened alright.' (Txt)
- b. *mámáde dá = ma = hé ngé'pu hai*
 yesterday 1.PST=SS/go.PL=PL.EXCL there down
 'Yesterday we went down there.' (Txt)
- c. *ya jö n-dí = jö = hé*
 PL mushroom IMPF-1.PRES.R=collect[3OBJ]=PL.EXCL
há ya t'óho
 LOC.P PL mountain
 'We used to collect the mushrooms in the wild.' (Txt)

In contrast to matrix clauses like those in (37), the juxtaposed dependent clause can only be marked for TAM and subject agreement. This suggests that it is a truncated clause with the reduced syntax of an I node, as represented in Figure 1.

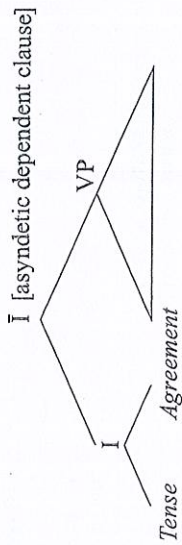


Figure 1. Structure of the asyndetic dependent clause.

The structural poverty of the Otomi asyndetic dependent clause was already mentioned in Section 3.2, where the restriction concerning negation was pointed out. It may be further illustrated with (38) and (39) by comparing the ungrammatical (b)-examples with the felicitous examples of matrix clauses in (37) above.

- (38) a. *ga = ho 's = é*
 1.PRES.IRR=boil[3OBJ].AS=PL.EXCL
[dá = dō xi = ñ-ho]
 PRES.IRR=get.cooked[3OBJ] PERF.R=ST-be.good[3OBJ]
 'We boil it so that it cooks alright.' (Txt)
- b. **ga = ho 's = é*
 1.PRES.IRR=boil[3OBJ].A=PL.EXCL
[ya xa = dá = dō xi = ñ-ho]
 P INT=PRES.IRR=get.cooked[3OBJ] PERF.R=ST-be.good[3OBJ]
 Intended reading: like (38a)
- (39) a. *mí = kad-i [mí = hóm-b = yá bojö]*
 3.IMPF.R=deceive[3OBJ]-F 3.IMPF.R=take.AS-3DAT=PL.3POSS money
 'He used to deceive them in order to rob them of their money.' (Txt)
- b. **mí = kad-i [yá bojö mí = hóm-b-i]*
 3.IMPF.R=deceive[3OBJ]-F PL.3POSS money 3.IMPF.R=take.AS-3DAT-F
 Intended reading: like (39a)

In clear contrast to juxtaposed dependent clauses, syndetic dependent clauses in Otomi are full CPs that may have constituents filling syntactic positions to the left of the predicate, just like a matrix clause. Their structure is represented in Figure 2.

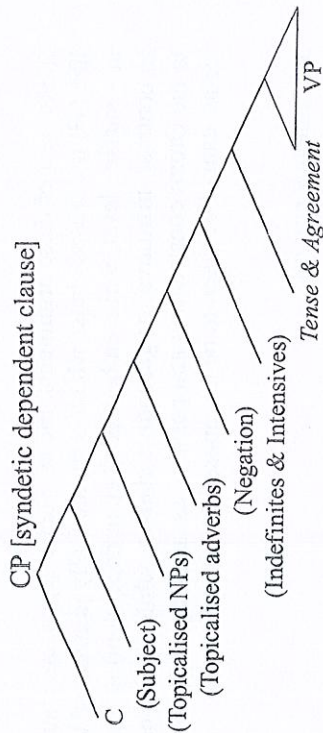


Figure 2. Structure of the syndetic dependent clause.

The structure of CPs is illustrated in (40) with fronted adverbials in a syndetic purpose clause headed by *pá* and in (41) with a fronted NP in a complement clause headed by the complementizer *ge* (which may be elided).¹²

- (40) a. *ntõnse da = g'úng-a = ká*
 then 3.PRES.IRR=SS/grind.A[3OBJ]-B=LOC.P flat.stone
jümi
 [pá ya da = k'é't'-i]
 PURP P 3.PRES.IRR=grind.com[3OBJ]-F

'So one grinds it on the flat stone so that now one may crumble it.'
 (Txt)

- ¹² In English, the complementizer can be elided if it functions as object, like in (ii).
 (ii) a. *I think [that you saw someone]*
 b. *I think [you saw someone]*

The situation in (ii) raises the question of whether (iia) is the same structure as (iib) or whether they should be treated as two different types, one syndetic and the other asyndetic. A closer inspection of the behaviour of these clauses in relation to the *coordinate structure constraint* discussed by Ross (1967) reveals that they indeed are two different constructs. This is illustrated in (iii) and (iv), from (Comrie 2008: 8). For example, while both the structures in (ii) and (iv), *ordinated*, (iib) is more embedded, because both its object and subject can be extracted for relativization, as shown in (iii) and (ivb), respectively. With type (iia), only the object can be extracted.

- (iii) a. *The person that I think [that you saw ____]*
 b. *The person that I think [you saw ____]*
 (iv) a. **The person that I think [that ____ saw you]*
 b. *The person that I think [____ saw you]*

- b. *ya ma = ga = k'õ's = é* *ma bóí ...*
 P IMM=1.PRES.IRR=have.a.look.AS[3OBJ]=PL.EXCL 1POSS cattle
 [pá xti ndé ga = mínts = é
 PURP in.the.afternoon 1.PRES.IRR=gather.AS[3OBJ]=PL.EXCL
má n'á_gí
 other one_time

'We'd go and have a look at the cattle ... so that in the afternoon we'd gather it again.' (Txt)

- (41) *nó = r* *'b'éhñö pöd-i*
 DEF.SG=SG woman [3.PRES.R]=know[3OBJ]-F
 [(ge) *nó = r* *bojõ bi = bē* *nó = r* *xi'tá*
 C DEF.SG=SG money 3.PST=SS/steal[3OBJ] DEF.SG=SG old.man
 'The woman knows that the man stole the money.'
 (lit. '... that the money the man stole.')

The reduced syntax of juxtaposed dependent clauses strongly suggests that these clauses are less finite than syndetic clauses. Alternatively, the facts open up the possibility for an analysis to treat them as non-finite clauses. As Otomi languages do NOT have non-finite verb forms, the juxtaposed dependent clause serves as the functional equivalent of deranked structures such as converbs or infinitives in other languages. However, an asyndetic dependent clause also shows a number of differences from gerunds or infinitives which may also justify its treatment as a different structure. The juxtaposed clause in Otomi is marked for TAM, so that assignment of this semantic value has a correlation in the configuration of the clause, especially when it functions as an adjunct (cf. [23] and [24]). It is also marked for subject agreement.¹³ Furthermore, the clause is not embedded within the main clause with the external syntax of an NP, which is otherwise common with gerunds across languages, and to a certain extent with infinitives too. Moreover, when it is used as a complement clause, it does not occupy the

¹³ In this respect, it could be argued that agreement with subject is necessary by the morphology because clitics are cumulative for TAM and subject cross-reference. However, in other instances of the grammar (mainly in patientive intransitive verbs), there is neutralization of subject in favour of a third person clitic, which works as an exponent of an empty nominative subject. In principle, nothing would preclude the asyndetic clause to receive a dummy person clitic of third person if agreement wasn't an active head in the structure.

structural slot of an argument, as for example in Turkish in (42), (adapted from Kornfilt 2007: 311).

- (42) *biz* [*snav-ı* *geç-ti-k*] *san-ıyor-uz*
 we test-ACC pass-PST-1PL believe-PRES.PROG-1PL
 'We believe we passed the test.'

In (42), the dependent clause bearing both tense and agreement is embedded in the O position, as Turkish is an SOV language, and this indicates that the dependent clause is not juxtaped. Since SIT Otomi is SVO, it could be argued that an asyndetic clause such as (43) functioning as object argument could in principle also occupy the position of O. However, there is some evidence against such an analysis. Object Ss can be topicalized, as shown in (41) above, or can be clefted as in (44). This can never happen with asyndetic dependent clauses as shown by the ungrammaticality of (45). Likewise, a dependent clause functioning as subject cannot occur before the verb, compare (46a) with (46b).

- (43) *bi = ne* [*da = ma*]
 3.PST=want[3OBJ] 3.PRES.IRR=SS/go
 'He wanted to go.'
- (44) *gó gé = r* *bojō nó bi = bē = 'ä*
 FOC IDEN.REL=SG money REL 3.PST.R=SS/steal[3OBJ]=3
 'It was the money what he stole.'
- (45) a. **[da = ma]* *bi = ne*
 3.PRES.IRR=SS/go 3.PST=want[3OBJ]
 Intended reading: 'To go, he wanted.'
- b. **gó gé* [*da = ma*] *nó bi = ne = 'ä*
 FOC IDEN.REL 3.PRES.IRR=SS/go REL 3.PST=want[3OBJ]=3SG
 Intended reading: 'It was to go what he wanted.'
- (46) a. *mí = tsá = [ga* *ma]*
 3.IMPF.R=be.possible=1.PRES.IRR SS/go
 'I could go.' (lit. 'It was possible that I would go.')
- b. **[ga = ma]* *mí = tsá*
 1.PRES.IRR=SS/go 3.IMPF.R=be.possible
 Intended reading: 'That I would go was possible.'

In the light of these arguments, the asyndetic dependent clause in examples like (43) is a tensed clause which is syntactically juxtaped to the main one at a surface level, while embedded (hence dependent) at a structural level. In contrast, infinitival or gerundive clauses, while functionally equivalent to the Otomi construction and possibly as similarly embedded in structural terms, cannot be taken to be juxtaped in the same sense.

3.4. Inherent control

The notion of "inherent control" in Stiebels (2007) proves useful for an understanding of another important syntactic property of the juxtaped construction (cf. also Schmidtke-Bode this volume on control in purpose clauses). According to Stiebels, this type of control is found when the predicate of a main clause requires the predicate of the dependent clause to have a co-referential argument, independently of the syntactic realization of the subordinate clause.¹⁴ In the juxtaped construction, the restriction is as follows: The subject argument of the juxtaped clause must be co-referential with one of the arguments of the main clause. For example, as in typical purpose clauses, both the main and the goal event can share the same actor, encoded as subject (SUB) in both clauses, as in (47).¹⁵

- (47) sub=[sub]
 a. *'né gá: = hōx = é*
 and 1.PRES.IRR=take.AS[3OBJ]=PL.EXCL
 [*gá: = pā = hé = 'yá*]
 1.PRES.IRR=sel[3OBJ]=PL.EXCL=P
 'And we take it to sell it.' (Txt)
- b. *bí: = 'bat' -i* [*bí: = k'ó't' -i*]
 3.PST=turn.around-F 3.PST=look.at[3OBJ]-F
 'He turned around to look at it.'

¹⁴ The co-referential argument most commonly functions as a subject of the dependent clause. "Structural control" is similar to "inherent control", but in the former case the dependent clause is non-finite.

¹⁵ The referent of the subject of the dependent clause may be semantically included in the referent of the subject of the main clause, as in (v).

(v) *náng-i* [*ga = mē*]
 [IMPERATIVE]get.up-F 1.PRES.IRR=SS/go.DU
 'Get up so that we may go.' (Txt)

The referent of the subject within the purpose clause may function as direct (DO) or indirect object (IO) of the main clause, as illustrated in (48a) and (48b), respectively:

(48) a. DO=[SUB]

ya gá thó bi = gü-i, gá thó nú = ya jö'i
 P all 3.PST=SS/expel[3OBJ]-F all DEF=PL person
 [*bí_i = 'bá-yá pu njamú*]
 3.PST=stand.B=INT-far.away that.way

'They expel all, all the people so that they would stand far away like that.' (Txt)

b. IO=[SUB]

'né ... ga = 'úm-p_i = hé n'á txí-tui
 and 1.PRES.IRR=give.to.3.AS-3DAT=PL.EXCL one DIM-bit
 [*dá_i = h'ñ-ó's -i*]
 3.PRES.IRR=SStake[3OBJ]-F

'And we give them a little bit to take home.' (Txt)

The juxtaped construction CANNOT however be used when the subject of the dependent clause is not co-referential with an argument of the predicate of the main clause, as shown by the ungrammaticality of (49) (with a purposive interpretation):

(49) a. **dá_{Pi} = náng-i [gí = h'ñ-ú-d-i]*

1.PST=get.up-F 2.PRES.IRR=<SS>sit.down-F

Intended reading: 'I stood up for you to sit down.'

b. **dá_{Pi} = 'eng-a = r*

1.PST=throw.A[3OBJ]-B=SG ball *pé'ó'tá [gí = m'ñ-i]*
 2.PRES.IRR=grasp[3OBJ]-F

Intended reading: 'I threw the ball for you to pick it up.'

In order to convey such readings, SIT Otomi uses alternative syndetic dependent clauses, for example in (50a-b) (with a purposive reading), clause or in (50c) (with a complement clause).

(50) a. SUB≠[SUB]

dá_i = náng-i dá_i = h'ny'ó'k-w-d = r
 3.PRES.IRR=get.up-F 3.PRES.IRR=<SS>make-3DAT-B.3POSS=SG
txí'txi [pá dá_i = zí]
 baby's.bottle PURP 3.PRES.IRR=SS/ingest[3OBJ]
 'She gets up to prepare him (the baby) his bottle so that he drinks it.' (Txt)

b. DO≠[SUB]

dá = s'x-k-a = gí ká = r dó'tór
 3.PRES.IRR=IS/take.animate-IOBJ-B=IOBJ LOC.P=SG doctor

[pá dá_i = 'óthe = gí]
 PURP 3.PRES.IRR=heal=IOBJ

'They take me to the doctor for him to heal me.' (Txt)

c. IO≠[SUB]

dá = 'á'h-p_i-a = bí_i [(ge) gí = má]
 1.PST=ask.for.A-3DAT-B=3DAT C 2.PRES.IRR=SS/go
 'I asked him that you would go.'

Such constructions are control-neutral, that is, they can also be used when there is argument co-reference between the clauses, as for example in (51).

(51) SUBJ=[SUBJ]

i asi dá_i = já = pu [pá gá_i = p'á]
 and so 1.PRES.R=do[3OBJ]=like.that PURP 1.PRES.IRR=sel[3OBJ]
 'And I do it so to sell them.' (Txt)

3.5. Summary of properties

A summary of the grammatical properties exhibited by dependent clauses in the juxtaped construction of Otomi is provided in Table 4. These properties are compared to those of dependent clauses headed by conjunctions. The asymmetries surveyed in Table 4 indicate that the asyndetic dependent clause in Otomi is a deranked syntactic structure.

Table 4. Properties of the two types of dependent clauses.

	Asyndetic dependent clause	Syndetic dependent clause
Intonation	One single unit with the main clause	Independent unit
TAM	Dependent and restricted	Independent
Negation	Cannot be negated	Can be negated
Complexity	Reduced structure: an \bar{I}	Full structure: a CP
Control	Inherent: Co-referential argument is realized as subject	Control-neutral

4. Syntactic integration

The juxtaposed construction at times surfaces as a grammatical structure which shows a high degree of syntactic integration between the main and the dependent clause. This phenomenon may be seen in two domains: (i) in the marking of a plural subject; and (ii) in a special morphological adjustment I call 'compaction' that verbs in the main clause may undergo.

4.1. Plural marking

In Otomi, a clause with a first or second person plural subject is obligatorily marked by means of a number enclitic. This enclitic is commonly hosted by the verbal predicate, as in (52a-c).

- (52) a. *ga = pã = hē*
 1.PRES.IRR=sell[3OBJ]=PL.EXCL IND.SG=SG DIM-female.turkey
 'We sell a turkey.'
- b. *ga = hoñ = hē*
 1.PRES.IRR=look.for[3OBJ].AS=PL.EXCL
tó \emptyset *da = max-k = hē = 'né*
 someone REL 3.PRES.IRR=SS/help.AS-IOBJ=PL.EXCL=also
 'We look for whom may help us.'

The number enclitic may alternatively be hosted by an NP having a plural possessor that is co-referential with the subject, as in (53):

- (53) *komo hī = 'ó thó*
 as NEG=[3.PRES]not.exist_DEL
té \emptyset *ga_i = tsí^h-p-a = [ma_i hmé = hē]_{NP}*
 something REL 1.PRES.IRR=ingest-3DAT-B=IPOSS tortilla=PL.EXCL
 'As there wasn't anything with which we could eat our tortillas.' (Txt)
 (lit. '... with which I could eat our tortillas.')

This coding pattern also applies to many instances of the juxtaposed construction, as shown in (54), where plural marking appears in each clause.

- (54) a. *'né dá = pēnk = hē = 'yá* [*dá = né = hē*]
 and 1.PST=return.AS=PL.EXCL=P 1.PST=dance.AS=PL.EXCL
 'And we returned to dance.' (Txt)
- b. *dá = jwá = hē* [*dá = né = hē*]
 1.PST=finish.AS[3OBJ]=PL.EXCL 1.PST=dance.AS=PL.EXCL
 'We finished dancing.' (Txt)

However, when there is subject co-reference between the main and the dependent clause in the juxtaposed construction, the plural enclitic can alternatively be hosted by the juxtaposed clause ONLY, as illustrated in (55a-b) for purpose clauses and in (56) for a complement clause.

- (55) a. *dá = náng-i* [*dá = né = hē*]
 1.PST=get.up-F 1.PST=dance.AS=PL.EXCL
 'We got up to dance.' (Txt)
 (lit. 'I got up we danced.')
- b. *ga = pēng-i*
 1.PRES.IRR=return-F
 [*ga = pēt = hē*]
 1.PRES.IRR=smash.AS[3OBJ]=PL.EXCL PL tortilla=P
 'We return to make the tortillas.' (Txt)
 (lit. 'I'd return we'd make tortillas.')

- (56) *kwándo dí = ne*
 when 1.PRES.R=want[3OBJ]
 [*ga = hō^h-a = n'á ma tráhé = hē*]
 1.PRES.IRR=make-B=one IPOSS dress=PL.EXCL
 'When we want to make our dresses.' (Txt)
 (lit. 'When I want I'd make one (of) our dress.')

In the juxtaposed construction, it is the main clause that drops grammatical marking; not the dependent one. If the enclitic were elided in the second clause, the structure would be interpreted as a sequence of two paratactically aligned matrix clauses, as in example (57). Note that the purpose reading is not possible.

- (57) [dá = nánk = hē] (#) [dá = nē-i]
 1.PST=get.up.AS=PL.EXCL 1.PST=dance-F
 a. 'We got up (and) I danced.'
 b. *'We got up to dance.'

4.2. "Compaction" of the main verb

Verbs in Otomi languages commonly show a morphological contrast conditioned by the verb's sensitivity to its syntactic environment. This contrast is known as "compaction". When a verb shows compaction it may appear in two inflectional shapes: a "free shape" (F-shape), illustrated in (58), and a "bound shape" (B-shape), shown in (59). The free shape is a pausal form which is used before a clausal and/or a prosodic boundary. The bound shape is a cliticized form, which is employed when the verb occurs intracausally (for more details see Palancar 2004, 2011).

(58) F-shape

- a. [bi = yút -i] #
 3.PST=SS/get.in-F
 'He got in.' (Txt)
 b. [nē bá = 'ěh-ē] #
 and 3.HITHER.PST=come-F
 'She came back.' (Txt)

(59) B-shape

- a. [bi = yút -a = ('na = r ndé)]_{NP} #
 3.PST=SS/get.in-B=IND.SG=SG man
 'A man got in.' (Txt)
 (lit. 'Got in a man.')
- b. [nē ga = 'ě = (má n 'á_gí)]_{ADVP} #
 and 1.PRES.IRR=come.B=other one_time
 'And she came again.' (Txt)

In the juxtaposed construction, as shown in (60), the verb in the main clause appears in F-shape because it occurs at the clausal boundary before the dependent clause. This behaviour is expected of all verbs occurring at clausal boundaries, and it may be seen as well in juxtaposed clauses which are paratactically linked, as in (61).

- (60) a. dá = dót -i [dá = gó'm-i]
 3.PRES.IRR=SS/tie[3OBJ]-F 3.PRES.IRR=SS/cover[3OBJ]-F
 'They'd tie it to cover it up.' (Txt)
 b. ya kwándo gi = jwád-i [gi = tsi]
 P when 2.PRES.IRR=finish[3OBJ]-F 2.PRES.IRR=ingest[3OBJ]
 ri desáyúnó
 2POSS breakfast
 'When you'd finish eating your breakfast.' (Txt)

- (61) [bi = zón-i] # [bi = ñü 'ts -a = nō = r xóni]
 3.PST=SS/arrive.there-F 3.PST=fill[3OBJ]-B=DEF.SG=SG pitcher
 'She arrived and filled her water jug.' (Txt)

F-shapes and B-shapes are NOT interchangeable. When a B-shape is used in the syntactic context where an F-shape is expected, the result is normally ungrammatical, as shown in (62a) (but see the examples in [63] below). The same is true for an F-shape used in the context of a B-shape, as shown in (62b).

- (62) a. B-shape in the context of a F-shape
 *dá = jwá'č't-a = [dá hō'k-i]
 1.PST=finish[3OBJ]-B=1.PST make[3OBJ]-F
 Intended reading: 'I finished doing it.'
- b. F-shape in the context of a B-shape
 *dá = jwá'č't-i gá'thó
 1.PST=finish[3OBJ]-F all
 Intended reading: 'I finished everything.'

Interestingly, despite the ungrammaticality of (62a), a number of verbs in the main clause may also appear in B-shape before the asyndetic dependent clause. When this happens, the verb serves as a host for the TAM clitics of the dependent clause. The phenomenon is illustrated in (63).

- (63) a. $xíá = 'é' = [gá]$ $tsá-i = ri$
 1.PERF.R=come.B=1.PRES.IRR take.animate.AS[3OBJ]-B=2POSS
 jódó] brother.of.man
 'I've come to take your brother.' (T_{xt})
- b. $péro nú = ya ngǔ dé nú = nu$ $Tlátélólkó$
 but DEF=PL house of DEF=there Tlatelolco
 $bi = thé' = [bi]$ $yót' -e]$
 3.PST=affect.completely.B=3.PST collapse.mass-F
 'But the houses in Tlatelolco collapsed completely.' (T_{xt})

The use of compaction in examples like (63) indicates the disappearance of the syntactic boundary between the main and the dependent clause as motivated by a growing degree of syntactic integration. The main verbs that undergo compaction are given in Table 5.

Table 5. Verbs allowing compaction in the juxtaposed construction.

Main verbs	Motion:	' $éhé$	'come'
Matrix verbs	Phasal:	$thége$ (intr.) $tége$ (tr.) $náyi$ (intr.) $jwádi$ (lab.)	'affect completely by doing' 'finish, consume by doing' 'start' 'finish doing'
	Modal:	$tsá$ (intr.)	'be possible'

The modal verb $tsá$ 'be possible' does not have a B-form *per se*. Because of this, this verb does not undergo compaction in the strict sense, but it behaves in a similar fashion as the other main verbs in Table 5 because it hosts the TAM clitic of the dependent clause. This is shown in (64):

- (64) ge $hí = mí$ $tsá = [n-dá]$ $'nón] = 'á$
 IDEN.REL NEG=3.IMPF.R be.possible=IMPF-3.PRES.IRR move.AS=3SG
 'That (his tongue) was what he couldn't move.' (T_{xt})
 (lit. 'That was what it was impossible for him to move.')

Finally, when a verb from Table 5 is used in the B-shape and the clause has a plural subject in the first or second person, the number enclitic can only occur in the dependent clause, as shown in (65).

- (65) $dá = ndú' = [dá]$ $zón = 'nǔ'$
 1.PST=start.B=1.PST weep.AS=PL(INCL)
 'We started crying.'

5. Conclusions

In this paper, I have claimed that clausal juxtaposition is a clause combining mechanism which can be used to convey both hypotactic and paratactic relations between the clauses involved. The case is shown with a construction in Otomi which consists of a main clause and a juxtaposed dependent clause. Such a dependent clause may function as an adjunct of the main predicate or as a clausal complement of specific complement-taking verbs. When it functions as an adjunct, it commonly conveys purposive semantics, though simultaneous readings are also possible. I have also shown that this dependent juxtaposed clause, though finite, is a deranked structure which is functionally and structurally equivalent to other deranked structures like infinitives or gerundives in other languages.

I have pointed out that finding clause juxtaposition being used as a productive clause combining mechanism in a given language does NOT entail that the clauses involved necessarily stand in a paratactic relation to each other. In those cases where a hypotactic relation between the juxtaposed clauses is indicated by the use of unitary prosody, the Otomi case shows that this fact does NOT entail that the dependent clause has a syntactic structure equivalent to that of a matrix clause, for it may covertly present the same morphosyntactic hallmarks that formally characterize dependent clauses in other languages.

Besides, the juxtaposed dependent clause in Otomi is not an instance of coordination because it does not obey the *coordinate structure constraint* of Ross (1967), as elements in its structure can be extracted for questioning, focus or relativization (see example [66]).

- (66) $ǔó$ $gó$ $nó$ $gí = ne$ $[gí = nu \text{ ---}]?$
 someone FOC REL 2.PRES.R=want[3OBJ] 2.PRES.IRR=see[3OBJ]
 'Whom do you want to see?' (lit. 'Who is it that you want to see?')

On the other hand, the juxtaposed construction of Otomi presented in this article is no doubt typologically related to other constructions in the world's languages which also use juxtaposition as a grammatical device to encode dependent inter-clausal relations. However, at least in the better known

cases reported so far the languages in question commonly have predicate or verbal, rather than clausal, juxtaposition. For example, in *bona fide* serial verb constructions, as characterized in Aikhenvald (2006), a sequence of verbs jointly acts as a single predicate in one clause.¹⁶ In other cases, we seem to have a sequence of predicates rather than of verbs, as for example in Tzotzil (Mayan), where there is a serial predicate construction identified in the literature (Haviland 1981, 1993; Gast 1998) which is used to express purpose and consequence.¹⁷ This construction consists of two juxtaposed predicates marked for person, but only the first predicate is marked for aspect while the second one is marked for the so-called “subjunctive” mood, a dependent mood. An example is given in (67), (Haviland 1993: 35, *apud* Gast 1998: 76).

- (67) *ch-i-may* *j-tuch'-Ø-Ø* *i* *tajchuch=e*
 IND/ICP-1ABS-climb IERG-cut-SUBJ-3ABS DET mushroom=CL
 ‘I climbed up to pick the lentinus mushroom.’

The Tzotzil construction represents an intermediate stage between a proper serial verb construction and the clausal juxtaposition construction presented for languages like Otomi, which consists of two clauses with predicates marked for TAM and subject agreement. Such structural similarities suggest the existence of a cline for the emergence of complex predication and auxiliariation, which has been proposed in different forms (for example DeLancey 1991; Givón 1991, 2009, etc.) and has been revisited by Anderson (2005: 4), as shown in (68), (stage [F] is adapted from “univerbated suffix” into “affix”).

- (68) (a) parataxis > (b) hypotaxis > (c) serialisation > (d) light verb >
 (e) auxiliary > (f) affix

¹⁶ Aikhenvald (2006: 1) defines such constructions as follows: “A serial verb construction (SVC) is a sequence of verbs which act together as a single predicate, without any overt marker of coordination, subordination, or syntactic dependency of any other sort. Serial verb constructions describe what is conceptualized as a single event. They are monoclausal; their intonational properties are the same as those of a monoverbal clause, and they have just one tense, aspect, and polarity value. SVCs may also share core and other arguments. Each component of an SVC must be able to occur on its own. Within an SVC, the individual verbs may have same, or different, transitivity values.”

¹⁷ The construction appears not to be productive synchronically, as it displays a number of collocational restrictions.

While the linearity of such a cline has been criticized by Bower (2008) as too simplistic, especially regarding the emergence of light verbs from serialisation, the literature is more in agreement about stage (a) > (b). However, it should be noted that parataxis and hypotaxis represent two fundamental types of clausal relations inherent to all kinds of clause combining devices, but are not grammatical constructions *per se*, like the other stages of the cline. Therefore, they should not be placed at the same level as the other stages in clines such as (68). Moreover, stage (a) > (b) is based on the general identification of clausal juxtaposition with parataxis as an encoder of asyndetic coordination. I hope to have shown in this paper that the three concepts – parataxis and hypotaxis, on the one hand, and clausal juxtaposition, on the other – should be taken to be distinct categories, and in this light, I propose an amendment of the cline in (68) as shown in (69).

- (69) (a) clausal juxtaposition > (b) serialisation > (c) light verb >
 (d) auxiliary > (e) affix

(paratactic > hypotactic)

In the cline in (69), clausal juxtaposition is given its right place as a grammatical device. It is typologically expected that if a language uses juxtaposition to encode hypotactic relations, it will also use it both productively and extensively to express paratactic relations between clauses. When used for this purpose, juxtaposition commonly renders a sequential and/or consequential reading between the events expressed in the clauses involved. It should be borne in mind, however, that such readings are not quintessential to clausal juxtaposition as a grammatical device *per se*, as is often assumed, but only to clausal juxtaposition when it is used to encode parataxis.

The Oto-Manguean case presented here with Otomi raises important questions, both for a general theory of language and for typological methodology, as instances of subordination encoded by clausal juxtaposition in other languages may conceal similarly complex underlying structures which are not apparent from surface phenomena.

Acknowledgements

This paper was written under the auspices of the ESRC/AHRC project RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguean languages”. I dedicate the paper to Judith Aissen as a way to thank her for her teachings. I also want to thank Roberto Zavala, Verónica

Vázquez, Gilles Polian, Paulette Levy and Sonia Cristofaro for their invaluable comments on earlier versions of this article. My sincere gratitude to the editors of this book, Volker Gast and Holger Diessel, for all their insightful observations and for their carefulness with details in editorial matters. I also want to thank the external reviewer of the book who spotted weak areas that needed attention.

References

- Aarts, Bas. 2006. Subordination. In Keith Brown (ed.), *The Encyclopedia of Language and Linguistics Vol. 5* (Second edition), 248–254. Oxford: Elsevier.
- Aikhenvald, Alexandra Y. 2006. Serial verb constructions in typological perspective. In Alexandra Y. Aikhenvald (ed.), *Serial Verb Constructions: A Cross-Linguistic Typology*, 1–68. Cambridge: Cambridge University Press.
- Anderson, Gregory. 2005. *Auxiliary Verb Constructions*. Oxford: Oxford University Press.
- Bartholomew, Doris. 1973. Otomi dependent clauses. In C. Corum, Thomas C. Smith-Stark & A. Weiser (eds.), *You Take the High Note and I'll Take the Low Note*. Chicago Linguistics Society 7, 1–8.
- Bowern, Claire. 2008. The diachrony of complex predicates. *Diachronica* 25 (2): 161–185.
- Comrie, Bernard. 2008. Subordination, coordination: Form, semantics, pragmatics. In Edward Vajda (ed.), *Subordination and coordination strategies in North Asian languages*, 1–16. Amsterdam: John Benjamins.
- Cosme, Christelle. 2006. Clause combining across languages: A corpus-based study of English-French translation shifts. *Languages in Contrast* 6 (1): 71–108.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Cristofaro, Sonia. 2005. Purpose Clauses (Chapter 125). In Martin Haspelmath, Matthew S. Dwyer, Hans-Jörg Bibiko, David Gil, Bernard Comrie & Claudia Schmidt (eds.), *World Atlas of Language Structures*, 506–509. Oxford: Oxford University Press. <http://wals.info/feature/description/125> (accessed 6 June 2012).
- Culicover, Peter & Ray Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28 (2): 195–218.
- Deane, Paul. 1992. Grammar in mind and brain: Explorations in cognitive syntax. Berlin & New York: Mouton de Gruyter.
- DeLancey, Scott. 1991. The origins of verb serialization in Modern Tibetan. *Studies in Language* 15 (1): 1–24.
- Dixon, R. M. W. 1995. Complement clauses and complementation strategies. In F. Palmer (ed.), *Grammar and Meaning: Essays in honor of Sir John Lyons*, 175–220. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 2006. Complementation Clauses and Complementation Strategies in Typological Perspective. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A Cross-Linguistic Typology*, 1–48. Oxford: Oxford University Press.
- Gast, Volker. 1998. On Clause Combining in Tzotzil. Johannes Gutenberg-Universität Mainz, unpublished MA Dissertation.
- Givón, Talmy. 1990. *Syntax: A functional-typological introduction, Vol. II*. Amsterdam: John Benjamins.
- Givón, Talmy. 1991. The evolution of dependent clause morpho-syntax in Biblical Hebrew. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization, Vol. 2: Focus on Types of Grammatical Markers*, 257–310. Amsterdam: John Benjamins.
- Givón, Talmy. 2009. *The Genesis of Syntactic Complexity*. Amsterdam: John Benjamins.
- Haiman, John. 1985. *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.
- Haviland, John. 1981. *Sk'op Soiz'leb: El Tzotzil de San Lorenzo Zinacantan*. Mexico City: Universidad Nacional Autónoma de México.
- Haviland, John. 1993. The syntax of Tzotzil auxiliaries and directionals. *Berkeley Linguistics Society* 19: 35–49.
- Hoeksema, Jacob & Donna Napoli. 1993. Paratactic and Subordinative So. *Journal of Linguistics* 29 (2): 291–314.
- Hyman, Larry. 1971. Consecutivization in Fefé? *Journal of African Languages* 10 (2): 29–43.
- Johannessen, Janne Bondi. 1998. *Coordination* [Oxford Studies in Comparative Syntax]. Oxford: Oxford University Press.
- Kornfilt, Jaklin. 2007. Verbal and nominalized finite clauses in Turkish. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*, 305–332. Oxford: Oxford University Press.
- Lakoff, George. 1986. Frame semantic control of the Coordinate Structure Constraint. *Chicago Linguistics Society* 22 (2): 152–167.
- Lanier, Nancy. 1968. Three structural layers in Mezquital Otomi clauses. *Linguistics* 43 (1): 32–85.
- Li, Charles N. & Sandra A. Thompson. 1973. Serial Verb Constructions in Mandarin Chinese: Coordination or Subordination? *Papers from the Comparative Syntax Festival, Chicago Linguistics Society* 9: 96–113.
- Longacre, Robert E. 1966. Trique clause and sentence: A study in contrast, variation, and distribution. *International Journal of American Linguistics* 32: 242–252.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Volume 2: Complex Constructions*, 42–140. Cambridge: Cambridge University Press.

- Noonan, Michael. 2006. Complement Clauses. In Keith Brown (ed.), *The Encyclopedia of Language and Linguistics* (Second edition). 694–697. Oxford: Elsevier.
- Palancar, Enrique L. 2004. Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70(3): 251–278.
- Palancar, Enrique L. 2009. *Gramática y Textos del Hñotĩhò: Otomí de San Ildefonso Tultepec, Querétaro. Volumen I: Gramática*. Mexico City: Plaza y Valdés.
- Palancar, Enrique L. 2011. Shape conditions in Otomi. Invited paper read at the III Morphology Meeting of the Southeast, University of Surrey, (8 April.)
- Pride, Kitty. 1965. *Chatino Syntax*. (Publications in Linguistics and Related Fields 12.) Norman: University of Oklahoma, Summer Institute of Linguistics.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Ross, John R. 1967. Constraints on variables in syntax. PhD dissertation, MIT.
- Schmerling, Susan. 1975. Asymmetric conjunction and rules of conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics 3: Speech acts*, 211–231. New York: Academic Press.
- Stiebels, Barbara. 2007. Towards a typology of complement control. In Barbara Stiebels (ed.), *ZAS Papers in Linguistics* 47: 1–80.
- Thompson, Sandra A. & Robert E. Longacre. 1985. Adverbial clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Volume 2: Complex Constructions*, 171–234. Cambridge: Cambridge University Press.
- Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Volume 2: Complex Constructions* [2nd Edition], 237–300. Cambridge: Cambridge University Press.
- Wacke, O. 1931. Formenlehre der Ono-Sprache. *Zeitschrift für Eingeborenen-Sprachen* 21: 161–208.

IX. Semántica depictiva sin predicación secundaria: el caso de la lengua otomí*

Enrique L. Palancar
UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

1. Introducción a la lengua

El otomí es una lengua otomange de la familia otopame que se habla en varios estados en el centro de México. La lengua tiene una diversificación dialectal tan considerable que puede resultar pertinente hablar de "lenguas otomíes", tal como lo hace Suárez (1983). Los datos sobre los que se basa mi análisis provienen del otomí hablado en San Ildelfonso Tultepec (STT), un pueblo dentro de la municipalidad de Amealco de Bonfil, en el estado de Querétaro. El fenómeno que se estudia aquí parece ser un rasgo diacrónico de la lengua puesto que aparece en otros dialectos, pero dada la variación interdialectal y algunas de las idiosincrasias del fenómeno mismo, es completamente esperable que existan diferencias estructurales y léxicas significativas.

2. Introducción al objeto de estudio

Por lo general las construcciones de predicación secundaria suelen expresar un tipo de semántica que se ha venido a denominar "depictiva". A grandes rasgos, y siguiendo la propuesta en Schultze-Berndt y Himmelmann (2004), Aissen (2004) y Himmelmann y Schultze-Berndt (2005), en la semántica depictiva típica se indica

* Quiero agradecer los valiosos comentarios a una versión anterior de este trabajo de los participantes del Taller sobre Predicación, que fue impartido en CESAS Sureste por Judith Aissen en agosto del año 2004. Infinitas gracias a los editores, Roberto Zavala y Judith Aissen, por su excelente ojo crítico y su paciencia, y en especial a Judith, por ofrecer luz para guiarse en los intrincados y fascinantes caminos de la predicación. Mil gracias de corazón de nuevo a Roberto Zavala, en calidad de generoso anfitrión en mi estancia de investigación en San Cristóbal de Las Casas, Chiapas, en agosto del año 2005, y también a Verónica Vázquez y a Valentín Peralta, con quienes compartí tan agradable y productiva estancia.

la existencia de una correlación temporal entre un cierto evento que protagoniza un participante y un cierto estado en el que este mismo participante se encuentra durante el proceso del evento referido. Un ejemplo ilustrativo se ilustra en la construcción en 1 en español:

(1) *El hombre comió parado.*

En esta expresión del español se nos informa, por un lado, de que un hombre protagonizó un cierto evento de ingerir comida y, por otro, de que este mismo hombre se encontraba de pie durante ese proceso. En 1 se indica que tanto el evento de comer como la condición de estar parado fueron simultáneos.¹ El ejemplo muestra un caso de semántica depictiva del tipo de estado o condición. Por otro lado, el ejemplo refleja una construcción gramatical especial que se puede llamar “construcción de predicación secundaria depictiva”, entre otras razones porque en una sola cláusula tenemos dos predicados, *v. g.* el predicado *comió*, que sería principal o primario y que refiere al evento, y el predicado *parado*, que es secundario o subsidiario del primario y que refiere al estado.

Según Schultze-Berndt y Himmelmann (2004), la semántica depictiva típica, además de estado o condición, también incluye etapa de vida (*fui a París de niño*) o rol (*trabaja como cirujano*). Estos autores proponen que es relativamente común que las lenguas que poseen una construcción depictiva genuina la empleen asimismo para expresar otras áreas semánticas de ámbito más adverbial, como son la cuantificación (*lo hicieron juntos; todos; en parejas; etc.*), la semejanza o comparación (*trabaja como un esclavo*), etc.² Al respecto, se propone una jerarquía implicacional en la que la construcción comienza primero por expresar estado o condición y se extiende después a otros ámbitos. Aunque las lenguas varían en el grado de amplitud semántico-funcional que conceden a su construcción de predicación secundaria, la propuesta predice que si esta construcción existe, ha de expresar al menos estado o condición.³

¹ La simultaneidad se aplica al menos en lo que respecta a la parte crucial de la dinámica del evento. Si comer implica una secuencia de acciones convencionales llevadas a cabo una detrás de otra durante un tiempo convencionalmente establecido, para que la proposición en 1 fuera cierta bajo criterios de verdad el hombre tendría que haber estado parado durante un tiempo sustancial en este proceso y no solamente, por ejemplo, al final del acto de comer.

² Los ejemplos del español dados en este párrafo pretenden ilustrar nada más aspectos semánticos, no se trata de ilustrar la construcción depictiva del español.

³ En virtud de esta jerarquía —basada en un amplio corpus de lenguas indoeuropeas y aus-

En este artículo muestro que el otomí de *srr* tiene una construcción de predicación secundaria, pero propongo que a pesar de la existencia de esta construcción la lengua no la usa de forma productiva y natural para expresar semántica depictiva. En su lugar, el otomí debe entenderse como una lengua que emplea para este propósito cláusulas independientes en una relación de yuxtaposición.

Antes de presentar la manera en que se codifica gramaticalmente la semántica depictiva en otomí, resulta conveniente primero ver algunos rasgos gramaticales relevantes de la predicación primaria. En la sección 4 se introducirá la construcción marginal de predicación secundaria, para mostrar luego en la sección 5 la forma natural que tiene la lengua de expresar semántica depictiva mediante el uso de cláusulas independientes. Esta sección es la más extensa del trabajo, en ella se caracterizan las expresiones depictivas, se muestra el alcance semántico de este tipo de expresión y se presenta una estructura de predicado complejo que sirve de alternativa léxica para expresar “rol”. Finalmente, en la sección 6 se presentan las conclusiones generales de este trabajo.

3. La predicación primaria en el otomí de San Ildefonso Tultepec

Un predicado primario en otomí se flexiona en las categorías de tiempo, aspecto y modo por medio de un complejo paradigma de proclíticos verbales, que a veces pueden recibir acento y funcionar como palabras funcionales. Estos proclíticos sirven también de manera acumulativa como exponentes morfológicos de concordancia de sujeto. Por esta razón los denomino marcadores de tiempo/aspecto/modo y persona, y se abrevian como TAMYP. En 2 doy ejemplos de predicados verbales:⁴

tralianas—, resulta un heurístico conveniente en el trabajo descriptivo comenzar primero por la exploración de expresión gramatical que reciben los estados en esta construcción.

⁴ En la representación ortográfica de la lengua se emplea la convención sugerida en Hekking (2002), con dos innovaciones: el uso de <ř> para /r/ y el uso de <ʎ> para representar una aspiración de índole fonética. Desviaciones del AFI en esta convención son las siguientes: Para consonantes, <ɔ> para /ʔ/ y eyección; <ñ> para /ɲ/; <x> para /ʃ/; <tx> para /tʃ/; <f> para /pʰ/; <ɟ> para /bʰ/; <y> para /j/; <r> para /r/; y <ř> para /r/. Para vocales, el diacrítico <˘> indica nasalización; <a> para /a/; <ɛ> para /ɛ/; <ɛ˘> para /ɛ/; <i> para /i/; <ɔ˘> para /ɔ/. El otomí es una lengua tonal, pero en este trabajo sólo se representa el tono en los proclíticos o palabras funcionales, en donde el contraste es siempre entre tono alto y bajo. El tono alto se presenta con el diacrítico <˘>, mientras que el tono bajo, siendo un tono base, no recibe representación.

- (2) a. *nuya txi jö i bi düü*
nu=ya txi-jö i bi düü (bi=<d>ü)
 DEF=PL DIM-gente 3.PSD=<TNP>morir
 'La pobre gente se murió.' {Txt}
- b. *nu ma kosa di usa pa ga hāt'i*
nu ma kosa (dī=usa=Ø) pa (ga=hāt'-Ø-i)
 DEF(PL) IPOS cosa I.PRES=usar=3OBJ para I.IRR=engañar-3OBJ-L
 'Estas cosas que uso para enganarla (a la gente).' {Ttxt}

En el ejemplo 2a tenemos el predicado primario *bi düü*, "se murió", en donde aparece el verbo intransitivo *tüü*, "morirse", flexionado en TAMYP mediante el proclítico *bi*. Este marcador indica tanto tiempo pasado y aspecto completivo como concordancia con una tercera persona (no hay distinción de número en tercera persona). En 2b tenemos dos predicados primarios, ambos flexionados en primera persona mediante dos proclíticos distintos: *dī*, que se emplea para tiempo presente, y *ga* que indica irrealis.

La lengua tiene en principio tres clases verbales distinguibles, entre otros diagnósticos, a razón de los clíticos de TAMYP que reciben. Se trata de los "verbos activos" (V-act), los "verbos incoativos" (V-inc) y los "verbos estativos" (V-est). Como referencia, en 3 ofrezco un paradigma de los clíticos de TAMYP más relevantes que aparecen en el otomí de STT, en relación con cada una de estas clases:

⁵ Las abreviaturas son: "-" ausencia de frontera morfológica gramatical en fuente; "-" frontera morfológica en fuente; " " elemento lexicalizado en fuente; en la glosa indica ausencia de frontera morfológica léxica; "=" clítico; A: forma ajustada verbal; ACT: actualizador; AP: antipasiva; CIT: citativo; D: forma dependiente verbal; DAT: dativo; DEF: definido; DEL: delimitativo; DEM: demostrativo; DEP: dependiente; DIM: diminutivo; ERG: ergativo; EST: estativo; EX: exclusivo; FI: flexión de verbo intransitivo "nasal"; FOC: foco; IMPF: imperfecto; IN: inclusivo; INF: infinitivo; INT: intensivo; INVOL: auxiliar de involuación; IRR: (presente de) irrealis; L: forma libre verbal; M: medio; NEG: negativo; OBJ: objeto; P: partícula; PF: perfecto; P-PF: plu-perfecto; PL: plural; P.LOC: preposición locativa; PN: predicación nominal; POS: posesivo; PRES: presente; PSD: pasado; REL: relativo; SC: sufijo categorial; SG: singular; SUB: subjuntivo; {TXT}: *corpus*; TNP: tema no presente; VE: verbo estativo.

	Modo <i>realis</i>						Modo <i>irrealis</i>		
	Presente		Imperfecto		Pasado		Presente		
	V-act	V-inc	V-act	V-inc	V-act	V-inc	V-act	V-inc	V-est
1 ^a	<i>dī</i>	<i>xtá</i>	<i>n-dī</i>	-	<i>dá</i>	<i>xtá</i>	-	<i>ga</i>	-
2 ^a	<i>gí</i>	<i>xká</i>	<i>n-gí</i>	-	<i>gá</i>	<i>xká</i>	-	<i>gi</i>	-
3 ^a	Ø	<i>xi</i>	<i>mí</i>	<i>már</i>	<i>bi</i>	<i>xki</i>	-	<i>da</i>	<i>dar</i>

Por otro lado, el verbo *hāt'i*, "engañar", en el predicado primario *ga hāt'i*, "lo engañaré", en 2b aparece en una forma morfológica especial a la que denomino "forma libre". Esta forma se emplea cuando el verbo aparece ante frontera de frase fonológica; es decir, se trata de formas pausales que se pueden enunciar antes de una pausa prosódica. No obstante, el uso de una forma libre está también condicionado sintácticamente ya que el verbo aparece en esta forma cuando le sigue otro predicado, tal como se ilustra en 4:

- (4) *pa da k'et'i da gāt'a hme*
pa (da=k'et'-Ø-i) (da=<g>gāt'-Ø-a)=hme
 para 3.IRR=moler-3OBJ-L 3.IRR=<TNP>echar_tortillas-3OBJ-D=tortilla
 'Para que lo muela (el trigo) y eche tortillas.' {Ttxt}

Tal como se ve en 4, el verbo *k'et'i*, "moler", aparece en forma libre porque le sigue el predicado *da gāt'a*, "echará tortillas". En contraste, en este último predicado el verbo *kāt'i*, "echar tortillas", no aparece en forma libre sino en otra forma que he denominado "forma dependiente" en Palancar (2004). En principio, una forma dependiente se emplea cuando al verbo no le sigue otro predicado dentro de la misma cláusula (para más detalles véase Palancar 2004).

En los ejemplos de hasta ahora se han mostrado solamente verbos activos. En 5 aparecen ejemplos de dos predicados primarios con un verbo incoativo y otro estativo, respectivamente:

Verbo activo incoativo:

- (5) a. *hinda japa daño porke xi dö xi ñho*
hin=da já-p-a=daño porke (xi=dö) xi=ñ-ho-Ø
 NEG=3.IRR hacer-3 DAT-D=daño porque 3.PF=COCER (3.)PF=EST-ser_bueno-3OBJ

'No hace daño porque está cocido bien.' {Ttxt}

Verbo estativo:

- b. *no'or* 'estivo', *no xi nis'ò'ä*
no = 'ò=r *zəne*, *no* (*xí=n-t's'ò-Ø*) = 'ä
 DEF.SG=DEM.SG.3=SG brujo REL (3.)PF=EST-ser_malo-3OBJ=3SG
 'Ese brujo sí que es malo.' {Txt}

Adicionalmente, en 5 se puede observar que el proclítico *xi* es un marcador de perfecto de tercera persona que se emplea para indicar tiempo presente con estas dos clases de verbos, tal como se puede comprobar en la tabla en 3.

Los verbos estativos en otomí expresan "conceptos de propiedad" a la Dixon (1982) y tienen un patrón de marcación de persona del tipo activo-estativo (Mithun, 1991), en donde se codifica con morfología de objeto a la persona a la que se adscribe la propiedad expresada por el verbo (para más detalles véase Palancar, 2006). Este patrón se puede ver más claro en 6, en donde aparece el clítico de objeto de segunda persona:

- (6) *xi ñh'eis'i'i*
xi=n-ñ-h'eis'i=i
 (3.)PF=EST-ser_alto=2OBJ
 'Eres alto.'

Aparte de los verbos, los cuantificadores también pueden funcionar en otomí como predicados primarios. Cuando lo hacen reciben la misma morfología de TAMYP que los verbos. Un ejemplo de ello aparece en 7a con el numeral *yoho*, "dos", el ejemplo en 7b muestra el uso del numeral en modificación:

- Predicación:
 (7) a. *ne de gepu xa bin jobya koma mi yoho*
ne de_g'e'pu xa=bi=n jobya koma (mi=yoho)
 Y después INT=3.PSD=FI contentarse porque 3.IMPF=dos
 'Y ya después se contentaron porque eran dos.' {Txt}

Modificación:

- b. *dá tsixa nu yoho ma böisi ya mi txi dōta*
dá=tsix-Ø-a=(nu) *yoho ma böisi*
 1.PSD=llevar_animado.A-3OBJ-D=DEF(PL) dos 1POS niño
 Ø *ya mi=tsi-dōta*
 REL P 3.IMPF=DIM-ser_grande
 'Me llevé a mis dos hijos que ya estaban grandecitos.' {Txt}

Finalmente, los sustantivos se pueden asimismo emplear como predicados primarios en una cláusula. A diferencia de los verbos y de los cuantificadores, los sustantivos reciben una flexión de TAMYP especial cuando aparecen en predicación. En el predicado de esta construcción de predicación nominal aparecen: 1) uno de estos marcadores especiales de TAMYP; 2) un determinante de número (singular o plural), y 3) el sustantivo predicativo en cuestión.⁶ La construcción se muestra en 8a, con el uso de dos predicados nominales, uno en singular y el otro en plural. En contraste, en 8b se ofrece el uso de uno de los mismos sustantivos con función referencial:

Predicativo:

- (8) a. *hingi már ge'ra sino ge má ya bē*
hingi (má=r ge'ra) sino ge (má=ya bē)
 NEG 3.IMPF.PN=SG guerra sino REL.FOC 3.IMPF.PN=PL ladron
 'No era la guerra, sino es que eran unos ladrones.' {Txt}

Referencial:

- b. *nu'nu mi ndur ge'ra*
nu'nu mi=ndur=(r ge'ra)
 cuando 3.IMPF=comenzar=SG guerra
 'Cuando comenzó la guerra.' {Txt}

3.1. CRITERIOS DE DEFINICIÓN DE CLÁUSULA

Una cláusula en otomí contiene un solo predicado primario. En este sentido, los ejemplos en 9 son ejemplos de cláusulas:

- (9) a. *nuga ya dá tinga ma bojō*
[nu=ga ya (dú=ting-Ø-a)=ma bojō]
 DEF=1 P 1.PSD=encontrar.A-3OBJ-D=1POS dinero
 'Yo ya encontré mi dinero.' {Txt}

⁶ A pesar de que Ecker (1952) ve los marcadores de TAMYP en esta construcción como manifestaciones flexivas de una cópula, en realidad no existe ninguna evidencia basada en criterios morfológicos que justifique tal tratamiento al menos para el otomí moderno.

- b. *pa da zok a i pu ri bojö 'nehe*
 [pa (da = <z> o 'k- 'a = 'i) = pu ri bojö = 'nehe]
 para 3.IRR=<TNP>dejar.A-2DAT=D=2DAT=allí 2POS dinero=también
 'Para que te dejen allí tu dinero también.' {Txt}

En realidad ejemplos como los que aparecen en 9 no resultan problemáticos para la delimitación de frontera de cláusula, ya que cada uno presenta un solo predicado primario. El problema aparece cuando encontramos en el discurso dos o más predicados contiguos, por ejemplo en 10:

- (10) a. *bi boni bi k'öni*
 (bi = on-i) (bi = k'ön-i)
 3.PSD=<TNP>salir-L 3.PSD=ir_a_ver-L
 'Salió a ver.' {Txt}
- b. *hin ö, ya dí zä 'wa dí 'buhu*
 hin ö, ya (dí = zä) = 'wa (dí = 'bu = hu)
 no P 1.PRES=acomodarse=aquí 1.PRES=vivir=PL(.IN)
 'No, yo estoy bien aquí viviendo juntos (tú y yo).' {Txt}
- c. *himí hek' a i ngi köi*
 (hi = mi hek' -a = 'i) (n-gi = kö-i)
 NEG=3.IMPF dejar.A-2OBJ-D=2OBJ IMPF-2.IRR=bajarse-L
 'No te dejaba bajar (del árbol).' {Txt}

Ante ejemplos como éstos en donde concurren dos predicados contiguos surge la pregunta de cuántas cláusulas tenemos en realidad: ¿se trata de una o de dos? Para resolver este dilema, en este trabajo voy a proponer varios criterios para delimitar cláusulas en el otomí de SIT. El primero y más relevante aparece en 11:

- (11) Cada cláusula tiene una sola marca de TAM.

Aplicando este criterio se puede concluir que todos los ejemplos en 10 representan en efecto casos de dos cláusulas. Por ejemplo, en los predicados *bi boni*, 'salió', y *bi k'öni*, 'fue a ver', en 10a se observa el marcador *bi* de TAM en cada uno de ellos, de forma independiente, por lo que de acuerdo con 11 el ejemplo 10a presenta en efecto dos cláusulas. Lo mismo sucede con los predicados en 10b y 10c, pero con diferentes marcadores.

Adicionalmente, se pueden proponer otros dos criterios de marcación que resultan convenientes para identificar cláusulas:

- (12) Con sujeto plural, cada cláusula tiene una sola marca de sujeto plural.
 (13) En polaridad negativa, cada cláusula tiene una sola marca de negación.

El criterio en 12 se muestra en 14:

- (14) *pe nu ihu xa gi 'öhu, ni xi gi tsghu*
 [pe nu = 'i = hu (xa = gi = 'ö = hu)] [ni xi (gi = tsä-Ø = hu)]
 pero DEF=2=PL INT=2.PRES=dormir.A=PL ni 2.PRES=sentir-3OBJ=PL
 'Pero ustedes están bien dormidos y ni cuenta se dan.' {Txt}

En este ejemplo tenemos dos cláusulas coordinadas que tienen un sujeto plural correferencial. Dado que cada cláusula tiene un sujeto plural, en cada una de ellas aparece el morfema de plural = *hu*, asociado esta vez a los predicados primarios.

No obstante, la marca de sujeto plural es facultativa en el predicado primario si ésta aparece en otro lugar dentro de la misma cláusula. Esto puede suceder por ejemplo cuando la cláusula contiene un poseedor interno plural que es correferencial con el sujeto. En circunstancias como éstas, la marca de número puede aparecer solamente en la FN que designa lo poseído. Esto se ilustra en 15:

- (15) *ga isi ma t'ëihu, ga māhā*
 [ga = tsi-Ø (ma t'ëi = hu)] [ga = māh-ā]
 1.IRR=ingerir-3OBJ 1POS atole=PL(.IN) 1.IRR=ir.PL.IN-L
 'Nos tomamos nuestro atole, (y) nos vamos.' {Txt}

En el ejemplo 10b arriba, repetido aquí como 16, la marca de plural = *hu* aparece solamente en el predicado *dí 'buhu*, 'vivimos', pero no así en el predicado *dí za*, 'estoy bien'. Por esta razón, y ya que tenemos dos cláusulas según el criterio en 12, las dos cláusulas en 16 no tienen un sujeto correferencial.

- (16) *hin ö, ya dí zä 'wa dí 'buhu*
 hin ö, [ya (dí = zä) = 'wa] [(dí = 'bu = hu)]
 no P 1.PRES=acomodarse=aquí 1.PRES=vivir=PL(.IN)
 'No, yo estoy bien aquí viviendo juntos (tú y yo).' {Txt}

Por otro lado, el criterio relativo a la negación en 13 se ilustra en 17:

- (17) *hinda zoni, himbi ma*
 [(him = da <z > on-i)]
 NEG=3.IRR <TNP>legar_alli-L NEG=3.PSD <TNP>ir
 'No llegará, no se fue.'

En 16 tenemos de nuevo dos cláusulas contiguas, que en esta ocasión comparten el hecho de que ambas aparecen en la polaridad negativa. A razón del criterio en 13, en sendas cláusulas aparece una marca de negación, que está asociada morfosintácticamente a los proclíticos de TAMYP. En términos semánticos, el criterio en 13 viene a indicar que el ámbito de la negación emplazada sobre un predicado primario no puede alcanzar a otro predicado primario. En otras palabras, la lectura semántica que se obtiene en el ejemplo 17 no es posible con la sola presencia de un marcador de negación, tal como se muestra en 18:

- (18) *hinda zoni, bi ma*
 [(hin = da <z > on-i)] [(bi = <m > a)]
 NEG=3.IRR <TNP>legar_alli-L 3.PSD=<TNP>ir
 'No llegará, se fue.' / * 'Llegará, no se fue.' / * 'No llegará, no se fue.'

Los criterios en 11, 12 y 13 resultan importantes en otomí para distinguir fronteras de cláusula cuando se tienen dos o más predicados contiguos en el discurso. Una vez que estos criterios se han presentado, resulta conveniente ahora presentar argumentos a favor de la existencia de una construcción de predicación secundaria en esta lengua.

4. Predicación secundaria en otomí de San Ildefonso Tultepec

En principio, para el otomí de SIT se puede decir que existe una genuina construcción de predicación secundaria depictiva. Esta construcción se ilustra en el ejemplo en 19:

- (19) *gá ntxaha nithemfo*
 [(gá = n-txah-a)_{pl}] (nithemfo)_{pl2} = tho]
 2.PSD=M-bañarse-L encuerado=DEL
 'Te bañaste encuerado (así no más).'

El ejemplo 19 representa una sola cláusula con dos predicados: uno primario (P1°) *gá ntxaha*, 'te bañaste', y otro secundario (P2°) *nithemfo*, 'encuerado'. A continuación doy evidencia que justifique este análisis aplicando los tres criterios de cláusula que presenté en la sección anterior.

En relación con el criterio en 11, los dos predicados contiguos en 19 reciben exclusivamente una marcación de TAM con el marcador *gá*. Por esta razón estos dos predicados están insertos dentro de una misma cláusula. Asimismo, el predicado *nithemfo*, 'encuerado', no tiene ningún marcador que indique TAM (ni concordancia de persona). Al no recibir marcación de TAM, el predicado *nithemfo* es un predicado no finito.

Los dos predicados en 19 son intransitivos y comparten el mismo argumento. Siguiendo a Schultze-Berndt y Himmelmann (2004), el ejemplo en 19 tiene un controlador *s*, ya que el argumento del predicado depictivo funciona como *s* para con el primario. Cuando este argumento es plural —para el caso de la primera y segunda persona—, la marca de plural recae sobre el predicado primario, tal como se muestra en 20, pero su ámbito referencial alcanza al argumento del predicado secundario:

- (20) *gá ntxahu (n)goxjiitho*
 [(gá = n-txa = hu)_{pl}] ((n)goxji)_{pl2} = tho]
 2.PSD=M-bañarse.A=PL encuerado=DEL
 'Ustedes se bañaron encuerados.'

Según el criterio expuesto en 12, si la construcción en 19 representara dos cláusulas, sería necesario el uso obligatorio de la marcación de plural en ambos predicados, sin embargo esto no es posible, tal como se muestra en 21:

- (21) * *gá ntxahu (n)goxjiithohu*
 * [(gá = n-txa = hu) (n)goxji) = tho = hu]
 2.PSD=M-bañarse.A=PL encuerado=DEL=PL
 Lectura buscada: 'Ustedes se bañaron encuerados.'

Cuando el predicado primario es transitivo, el controlador puede ser *a* u *o*. Un ejemplo en el que la interpretación por defecto es *o* aparece en 22:⁷

⁷ Por alguna razón, la construcción depictiva que ejemplifico no se puede emplear cuando se implica a una primera persona como controlador *o*.

(22) *ha dá hand 'a i nthemfotho*

[*ha* (*dá* = *hand*-²*a* = ¹*i*)_{ppz} (*nthemfo*)_{ppz} = *tho*]
 y entonces 1.PSD=ver-2OBJ-D=2OBJ encuerado=DEL
 'Y es que te vi encuerado (así no más).'

Por otro lado, la negación es un diagnóstico importante para determinar la existencia de una construcción de predicación secundaria. Daré un ejemplo ilustrativo en español antes de mostrar los datos del otomí. Por ejemplo, considérese el ejemplo en 23:

(23) *El pájaro no salió volando.*

En 23, la negación alcanza semánticamente al predicado depictivo (*volando*). Es decir, en este ejemplo no se niega el hecho de que el pájaro saliera sino el hecho de que volara al hacerlo. El fenómeno que ilustra 23 demuestra que el predicado *volando* es lo suficientemente independiente como para que su evento se niegue aislado del evento principal que denota el predicado eventual (Aissen, 2004). De la misma forma sucede en otomí con ejemplos de la construcción depictiva, tal como se muestra en 24:

(24) *hin ö, hindá ma ngoxjiho*
hin ö, [(hin=dá < m > a)]_{ppz} = tho
 no NEG=1.PSD <TNP> it encuerado=DEL
 'No, no me fui encuerado.'

En el ejemplo 24, la marca de la negación recae sintácticamente sobre el predicado primario pero alcanza semánticamente al depictivo (*n)goxji*, "encuerado." Es decir, lo que se niega en 24 no es el hecho de que el hablante fuera a algún lugar, sino el que lo hiciera estando desnudo.

4.1. DISTRIBUCIÓN DE LA CONSTRUCCIÓN DE PREDICACIÓN SECUNDARIA

En la sección anterior he mostrado ejemplos de predicados secundarios con los lexemas *nthemfo* y (*n)goxji*, que son dos sinónimos que significan "encuerado" y que pertenecen a un registro muy informal.⁸ Hasta donde alcanza mi conocimiento en

el momento de escribir este trabajo, *ningún* otro elemento léxico en el otomí de sir puede emplearse como depictivo en esta construcción, ora como verbo, ora como sustantivo. Esto se muestra en la agramaticalidad de ejemplos, como en 25:

(25) *dá isoho nilitunhülgüt'ëi*, etc.
 * [*dá* = *isoh*-*o* *nilitunhülgüt'ëi*/etc.]
 1.PSD=llegar_aquí-L emborracharse/estar_hambriento/(un)celoso/etc.
 Lectura buscada: 'Llegué borracho/hambriento/celoso/etc...'

En resumen, la construcción depictiva o de predicación secundaria en otomí de sir es una construcción insólita ya que no permite la inclusión de otros elementos léxicos. No obstante, y a pesar de la escasa productividad, su existencia misma permite pensar al menos tres posibilidades: a) que la construcción contenga en efecto más miembros (dato que en el momento presente de la investigación se desconoce); b) que haya variedad dialectal respecto de este fenómeno y/o respecto de los elementos léxicos que se permiten en la construcción; y c) que haya variedad individual y los hablantes integren nuevos elementos léxicos con semántica depictiva en la construcción provenientes de otros ámbitos.

Es importante recalcar también que los elementos *nthemfo* y (*n)goxji*, "encuerado", presentan además un comportamiento léxico insólito en la lengua, ya que pueden funcionar sin restricción alguna tanto como verbos o como sustantivos. Esto se muestra en 26, en donde presento tanto su uso en predicación verbal y en nominal:⁹

Predicación verbal:

(26) a. *gi nthemfo/ngoxji*
gi = *nthemfo/ngoxji*
 2.PRES=estar_encuerado
 'Estás encuerado.'

Predicación nominal:

b. *gár nthemfo/ngoxji*
gá = *r* *nthemfo/ngoxji*
 2.PRES.PN=SG (un.)encuerado
 'Eres un encuerado.'

⁹ El elemento *goxji* aparece siempre con una nasal temática en *n-goxji* tanto cuando funciona como verbo como cuando lo hace como sustantivo. La nasal es aparentemente facultativa cuando aparece en posición de depictivo.

⁸ La semántica de estos elementos no implica la completa desnudez del individuo: el estar en ropa interior o en traje de baño en público puede también calificarse como un ejemplo de *nthemfo* o (*n)goxji*. Alternativamente, el verbo incoativo *moge*, "desvestirse/desnudarse", se emplea como una alternativa léxica pragmáticamente neutra para expresar la misma semántica depictiva.

En otomí los lexemas muestran por lo general una pertenencia bastante rígida a una determinada clase de palabra, por lo que el fenómeno de conversión que presentan estos elementos es bastante excepcional. Cuando estos ítems funcionan como predicados secundarios lo más probable es que lo hagan en calidad de sustantivos y no como formas verbales no finitas. Los verbos del otomí carecen de formas no finitas, y resultaría extremadamente raro tipológicamente que éstas fueran precisamente las dos únicas formas de verbales de este tipo.

Por otro lado, en Palancar (2006) he propuesto que el otomí no tiene adjetivos como clase de palabra independiente. Estos elementos léxicos no pueden tampoco tratarse como los únicos adjetivos de la lengua, ya que cuando funcionan atributivamente dentro de una FN —el contexto no marcado por excelencia de un adjetivo— requieren sin embargo de una cláusula de relativo. Por otro lado, los ejemplos en 26 demuestran que estos elementos no toman cópula para funcionar como predicados primarios, sino que se comportan como verbos o como sustantivos.

Surge entonces una pregunta importante: si *nihemfo* y (*n*)*goxji*, “encuerado” son los dos únicos elementos que aparecen en esta construcción de predicación secundaria depictiva, ¿cómo expresa el otomí el amplio rango de la semántica depictiva? El objetivo de lo que resta del trabajo es dar una respuesta satisfactoria a esta pregunta.

5. Expresiones depictivas en otomí de San Idefonso Tultepec

A pesar de la existencia de la construcción de predicación secundaria que se introdujo en la sección anterior, en la lengua otomí “la forma natural” de expresar semántica depictiva se realiza a través de una secuencia de cláusulas en una relación de yuxtaposición. Como esta secuencia de cláusulas no conforma una construcción depictiva propiamente en la lengua, para los propósitos de este trabajo me referiré a ella como “expresión depictiva”. No obstante, es necesario aclarar que esta estructura en cuestión se debe observar como un fenómeno gramatical de seriación de índole amplia en la lengua que no está circunscrito a la semántica depictiva propiamente.

Una expresión depictiva tiene que tener al menos dos cláusulas: a) una “cláusula eventual” (CL-EV) en donde se expresa una eventualidad y b) al menos una “cláusula depictiva” (CL-DEPTV) en donde se indica el estado simultáneo o concomitante en el que se encuentra un participante de la eventualidad predicada en la cláusula eventual, y que va a funcionar en términos sintácticos como una cláusula de adjunto. Ejemplos de expresiones depictivas aparecen en 27:

- (27) a. *xa bi mafi mi hüm'pü'na*
 [xa = bi = maf-i]_{CL-EV} [mi = hüm = 'pü = 'na]_{CL-DEPTV}
 INT=3.PSD=gritar-L 3.IMPF=sentarse.A=allí=CIT
 ‘Dizque gritaba mucho sentado allí.’
 (Lit.: ‘Dizque gritó mucho, estaba sentado allí.’) {Txt}
- b. *ya de 'rato nde ga tsohe ya xi zaje*
 [ya de 'ra'ro nde ga = iso = he]_{CL-EV}
 p de ocho tarde 1.IRR=llegar_aquí.A=1 PL.EX
 [ya xi = < z > a^h-k = he]_{CL-DEPTV}
 p 3.PF=<TNP>cansarse.A-1OBJ=1 PL.EX
 ‘Ya a las ocho de la tarde llegamos cansados.’
 (Lit.: ‘Ya a las ocho de la tarde llegamos, ya estábamos cansados.’) {Txt}
- c. *ja dá tsogi mi 'bq'pütho*
 [ja dá = isog-Ø-i]_{CL-EV} [mi = 'bq' = 'pü = 'tho]_{CL-DEPTV}
 ACT 1.PSD=dejar-3OBJ-L 3.IMPF=estar_parado.A=allí=DEL
 ‘La dejó allí parada no más (como si no se quisiera sentar).’
 (Lit.: ‘La dejó [en algún lugar], estaba parada allí no más.’) {Txt}

En 27 se muestran tres casos de expresión depictiva en la que hay dos cláusulas en cada ejemplo. En 27a, la primera cláusula es la eventual e informa a través del predicado *xa bi mafi*, “gritó mucho”, que un cierto individuo protagonizó la acción de gritar mucho. La segunda cláusula, la depictiva, nos informa que este mismo individuo estuvo sentado durante todo el proceso en que duró la acción de gritar, y esto se hace a través del predicado primario *mi hüm'pü*, “estaba allí sentado”. En 27a tenemos dos verbos monovalentes y las cláusulas tienen un sujeto correferencial. Lo mismo sucede en 27b, en donde el predicado depictivo es un verbo estativo. En 27c, en cambio, el sujeto de la cláusula depictiva *mi 'bq'pütho*, “estaba parada allí no más”, es correferencial con el objeto de la cláusula eventual *ja dá tsogi*, “la dejó (ahí)”. Otro ejemplo del mismo tipo de expresión depictiva aparece en 28:

- (28) *bi dingi ndi 'bem'pü ndi 'öhö*
 [bi = < d > in-g-i]_{CL-EV} [n-di = 'bem = 'pü]_{CL-DEPTV}
 3.PSD=<TNS>encontrar-1OBJ-L IMPF-1.PRES=estar_útrado.SG.A=allí
 [n-di = 'öh-ö]_{CL-DEPTV}
 IMPF-1.PRES=dormir-L
 ‘Me encontraron acostado durmiendo.’
 (Lit.: ‘Me encontraron, estaba acostado allí, dormía.’)

A diferencia de las expresiones depictivas en 27, en el ejemplo 28 tenemos dos cláusulas depictivas en vez de una sola. Tanto 27 como 28 ilustran casos en donde los predicados primarios de las cláusulas depictivas son verbos. Sin embargo, se pueden encontrar asimismo sustantivos en predicación, tal como se ilustra en 29:

- (29) *dá penjí ndár güt'ëi*
 [dá = peng-1]_{CL-EV} [n-dá = r] *güt'ëi*_{CL-DEPTV}
 1.PSD=regresar-1 IMPF-1.PRES.PN=SG (un)celoso
 'Regresé celoso.'
 (Lit.: 'Regresé, yo era un celoso.')

5.1. EVIDENCIA DEL TRATO DE LA EXPRESIÓN DEPICTIVA COMO CLÁUSULAS INDEPENDIENTES

En los ejemplos de la sección anterior he presentado ejemplos de expresiones depictivas en otomí que constan de dos o más cláusulas. En esta sección se ofrece evidencia de que el análisis propuesto es en efecto el más pertinente. La evidencia se basa principalmente en los resultados que se obtienen de la aplicación de los tres diagnósticos que presenté en 11, 12 y 13 en la sección 3.1.

5.1.1. FLEXIÓN DE TAM

Los predicados implicados en expresiones depictivas son siempre predicados finitos, ya que reciben obligatoriamente flexión de TAM. El criterio en 11 es un diagnóstico fundamental para tratarlos como pertenecientes a cláusulas independientes. Al respecto considérese el ejemplo en 30:

- (30) *nu'mü mí nde, mí zohó xítá nítí*
 [nu'mü mí = nde], [mí = <z>oh-ø]_{CL-EV} [xítá = nítí]_{CL-DEPTV}
 cuando 3-IMPF=tarde 3-IMPF=<TNS>llegar_aquí-1 3.P-PF=emborracharse
 'Cuando anochece, llegaba borracho.'
 (Lit.: 'Cuando anochece, llegaba, estaba borracho.')

En 30 tenemos tres cláusulas, ya que cada predicado aparece obligatoriamente con morfología de TAM. Una hipotética estructura en la que el predicado depictivo no recibiera esta marcación resultaría agramatical, tal como se ilustra por ejemplo en la imposibilidad de 31:

- (31) * *mí zohó nítí*
 * [mí = <z>oh] *nítí*
 3-IMPF=<TNP>llegar_aquí emborracharse/estar_borracho
 Lectura buscada: 'Llegaba borracho.'

La agramaticalidad de 31 demuestra que el predicado *nítí*, 'emborracharse/estar borracho', no puede funcionar como elemento predicativo si no recibe flexión de TAM. Esta imposibilidad muestra que este elemento depictivo ha de funcionar como predicado primario en una cláusula independiente.

5.1.2. CONCORDANCIA DE SUJETO PLURAL

Cuando en una expresión depictiva se tiene un sujeto plural correferencial, la marca de plural aparece obligatoriamente en las dos cláusulas, tal como se muestra en 32:

- (32) a. *dá penjí xítá nítíhe*
 [dá = penk = he]_{CL-EV} [xítá = nítí = he]_{CL-DEPTV}
 1.PSD=regresar.A=1 PL.EX 1.PF=emborracharse=1 PL.EX
 'Regresamos borrachos.'
 (Lit.: 'Regresamos, nos hemos emborrachado.')
- b. *dá isa ma ngohü ndá hñihü*
 [dá = isa-ø] *ma ngohü = hñü*_{CL-EV} [n_dá = hñü = hñü]_{CL-DEPTV}
 1.PSD=comer/morder-3 OBJ 1 POS carne=PL(IN) 1-IMP.F.DEP=sentarse.A=PL(IN)
 'Comimos nuestra carne sentados.'
 (Lit.: 'Comimos nuestra carne, estábamos sentados.')

Las dos cláusulas yuxtapuestas en cada uno de los ejemplos en 32 tienen un sujeto plural en correferencia y en cada una de ellas aparece una marca de concordancia de plural de sujeto. Esto es esperable de acuerdo con el criterio en 12. Si la marca apareciera solamente en una de las cláusulas se perdería la lectura de correferencia de sujeto, tal como se muestra en 33:

- (33) a. *dá penjí xítá nítí*
 [dá = penk = he]_{CL-EV} [xítá = nítí]_{CL-DEPTV}
 1.PSD=regresar.A-1 PL.EX 1.PF=emborracharse
 'Regresamos, estando yo borracho.'
 Lectura buscada: * 'Regresamos borrachos.'

b. *dá pengi xta ntihu*

[*dá = peng-i*]_{CL-EV} [*xtá = nti = hu*]_{CL-DEPTV}
 1.PSD=regresar-L 1.PF=emborracharse=PL(IN)
 'Regresé, estando (tu y yo) borrachos.'
 Lectura buscada: * 'Regresamos borrachos.'

Por otro lado, Aissen (2004) menciona que en predicaciones secundarias sin marca de concordancia suele existir ambigüedad respecto de cuál es el controlador cuando se tienen dos participantes distintos. Compárense al respecto los ejemplos en 34 en español:

- (34) a. *el policía me arrestó tomado*
 b. *los policías nos arrestaron tomados*

En los ejemplos en 34 es común interpretar que el controlador de los predicados secundarios *tomado* y *tomados* sea el argumento objeto (o) del predicado principal, *v. g.* los hablantes. No obstante, se puede interpretar de la misma forma que el controlador es A, *v. g.* los policías.

Ambigüedades de este tipo se observan también en otomí cuando los predicados tienen argumentos en tercera persona. No obstante, cuando se trata de una primera o segunda persona la codificación gramatical no permite la ambigüedad, ya que la marca de concordancia de objeto de primera y segunda se tiene que hacer explícita en la cláusula eventiva. Esto se observa en 35, un equivalente semántico en otomí de 34:

- (35) *nor is'ú' tvi bi mije xta ntihe*
 [*no = r is'ú' tvi bi = mⁱ-k_i = he*]_{CL-EV} [*xtá = nti = he*]_{CL-DEPTV}
 DEF.SG=SG autoridad 3.PSD=agarrar.A-1OBJ=PL.EX 1.PF=emborracharse=1PL.EX
 'La policía nos agarró tomados.'
 (Lit.: 'La autoridad nos agarró, nos hemos emborrachado.')

5.1.3. MARCA DE NEGACIÓN

La negación es un diagnóstico importante para mostrar que la expresión depictiva en otomí está en efecto conformada por cláusulas independientes. Cuando en la lengua se quiere negar el estado que denota la cláusula depictiva, la marca de negación ha de recaer morfosintácticamente sobre el predicado primario de esta cláusula, tal como se ilustra en 36:¹⁰

¹⁰ Si la marca de negación recayera sobre ambos predicados, se negaría entonces de forma independiente tanto el evento como el estado que cada uno de ellos expresa.

- (36) a. *bi bonga nor 'behñö himi jolya*
 [(*bi = ong-a*) = *no = r 'behñö*]_{CL-EV} [(*hi = mi jolya*)]_{CL-DEPTV}
 3.PSD=<TNS>salir.A-D=DEF.SG señora NEG=3.IMPF contentarse
 'La señora no salió contenta.'
 (Lit.: 'La señora salió, no estaba contenta.')

- b. *bi zohö hixki nti*
 [(*bi = <z> oh-o*)]_{CL-EV} [(*hi = xki nti*)]_{CL-DEPTV}
 3.PSD=<TNS>llegar_aquí-L NEG=3.P-PF emborracharse
 'No llegó borracho.'
 (Lit.: 'Llegó, no se había emborrachado.')

Por otro lado, si la marca de negación recae sobre el predicado de la cláusula eventiva, se niega solamente el evento y nunca alcanza semánticamente al predicado de la cláusula depictiva. Esto se muestra en 37:

- (37) a. *himi bñi mi jolya*
 [(*him = bi ñ-i*)]_{CL-EV} [(*mi = jolya*)]_{CL-DEPTV}
 NEG=3.PSD <TNS>salir-L 3.IMPF=contentarse
 'No salió, (porque) estaba contenta.'
 Lectura buscada: *'No salió contenta.'
- b. *himi zohö xki nti*
 [(*xki = nti*)]_{CL-DEPTV}
 NEG=3.PSD <TNS>llegar_aquí-L 3.P-PF=emborracharse
 'No llegó, (porque) estaba borracho.'
 Lectura buscada: *'No llegó borracho.'

En esta sección he mostrado que los resultados de aplicar cada uno de los tres criterios de definición de cláusula en la sección 3.1 evidencian que la expresión depictiva en otomí está en efecto compuesta de cláusulas yuxtapuestas. Existe otra prueba auxiliar para determinar este mismo comportamiento. Se trata de la posibilidad de insertar una FN sujeto entre ambos predicados (Aissen, 2004). En otomí esta posibilidad es viable, aunque no común. Un ejemplo se muestra en 38:

- (38) *xa bi ma' nor txi Hulia mi hüm 'pü' na*
 [*xa = bi = ma' = (no = r txi-Hulia)*]_{CL-EV}
 INT=3.PSD=gritar.A-D=DEF.SG DIM-Julían

[mí = hüüm = 'pu = 'na]_{CL-DEFTV}
 3.IMPF=sentarse.A=allí=CIT
 'Dizque hizo un griterío el Julián sentado allí.'

En este ejemplo se puede observar que la FN *nor txi Hulia*, "el Julián(cito)", se inserta detrás del predicado primario dentro de la cláusula eventual. Esto permite que se mantenga la correferencia del sujeto. Si la FN aparece en la cláusula depictiva, se rompe la interpretación con correferencia y el resultado no es una expresión depictiva. Esto se ilustra en 39:

(39) *xa bi mafí, nor txi Hulia mí hüüm 'pu 'na*
 [xa = bi = maf-í], [(no = r txi-Hulia) mí = hüüm = 'pu = 'na]
 INT=3.PSD=gritar-L DEF.SG=SG DIM-Julián 3.IMPF=sentarse.A=allí=CIT
 'Gritó mucho, (y) dizque el Julián estaba sentado allí.'

5.2. OTRAS CARACTERÍSTICAS DE LA EXPRESIÓN DEPICTIVA

En esta sección presento una serie de características adicionales de la expresión depictiva que la relacionan con otras estructuras de la lengua. En la sección 5.2.1 muestro que la cláusula depictiva es una cláusula de adjunto que aparece generalmente en el mismo orden que otras cláusulas del mismo tipo pero que presentan un marcador hipotático. Por otro lado, en la sección 5.2.2 muestro que las cláusulas implicadas en la expresión depictiva comparten una temporalidad que permite que la cláusula depictiva aparezca con morfología de presente, independientemente del TAM que reciba la eventual. Se observa adicionalmente que éste es un fenómeno que acontece frecuentemente en el discurso otomí cuando los eventos narrados comparten cierto grado de cotemporalidad.¹¹

5.2.1. LA CLÁUSULA DEPICTIVA Y LAS CLÁUSULAS DE ADJUNTO

Por lo general, las cláusulas de adjunto en otomí aparecen después de la cláusula matriz o principal. Esto se muestra en 40:

¹¹ Le agradezco mucho a Judith Aissen (c. p.) esta observación.

Final:

(40) a. *el, mpefi pa gi tór bojó pa gi wingi ri bōtsi*
el, [m-pef-i] [pa gi = tō-Ø = r bojó
 eh AP-trabajar-L para 2.IRR=ganar.D-3OBJ=SG dinero
 [pa gi = wing-Ø = ri bōtsi]]
 para 2.IRR=alimentar.D-3OBJ=2POS niño
 'Trabaja para que ganes dinero para alimentar a tus hijos.' {Txt}

Causal:

b. *ya bi gwadi 'na ko bi zonga ga 'tho ya xōñju*
 [ya bi = <g> wad-i = 'na] [ko bi = <z> ong-a = ga 'tho ya xōñju]
 P 3.PSD=<TNP>terminar-L=CIT porque 3.PSD=llagar_ahí.A-D=todo PL hormiga
 '(Y) ya terminó porque llegaron todas las hormigas.' {Txt}

Manera:

c. *ha nugá ngu ne ga hwangi*
 [ha mu = gá][ngu ne ga = hwang-i]
 y_entonces DEF=1SG como querer 1.IRR=caerse_desde_altura-L
 'Y yo, como que me quería caer.' {Txt}

En los ejemplos en 40 se pueden ver varios tipos de cláusulas de adjunto que aparecen después de la matriz. Este orden refleja el patrón general de subordinación en la lengua, salvo el caso de las temporales y condicionales que preceden a la principal. El mismo orden se puede ver también en las cláusulas de relativo, tal como se muestra en 41:

(41) *muya nobio mí 'ñēhē 'ya himí japa nkasō*
 [(nu = ya nobio) [Ø mí = 'ñēh-ē-Ø]]
 DEF=PL novio REL 3.IMPF=tener_animado-L-3OBJ
 'ya 'i = mí ja^h-p-a = nkasō]
 P NEG=3.IMPF hacer-3DAT-D=caso
 'Los novios que tenía ya no le hacían caso.' {Txt}

La cláusula depictiva por lo general sigue también a la eventual en la expresión depictiva. Cuando este orden neutro se invierte, la depictiva se pone en una posición de foco, y en estos casos se tiene que emplear el enclítico =*tho*, un delimitativo que se puede

traducir como "así no más".¹² Si este clítico no se empleara, la predicación resulta algo forzada. El caso se ilustra en 42b en contraste con 42a, que es la estructura natural:

- (42) a. *xki nĩtho, bi zohq*
 [*xki = nĩ = tho*]_{CL-DEPTV} [*bi = <z> øh-ø*]_{CL-EV}
 3.P-PF=emborracharse=DEL 3.PSD=<TNS>llegar_aquí-L
 '¡Llegó borracho (así no más)!'

- b. ? *xki nĩ, bi zohq*
 ? [*xki = nĩ*]_{CL-DEPTV} [*bi = <z> øh-ø*]_{CL-EV}
 3.P-PF=emborracharse 3.PSD=<TNS>llegar_aquí-L
 Lectura natural buscada: 'Llegó borracho.'

Es necesario apuntar que, en contraste con 42a, el empleo del enclítico en el orden neutro es completamente facultativo. Por otro lado, el cambio de orden no es posible cuando la cláusula depictiva aparece en polaridad negativa, tal como se muestra en la agramaticalidad de 43:

- (43) * *hixki nĩtho bi zohq*
 * [*hĩ = xki nĩ = tho*]_{CL-DEPTV} [*bi = <z> øh-ø*]_{CL-EV}
 NEG=3.P-PF emborracharse=DEL 3.PSD=<TNS>llegar_aquí-L
 Lectura buscada: 'No llegó borracho.'

En este trabajo he tomando el orden de cláusulas de la expresión depictiva como un rasgo fundamental que apoye un análisis de la cláusula depictiva como un adjunto insertado dentro de una estructura oracional mayor. Queda encontrar en una investigación futura otros diagnósticos que puedan apoyar este análisis.

5.2.2. LA TEMPORALIDAD EN LA EXPRESIÓN DEPICTIVA

Las cláusulas depictivas en la expresión depictiva en otomí expresan típicamente el estado en el que se encuentra cierto participante a la hora de la realización del evento principal. Este estado es en la mayoría de los casos relevante para el modo en que se realiza el evento designado por la cláusula principal —si alguien está borracho cuando

¹² El hablante expresa cierta contraexpectativa o admiración acerca del evento cuando emplea el enclítico =*tho*. No obstante, su semántica es compleja y difícil de caracterizar. Roberto Zavala (c. p.) me ha sugerido que el enclítico puede ser un tipo de "mirativo".

sale de un lugar, es muy probable que este estado afecte al modo como sale de ese lugar—. No obstante, aparte de poder expresar la manera como se realiza el evento, la expresión depictiva establece una relación de contemporaneidad relevante entre un evento y un estado. Esta relación de temporalidad se puede después entender en un contexto adecuado como causal o temporal.

En otomí la relación semántica que expresa la cláusula depictiva como adjunto se recupera en contexto, pero si se quiere hacer explícito el componente causal o temporal se puede entonces recurrir a una expresión alternativa en la que se emplea una cláusula con un subordinador explícito, tal como se muestra en 44, con ejemplos de adjuntos causales en un tipo de expresión que ya no es depictiva:

Causal:

- (44) a. *himbi zohq ko xki nĩ*
 [(*him = bi <z> øh-ø*)] [*ko*] (*xki = nĩ*)
 NEG=3.PSD <TNS>llegar_aquí-L porque 3.P-PF=emborracharse
 'No llegó porque estaba borracho.'

Temporal:

- b. *nu 'mũ bi boni, mi seya*
 [*nu 'mũ bi = øn-i*] [*mi = seya*]
 cuando 3.PSD=<TNS>salir-L 3.IMPF=enojarse
 'Cuando salió, estaba enojada.'

Por otro lado, el tipo de contemporaneidad que se establece entre el evento y el estado que denota la cláusula depictiva favorece muchas veces que aparezca una asimetría tiempo-aspectual entre la cláusula eventual principal y la depictiva de adjunto. Esta asimetría se manifiesta de la siguiente manera:

El predicado primario de la cláusula eventual recibe morfología de TAM apropiada respecto de la temporalidad del discurso (sea éste pasado, irrealis, imperfecto, etc.). En el predicado primario de la cláusula depictiva se usa la morfología de TAM que se emplea para denotar estados en el tiempo presente. El fenómeno se ilustra en 45:

- (45) a. *bi boni xa tunthũ*
 [*bi = øn-i*] [*xa = ø = tunthũ*]
 3.PSD=<TNS>salir-L INT=3.PRES=estar_hambriento
 'Salió muy hambriento.'
 (Lit.: 'Salió, está muy hambriento.')

- b. *dá ntxaha xa dí johya*
 [dá = n-txah-a] [xa = dí = johya]
 1.PSD=M-bañarse-L INT=1.PRES=contentarse
 'Me bañé muy alegre.'
 (Lit.: 'Me bañé, estoy muy contento.')

Los ejemplos en 45 muestran verbos activos en la cláusula depictiva, y pueden incluir otros como *tuthe*, "estar sediento", *teho*, "estar vivo" o *seya*, "estar enojado". Todos estos verbos expresan estados psicológicos. Otros verbos activos que expresan cambio postural no pueden aparecer en esta asimetría de marcación tiempo-aspectual, tal como se muestra en 46:

- (46) * *xa bi mafi hüüm 'pu*
 * [(xa = bi = maf-i)] [(Ø = hüüm = 'pu)]
 INT=3.PSD=gritar-L 3.PRES=sentarse.A=allí
 Lectura buscada: 'Gritó mucho sentado allí.'

Sin embargo, la asimetría se emplea frecuentemente cuando en el predicado de la cláusula depictiva se usan un verbo incoativo o estativo. Los ejemplos siguientes en 47 sirven de ilustración. Recuérdese, tal como se vio en 3 en la sección 3, que para esta clase de verbos se utiliza morfología de perfecto para denotar estados en el tiempo presente:

- (47) a. *ma döme bi zoho xi nit*
 [ma döme bi = <z> oh-o] [xi = nit]
 1POS esposo 3.PSD=<TNS>llegar_aquí-L 3.PF=emborracharse
 'Mi esposo llegó borracho.'
 (Lit.: 'Mi esposo llegó, está borracho.')
- b. *bi hyokar t'ëi xi ñ'ixi*
 [bi = h <y> o 'k-Ø-a = r] [xi = ñ-ixi = Ø]
 3.PSD=<TNS>preparar-3OBJ-D=SG atole (3.)PF=EST-ser_agrio=3OBJ
 'Hizo el atole agrio.'
 (Lit.: 'Hizo el atole, está agrio.')

Esta asimetría de marcación "no" es una característica propia de la expresión depictiva. Encontramos en el discurso otomí muchos casos en los que esto mismo sucede, y son casos que no tienen que ver con semántica depictiva, aunque sí comparten cierta

cotemporalidad entre los eventos que se expresan. En 48 se ofrecen dos ejemplos que sirven para ilustrar el fenómeno:

- (48) a. *bi zoho, xa ñöhö 'pü ma n'agi*
 [bi = <z> oh-o] [xa = Ø = ñöh = lho = 'pu ma n'ag']
 3.PSD=<TNS>llegar_aquí-L INT=3.PRES=hablar.A=DEL=allí otro uno-vez
 'Llegaron y se pusieron a platicar mucho otra vez.' {Txt}
- b. *bi hyandi go xi zipya 'ä, ko tur fugir sëi kär ne*
 [bi = h <y> and-Ø-i] [go Ø xi = <z> i-Ø = pya = 'ä]],
 3.PSD=<TNS>ver-3OBJ-L FOC REL 3.PF=<TNS>ingerir-3OBJ=ahora=3SG
 [ko Ø = tu = r] [fugi = r] [sëi kär = r] [ne]
 porque 3.PRES=estar_sujeto=SG espuma=SG pulque P.LOC.3POS=SG boca
 'Vio que era él el que se lo había tomado porque tenía espuma de pulque en la boca.' {Ttxt}

En los ejemplos en 48 se muestran eventos que comparten cierto grado de cotemporalidad. En 48a el evento de platicar sucede de forma inmediata al evento de hablar, y el verbo que *ñöh*, "hablar" aparece en presente aunque el evento aconteciera en el pasado. Lo mismo sucede en 48b, en donde el acto de verle a alguien espuma en la cara es contemporáneo con el hecho de tenerla y el verbo *tu*, "estar pegado" aparece con morfología de presente. El mismo fenómeno sucede en la expresión depictiva, tal como se vio en los ejemplos anteriores en esta sección, pero aún no se entiende muy bien si la asimetría sucede dentro del ámbito de los adjuntos exclusivamente y se debe tomar como un rasgo de estructura oracional o si posiblemente manifiesta un fenómeno discursivo de ámbito mayor que la oración. El fenómeno debe quedar por el momento como objeto de estudio en una investigación futura.

5.3. ALCANCE SEMÁNTICO DE LA EXPRESIÓN DEPICTIVA

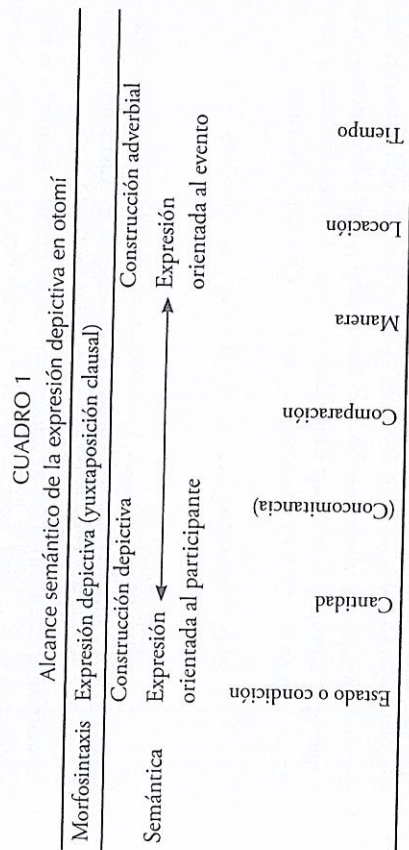
En algunas lenguas de Mesoamérica como el oluteco (véase Aissen y Zavala, en este volumen), la construcción de predicación secundaria abarca en la jerarquía de Schultze-Berndt y Himmelmann (2004) desde el extremo de estado hasta el extremo de manera o tiempo. En otras, como en Cora (véase Vázquez, en este volumen), la lengua emplea su construcción depictiva para los estados y cantidad. La expresión depictiva en otomí se usa casi exclusivamente para expresar estados y sólo ocasionalmente se encuentra yuxtaposición de cláusulas en la lengua para expresar también cantidad, tal como se ilustra en 49:

- (49) a. *jáhní gir ma, menta ga huntabe dí ga 'thohe pa ga m̃ahe 'ne*
 [jáhn-i gir < m > a], [menta ga = huntá = he],
 apurarse-1 2.PRES.SUB <TNP>ir mientras 1.IRR=]untar=1 PL.EX
 [dí = ga 'tho = he] [pa ga = < m > a = he = 'ne]
 1.PRES=todo=1 PL.EX para 1.IRR=<TNP>ir.PL.A=1 PL.EX=también
 'Apúrate a irte, mientras nos juntamos todas para ir también.' {Txt}

Por otro lado, otras áreas semánticas como la comparación se suelen también codificar a través de la expresión depictiva, tal como se observa en 50, en donde una piedra se compara con una mesa.¹³

- (50) *hont'ä nor do mí 'bü' njabu, mí nja ngur mexa*
 [hont'ä no = r do mí = 'bü-' = njabu]
 FOC DEF.SG=SG piedra 3.IMPF=estar.A-D=así
 [mí = (n)ja ngu = r mexá]
 3.IMPF=existir como=SG mesa
 'Había una piedra así como mesa.'
 (Lit.: '...era como mesa.') {Ttxt}

En general podríamos decir que la expresión depictiva en otomí, consistente de cláusulas en yuxtaposición, tiene el alcance que se presenta en la figura 1 dentro del continuo entre semántica depictiva y adverbial propuesto por Schultze-Berndt y Himmelmann (2004):



¹³ Como proposición *ngu* aparece casi siempre en colocación con el préstamo *komo* del español, v. g. *komo ngughe (komo ngu-gi = he)* (como como-1=PL.EX) 'como nosotros.'

5.4. OTRAS ESTRUCTURAS CON SEMÁNTICA DEPICTIVA: PREDICADOS COMPLEJOS

Por último, es preciso destacar que ciertos elementos que son semánticamente depictivos y que expresan rol pueden aparecer en otomí como miembros de compuestos verbales formando un predicado complejo. Considérese el ejemplo siguiente:

- (51) a. *dá teñhoya*
dá = te-ñhoya
 1.PSD=crecer-ser_pobre
 'Crecí pobre.'
- b. *má to dí janönö nge'ä xi ra hojã' i ko ngeke*
má to dí = ja-nönö
 1 POS suegra 1 PRES=hacer-madre
nge'ä xi = Ø = ra ho-jã' i ko nge-ke
 porque ENF=3.PRES=SG ser_bueno-gente con FOC-1SG
 'Hago de cuenta que mi suegra es mi madre, porque es muy buena conmigo.'
 (Lit.: 'Tengo a mi suegra por mi madre...')
 (Otomí del Valle del Mezquital, Hernández Cruz *et al.*, 2004: 117, glosas ELP)

Una prueba importante que demuestra que las estructuras en 51 son compuestos verbales es poder emplazar el clítico de plural de sujeto al final de la secuencia, tal como se muestra en 52a, un ejemplo típico de este tipo de compuestos. El ejemplo 52b muestra que esto es posible:

- (52) a. *dá m̃fuxhwöhihe*
dá = m-fux-hwöhi = he
 1.PSD=AP=cultivar.A-milpa=1 PL.EX
 'Cultivamos la milpa.'
- b. *dá teñhoyahe*
dá = te-ñhoya = he
 1.PSD=crecer-ser_pobre=1 PL.EX
 'Crecimos pobres.'

Este comportamiento contrasta con la imposibilidad de emplazar el clítico después de la raíz verbal, tal como se demuestra en la agramaticalidad de 52:

- (53) a. * *dá m̄f̄axe hwōhi*
 * *dá = m-_Jax = e* *hwōhi*

1. PSD=AP=cultivar.A=1 PL.EX milpa
 Lectura buscada: 'Cultivamos la milpa.'

- b. * *dá tehe ñhoya*

* *dá = te = he* *ñhoya*

1. PSD=crecer=1 PL.EX pobre
 Lectura buscada: 'Crecimos pobres.'

Si la estructura en 54 hubiera sido un ejemplo más de la marginal construcción de predicación secundaria depictiva del otomí, el ejemplo 53b hubiera sido gramatical, de la misma forma que el ejemplo 2, repetido aquí como 54:

- (54) *gá n̄xah̄u (n)ḡox̄jiho*

gá = n-txa = h̄u *(n)ḡox̄ji₁₂ = ho*

2. PSD=M-bañarse.A=PL encuerado=DEL
 'Ustedes se bañaron encuerados.'

Como codificación alternativa de la semántica depictiva en los compuestos en 51, se puede optar en otomí por una expresión depictiva con dos cláusulas, del tipo en 55, siendo el sujeto 55a uno en singular y el de 55b uno en plural:

- (55) a. *dá te n̄dí ñhoya*

[*dá = te*]_{CL-EV} [*n-dí = ñhoya*]_{CL-DEFTV}

1. PSD=crecer IMPF-1. PRES=ser_pobre

'Crecí pobre'

(Lit.: 'Crecí, era pobre.')

- b. *dá tehe n̄dí ñhoyah̄e*

[*dá = te = he*]_{CL-EV} [*n-dí = ñhoya = he*]_{CL-DEFTV}

1. PSD=crecer=1 PL.EX IMPF-1. PRES=ser_pobre=1 PL.EX

'Crecimos pobres.'

(Lit.: 'Crecimos, éramos pobres.')

6. Conclusiones

En este trabajo he propuesto que el otomí de SIT es una lengua que posee una construcción de predicación secundaria depictiva, pero que contrasta sorprendentemente respecto de las construcciones de predicación secundaria que se encuentran en otras lenguas de Mesoamérica en que en esta lengua la construcción depictiva sólo acepta dos elementos léxicos de registro muy informal, que son, además, sinónimos: se trata de *ñhemo* y *(n)ḡox̄ji*, "encuerado".

Ya que la construcción depictiva tiene un ámbito de uso tan insólitamente reducido, la lengua hace uso de cláusulas yuxtapuestas para dar expresión gramatical al amplio espectro de la semántica depictiva, en especial a la expresión de estados concomitantes del evento principal. A esta estructura de cláusulas la he llamado "expresión depictiva" y he mostrado que la cláusula que refiere al estado o condición en esta expresión se comporta en general como otro adjunto.

La emergencia de la construcción de predicación secundaria en otomí de SIT es todo un misterio, ya que no se discierne en la lengua ninguna clave que apunte hacia la existencia de un fenómeno de mayor integración estructural entre las cláusulas que aparecen en la expresión depictiva, como es el caso por ejemplo del tselal (véase Polian y Sánchez Gómez, en este volumen). Resta que un futuro trabajo de investigación en otras lenguas otomíes arroje nueva luz sobre las áreas que aún no se comprenden en su totalidad.

JUXTAPOSED ADJUNCT CLAUSES IN OTOMI: EXPRESSING BOTH DEPICTIVE AND ADVERBIAL SEMANTICS¹

ENRIQUE L. PALANCAR

UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

In this article, I study the grammatical expression of “depictive” content (Schultze-Berndt and Himmelmann 2004) (‘he arrived drunk’) and adverbial content (‘he arrived well’) in Otomi. I show that speakers of this Mesoamerican language encode both depictive and adverbial content by means of juxtaposed adjunct clauses. While this structure is widely and productively used to express depictive semantics, it shows a number of distributional restrictions when encoding several other notions including manner, extent, and time, which are generally expressed by means of adverbs. I claim that a substantial number of such adverbs emerged as the lexicalized outcome of former verbs functioning as predicates in adverbial clauses.

[KEYWORDS: Otomi, Otomanguean, depictive, adverbs, secondary predication]

1. Introduction to the object of study. Otomi is a Mesoamerican language family spoken in Central Mexico belonging to the Otopamean group of Otomanguean.² In this article, I present data from the Otomi spoken in the village of San Ildefonso Tultepec, in the state of Querétaro, Mexico. My data

¹ This study has been financed by the CONACyT (project no. 47475). An earlier version of this article was presented at the Syntax of the World’s Languages Conference II, held in September 2006 at Lancaster University. The analysis concerning depictive clauses resulted from a workshop coordinated by Judith Aissen on secondary predication that took place in Mexico, in which other Mesoamericanists actively participated. I am heartily grateful to Judith Aissen for her immense generosity, help, and wisdom in this matter and others. I want to thank Roberto Zavala for inviting me to that inspiring workshop and for always helping me pick out the obvious from the invisible. I am also very grateful to Eva Schultze-Berndt and Albert Bickford for all their wise and helpful observations on the article as reviewers, and to *IJAL*’s associate editor for his/her comments. My most sincere thanks go also to Mrs. Anastacia Cruz Vázquez for sharing her ancient Otomi wisdom with me. All errors and deficiencies remain my own responsibility.

² The name “Otomi” has been challenged lately by speakers of Mezquital Otomi—the largest speaking community—who prefer to call the language “Hñähñu” (with various spellings), itself the native denomination from Mezquital. A growing number of authors have started to use the name Hñähñu instead of Otomi. In this article, I stick to the old label for two reasons. On the one hand, speakers of SI Otomi do not recognize the Mezquital word Hñähñu as a representative word for their own language, Hñöñho being their native denominative. On the other hand, “Otomi” is by far the term with a wider history in both linguistics and anthropology, and to date it is not regarded, to my knowledge, as a pejorative term by the majority of native indigenous people belonging to this ethnic group.

come from my own fieldwork in the language, both from texts and from elicitation with native speakers.³ I refer to this dialect as SI Otomi. Nevertheless, the phenomenon I study here is found in other varieties too (Hekking and Andrés de Jesús 1984, Hekking 1995, Voigtlander and Echegoyen 1985, and Lastra 1992; 1997) but has not been provided a unified treatment in the literature. Given the large dialectal diversity of Otomi, it is expected that other Otomi languages display different lexical and constructional idiosyncrasies.

In this article, I study the grammatical expression of depictive content ('he arrived drunk') and adverbial content ('he arrived well') in SI Otomi, and I show that speakers of this Mesoamerican language encode both depictive and adverbial content by means of juxtaposed adjunct clauses, which despite a number of differences, could be analyzed as representing the same structure in the grammar of this language. This encoding strategy is reminiscent of the "General Adjunct Construction" proposed in Schultze-Berndt and Himmelmann (henceforth S-B&H) (2004) for other languages, but it also has important differences, as I shall show.

Depictive semantics is often encoded cross-linguistically by means of secondary predication. S-B&H (2004) review the constructional notion of secondary predication in linguistic theory and propose a number of cross-linguistically valid criteria to identify a given secondary predication construction devoted to the expression of depictive content. The authors refer to the instances of such a construction as "depictive secondary predicates" or just as "depictives."⁴ The adjectival phrase (AP) *drunk* in the clause (CL) in (1) serves as a good example of a depictive in English:

- (1) [*The man arrived (drunk)*_{AP}]_{CL}

In (1), adjectival phrase *drunk* qualifies as depictive by virtue of fulfilling a number of grammatical criteria: (a) it is a predicate (i.e., it predicates the state of being drunk of the man); (b) it is an adjunct (i.e., it is optional and the main predicate does not change its basic meaning in its absence); (c) it is nonfinite (i.e., it does not receive tense or aspect morphology); (d) it is a structurally subsidiary to the main predicate of the clause (i.e., it belongs to the same clause of the main predicate); (e) it occurs within the prosodic boundaries of the clause; and (f) it expresses depictive semantics. This last feature plays a fundamental role in the identification of depictives.

Depictive semantics is defined by S-B&H (2004:65) as the type of semantics that expresses a state of affairs "which holds true at the same time as the eventuality encoded by the main predicate." In this sense, the state of being

³ Otomi examples are from my personal archives.

⁴ More recently, Aissen and Zavala (forthcoming) have contributed further to the characterization of depictives with the application of a strict syntactic methodology in the analysis of a variety of Mesoamerican languages.

drunk predicated of the man by *drunk* is simultaneous or co-temporal to the event of his arrival; that is, the state of being drunk holds true at the same time as the arriving event encoded by the verb *arrived*. The depictive secondary predicates in (2) are further examples of depictive semantics:

(2) *The man arrived happy, disturbed, weeping, talking, etc.*

Depictive secondary predicates are common across the world's languages, but there are languages that lack them. Such languages employ other grammatical means to express similar depictive content. I argue in 3.1 that SI Otomi is one of these languages. In SI Otomi, depictive content is commonly expressed by means of an independent adjunct clause, which is juxtaposed to the main clause, which encodes the main event. An example is given in (3); the clause expressing the depictive content is given in square brackets:⁵

(3) *nǒ=r ndǒ bi=<z>óhó [xki=n-tí]*
 DEF.SG=SG man 3.PST=<NPS>arrive.here.F 3.PPRF=M-get.drunk

'The man arrived drunk' (lit., 'The man arrived, he had got drunk').

The adjunct juxtaposed clause *xki ntí* 'he had got drunk' in (3) expresses a very similar conceptual meaning to the semantics of *drunk* in (1). The clause not only informs us that the man had gotten drunk but, more crucially, that he was indeed in a drunk state at the time of his arrival.

More recently, following up on the proposal in their 2004 article, Himmelmann and Schultze-Berndt (2005) have proposed that depictive secondary predicates of the type in (1) can be found across languages expressing

⁵ In the transcription of Otomi, I have used the orthography in Hekking (2002) with only two innovations, mainly the use of <ř> for trill /r/, and a superscript ^h for salient phonetic aspiration. Deviations from the IPA convention are the following: <ʔ> /ʔ/; <ɟ> /p^h/; <ɟ> /k^h/; <ɣ> /j/; <x> /ʃ/; <ɬ> /tʃ/; <ñ> /ɲ/; <ɾ> /r/; <ř> /r/; <a> /a/; <ã> /ɔ̃/; <ẽ> /ẽ/; <õ> /ɔ̃/; <ü̃> /ĩ/. Umlaut indicates nasalization (i.e., <ä̃> /ä̃/; <ü̃> /ü̃/, etc.). Otomi has three tones: ascending, Ǻ; high, Ǻ; and low, which is not represented. Given that tone differences between SI Otomi and Mezquital Otomi are minimal, I have consulted the excellent transcriptions in Wallis (1956) and Hernández Cruz et al. (2004) for uncertain cases in SI Otomi. Abbreviations used are: < > nonconcatenative morpheme; – lexicalized morpheme; = clitic; A A-type person-marking paradigm; ABS absolutive; ACT activity marker; ANIM animate; ANTI antipassive; ART article; ADV adverbial/correlative; B Bound-form; B B-type person-marking paradigm; D dependent; DAT dative; DEF definite; DEL delimitative; DEM demonstrative; DIM diminutive; EX exclusive; F Free-form; I independent; IMP Imperfect; IMPER impersonal; IMPERA imperative; INCOM incomplete; INT intensive; IRR Irrealis; LOC locative; LOC~P locative preposition; M middle voice; MIRAT mirative; N neuter; NEG negation; NF nonfinite; NI nasal intransitive; NP nominal predication; NPS nonpresent stem; OBJ object; P particle; PCPL participle; PL plural; POSSR possessor; POSS possessive; PPRF Pluperfect; PRF Perfect; PRS Present; PST Past (completive); PURP purposive; QUOT quotative; REL relative; s suffixal form; SAP speech-act participant; SG singular; ST stative; SUBJ subjunctive; TAM tense/aspect/mood; TR transitive. In addition: S-B&H = Schultze-Berndt and Himmelmann; ZM = Zavala Maldonado.

adverbial semantics. The authors distinguish here two types of adverbs with regard to their semantics: “participant-oriented adverbs,” such as *impatiently* in (4a), and “event-oriented adverbs” like *quickly* in (4b) (examples are from Crystal 1996:141):

(4a) *Jack growled impatiently at the delay*

(4b) “Of course,” *she said quickly*

Participant-oriented adverbs, like depictives, are semantically predicative because they express the state a given participant is in when carrying out the action. What is important in (4a) is not so much that the growling was performed in an impatient manner, but rather that Jack was impatient when he growled. Such a semantic implication is not inferred in the use of *quickly* in (4b), an event-oriented adverb. This adverb expresses only the manner in which the action of speaking was performed; nothing is indicated about the state of the subject. The intricate semantic relation between adverbial and depictive means that, in some languages, both meanings are at times indistinguishable for grammatical purposes. Consider the German example in (5), where *wütend* ‘angry’ could be analyzed as the head of an adverbial phrase or of an adjectival phrase in secondary predication; there is no clear formal distinction between the two constructions (adapted from S-B&H 2004:61):

(5) German (Indo-European)

Claire hat wütend das Zimmer verlassen
 Claire has ANGRY ART.N.SG room(N) left

‘Claire left the room angry/angrily’.

S-B&H (2004) and Himmelmann and Schultze-Berndt (2005) propose that languages may take to an extreme the lack of formal distinction between depictives and adverbials. In such situations, the grammatical construction involved must be analyzed as a “General Adjunct Construction” rather than as secondary predication. Many Australian languages make use of such a General Adjunct Construction (S-B&H 2004). Zavala Maldonado (henceforth ZM) (2004) has convincingly shown that Olutec (Mixe-Zoquean) is one such language in Mesoamerica. Consider the examples in (6), which show two instances of a secondary predication construction (both the glosses and the translation are adapted into English from the Spanish original):

(6) Olutec (Mixe-Zoquean)

(6a) *ku:xu tan=wit-an-e:t=na ja:=na=koj ko:xo*
 lame A1(ABS)=walk-IRR.D-PL.SAP=still another=still=just day

‘We’re going to walk lame just for a couple of days more’.

(ZM 2004:294)

- (6b) *ya:xʔ-pa-k* *?i=mi:nʔ-nü-e*
 scream-NF=ANIM A3(ABS)=come-already-INCOM.D
 ‘He comes screaming’. (ZM 2004:299)

In (6a), the adjective *ku:xu* ‘lame’ occurs in preverbal position. The verb in (6a) receives a special aspectual morphology, called “dependent” (D), which marks the existence of a syntactic dependency relationship between the verb and another phrasal constituent. The same phenomenon is shown in (6b), but with a nonfinite verb instead of an adjective. Secondary predicates in Olutec may also follow the verb,⁶ and when they do so, the verb is no longer marked with dependent aspect but with the so-called independent (i) one, as shown in (7):

- (7) Olutec (Mixe-Zoquean)
 (7a) *jeʔ=k* *tan=kay-pe-:t-ütz* *tza:yʔ-ek*
 that=ANIM A1(ERG)=eat-INCOM.I.TR-PL.SAP-EX roast-PCPL
 ?i=tzuʔchi
 A3(PSR)=meat
 ‘We eat the meat roasted’. (ZM 2004:306)

- (7b) *jaykak* *∅=mi:nʔ-u* *piyüʔk-pa* *?i=jüʔna-pi*
 people A3(ABS)=come-COM.I run-NF A3=behind-LOC
 ‘The people came running after him’. (ZM 2004:306)

Adverbs are treated in Olutec grammar like depictive secondary predicates. They may either precede or follow the verb. Like in (6), when the adverb precedes the verb, the verb receives dependent aspect, as shown in (8a). Like in (7), when the adverb follows the verb, the verb receives independent aspect (8b):

- (8) Olutec (Mixe-Zoquean)
 (8a) *pi:nak=ak* *?i=?e:p-e*
 little=ANIM A3(ABS)=see-INCOM.D
 ‘He sees a little bit’. (ZM 2004:299)
 (8b) *∅=ka:=ja=?e:p-pa=k* *ni* *pi:nak*
 B3(ABS)=NEG=MIRAT=see-INCOM.I=ANIM NEG little
 ‘He doesn’t see a thing’ (lit., ‘He doesn’t see not little’).⁷
 (ZM 2004:307)

⁶Zavala Maldonado (2004:306) reports that this order is not as common as the one illustrated in (6).

⁷Notice that the adverb *p:inak* in (8b) does not function as a direct object despite the literal translation.

There is no formal distinction between the constructs in (8) and the ones in (6) and (7). Accordingly, ZM (2004) takes them to represent the same grammatical construction. Nonetheless, the construction is neither depictive nor adverbial in its semantics; rather, it encodes a type of syntactic adjunct relation embedded within the same clause. Such a construction is analyzed by ZM as a General Adjunct Construction.

I claim that SI Otomi is a language that has an equivalent structure to the General Adjunct Construction in Olutec, but one which is a structural “mirror image,” as it were. In Otomi, depictive and adverbial semantics may also be treated alike in grammatical terms; but instead of a monclausal construction, as in Olutec, the Otomi structure consists of a main clause and a juxtaposed adjunct clause expressing those meanings. Recall that the language commonly expresses depictive semantics by means of an adjunct clause, not by a secondary predication construction, as was shown in (3) above (repeated here as 9a). (9b) further illustrates that adverbial meanings may also be encoded by means of a similar juxtaposed clause:

(9a) *nǒ=r ndǒ bi=<z>óhó [xki=n-tí]*
 DEF.SG=SG man 3.PST=<NPS>arrive.here.F 3.PPRF=M-get.drunk
 ‘The man arrived drunk’ (lit., ‘The man arrived, he had got drunk’).

(9b) *tó’be bi=sígi [bi=njapu] bi=n=ka^h-té*
 still 3.PST=go.on 3.PST=be.so 3.PST=NPS=deceive.S-ACT
 ‘He still went on thus deceiving people away’ (lit., ‘He still went on, it was like that, he deceived’).

The examples in (9) show that both depictive and adverbial semantics may be expressed in Otomi by juxtaposed clauses, which pertain to the realm of subordination. Interestingly, while depictive semantics is widely and productively expressed in Otomi by means of adjunct clauses, the type in (9b) shows a number of distributional restrictions when encoding other adverbial meanings (mainly manner, extent, and temporal meanings). The restrictions reveal that the structure in question is no longer a productive mechanism. In contrast, Otomi now encodes such adverbial notions by means of adverbs. In 5 below, I claim that a number of such adverbs emerged from the lexicalization of former verbs, which once functioned as predicates in adverbial clauses.

In 2, I give a basic introduction to verbs in SI Otomi. In 3, I introduce the way that depictive semantics is encoded in this language, by means of both a secondary predication construction and an adjunct clause, as in (9a). In 4, I compare this structure with the juxtaposed clauses used to express a miscellanea of adverbial semantics. In 5, I study this adverbial clause in further

detail and show that a number of adverbs in the language were once verbs. The general proposal of the article is presented as a conclusion in 6, where I compare the two types of juxtaposed clauses presented in this paper and propose that they represent a unified subordination structure in Otomi that encodes adjunct relations, regardless of the semantic type involved.

2. Introduction to the language. SI Otomi is an accusative language with a basic SVO order. As for lexical classes, the language has verbs, nouns, adverbs, and particles, but it does not have adjectives (see Palancar 2006a). Predicates may be nominal or verbal. Verbs inflect by means of function words for a number of inflectional dimensions which include—among others—tense, aspect, mood, and agreement with the subject NP. Such function words are proclitics, but they may receive stress. A number of common paradigms are illustrated in (10):

(10)		First	Second	Third
		Person	Person	Person
	Indicative			
	Present (~ Habitual)	<i>dí=</i>	<i>gí=</i>	$\emptyset=$
	Imperfect	<i>n-dí=</i>	<i>n-gí=</i>	<i>mí=</i>
	Past (~ Completive)	<i>dá=</i>	<i>gá=</i>	<i>bí=</i>
	Perfect	<i>(x-)tá=</i>	<i>(x-)ká=</i>	<i>xí=</i>
	Pluperfect			<i>(x-)kí=</i>
	Irrealis			
	Irrealis (~ Future)	<i>ga=</i>	<i>gi=</i>	<i>da=</i>
	Imperfect Irrealis	<i>n-ga=</i>	<i>n-gi=</i>	<i>n-da=</i>

Verbs are classified as transitive or intransitive (only a few verbs exist that can be used in both ways). Intransitive verbs are in turn organized into five classes: active verbs; nasal intransitive verbs, which encode human activities (Palancar 2006b); inchoative verbs expressing inchoative states (Palancar [forthcoming]); patient verbs indicating inactive state of affairs (Palancar 2008); and stative verbs, which depict property concepts (Palancar 2006a).

2.1. Inflected forms of verbs. Most verbs in Otomi have two inflected forms: a “Free-form” (F) and a “Bound-form” (B). Apart from being prosodically conditioned,⁸ the forms are syntactically conditioned (for more details, see Palancar 2004). A Free-form is used when the verb occurs at a clausal boundary before another predicate, as in (11a), or before another clause, as in (11b), but also when the verb is followed by a certain class of adverb

⁸Free-forms are phonologically free and function as pausal forms used at the boundary of a phonological phrase. Bound-forms are cliticized forms which cannot occur at a prosodic boundary.

phrases within the same clause, as in (11c). The latter phenomenon is discussed in 5.3 below.⁹

- (11a) [pwes yá dá=jüt'- \emptyset -i] [bi=<m>á]
 so P 1.PST=pay-3OBJ-F 3.PST=<NPS>go
 'So I paid him (and) he left'.

- (11b) [jánge méhó yǒ gi=xǎx'- \emptyset -i] [pá
 because.of.that better IMPERA.NEG 2.IRR=provoke-3OBJ-F PURP
 hín-té gi=thóg'- \emptyset -i]
 NEG-something 2.IRR=experience.F- \emptyset -F
 'Because of that, it's better you don't provoke them, so that nothing
 (bad) happens to you'.

- (11c) [kabu \emptyset =tsǒn-i] (hyáxtho)_{ADVp}]
 after.all 3.PRS=arrive.there.F-F daily
 'After all, they get there daily'.

In contrast, a Bound-form is used when the verb occurs intraclausally, as in (12):¹⁰

- (12a) [yá bi=t-hǒ^hk- \emptyset -a=(nǒ=r kártérá)_{NP}]
 P 3.PST=IMPER-make-3OBJ-B=DEF.SG=SG road
 'They built the road'.
- (12b) [bi=h<ñ>ú-'=(ká gǒxthi)_{pp}]
 3.PST=<NPS>sit.B-B=LOC~P door
 'He sat at the door'.
- (12c) ha nǒ='ǒ [\emptyset bí=hǔx-a=(máñǒ)_{ADVp}]
 and DEF.SG=DEM.SG.III REL 3.PRS=sit.on-B=above
 'And the one sitting above'.

2.2. Clausal boundaries. Here, I establish three morphosyntactic criteria to define a clausal boundary for SI Otomi. Such criteria are useful for understanding secondary predication in the following section.¹¹ The criteria are as follows:

⁹ A Free-form is also used when the verb is followed by an adjunct secondary predicate, as in (16) in 3.

¹⁰ A Bound-form is always a cliticized form of the verb. In Bound-forms, a verb is hosted by the first phonological word of the following phrase: in (12a) the hosting word is the determiner cluster *nǒr*; in (12b) it is the general locative preposition *ká*; and in (12c) it is the adverb *máñǒ* 'above'.

¹¹ The arguments I present here follow the rigorous syntactic methodology proposed in Aissen (2004).

(A) A clause is always marked for tense/aspect/mood (TAM):

- (13a) [**gi**='ěhě] [gi=tsǎn-g-a=gi=wá ma ngǔ]
 2.IRR=come.F 2.IRR=scold-1OBJ-B=1OBJ=here 1POSS house

'You'd come (and) scold me here at my house'.

- (13b) [**hí-mí**=hek-'-a='í] [**n-gi**=kö-i]
 NEG-3.IMP=let.s-2OBJ-B=2OBJ IMP-2.IRR=go.down.L-L

'He wasn't letting you come down (from the tree)'.

The items in boldface in (13) are examples of TAM markers in (10). The examples in (13) show two verbs aligned in discourse sequence. These verbs function as main predicates of different clauses by virtue of being marked for TAM categories.

(B) When a clause has a plural subject of first or second person, the clause is morphologically marked for plural subject by means of a number enclitic to the verb. This is shown in (14), where each of the two clauses has a plural (coreferential) subject, and consequently each one receives the appropriate plural subject marking:

- (14) [pé nú='í=hú] xa=gi='ó=**hú**] [níxi gi=tsǎ=**hú**]
 but DEF=2=PL INT=2.PRS=sleep.S=PL nor 2.PRS=feel=PL

'But you are sound asleep and do not notice it'.

(C) Finally, when semantically appropriate, each clause is marked separately, and only once, for negative polarity, as in (15):

- (15) [**hín**-da=<z>ón-i] [**hím**-bi=<m>á]
 NEG-3.IRR=<NPS>arrive.there-F NEG-3.PST=<NPS>go

'He wouldn't arrive there, he didn't go'.

Next, I present the different structures that convey depictive semantics in the language.

3. Expressing depictive semantics in SI Otomi. In Palancar (forthcoming), I propose that SI Otomi has a secondary predication construction which is employed to express a limited range of depictive content. An example of such a construction is given in (16) with the predicate *nthěmpó* 'naked':

- (16) gá=n-txáhá nthěmpó=thǒ
 2.PST=M-bathe.F naked=DEL

'You had a bath naked'.

The predicate *nthěmpó* 'naked' functions in (16) as a secondary predicate which expresses depictive semantics; it indicates that the man was naked at

the time he bathed. In order to show that this phrase is indeed an adjunct predicate which occurs within the same clause as the main predicate, I apply here the three morphosyntactic criteria given in 2.2 to define a clausal boundary. Observe in (16) that the predicate *nthěmpó* ‘naked’ is not marked for TAM, so it does not form a clause by itself. Similarly, a secondary predicate is not marked for a plural subject (or controller), as shown in (17a), which contrasts with the ungrammatical (17b):

(17a) *gá=n-txá=hú* (n)*gǒxji=thǒ*
 2.PST=M-bathe.S=PL naked=DEL
 ‘Y’ all had a bath naked’.

(17b) **gá=n-txá=hú* (n)*gǒxji=thǒ=hú*
 2.PST=M-bathe.S=PL naked=DEL=PL
 Intended meaning: ‘Y’ all had a bath naked’.

Similarly, a secondary predicate cannot receive negation. This is illustrated in (18a), contrasted with the ungrammatical (18b):

(18a) *hín-n-dí=má* (n)*gǒxjí=thǒ*
 NEG-IMP-1.PRS=go naked=DEL
 ‘I wasn’t going around naked’.

(18b) **n-dí=má* *híngi* (n)*gǒxjí=thǒ*
 IMP-1.PRS=go NEG naked=DEL
 Intended meaning: ‘I wasn’t going around naked’.

Notice that, in (18a), negation occurs in the main predicate but has semantic scope over the secondary predicate (Aissen 2004). In other words, the sentence indicates that the speaker was indeed going around, but not in a naked state.

The examples in (16)–(18) above show that SI Otomi has a secondary predication construction. However, the two synonymous lexemes *nthěmpó* and (n)*gǒxjí* ‘naked’ are the only two words in this language that instantiate the construction. In other words, NO other predicate in the language can occur to my knowledge in this syntactic slot.¹² This may be seen by the ungrammaticality of (19):

¹² My knowledge on this point is based on having no corpus evidence of the occurrence of other words in the predicative slot of the depictive construction, and in not having found another example in the extensive elicitation I carried out specifically on the matter. However, as for any other realms of the grammar, some degree of variation across speakers is certainly a possibility.

- (19) **dá=tsóhó*
 1.PST=arrive.here.F
n-tĩ/tú_nthũ/n-txá_nǎ, etc.
 M-get.drunk/have_hunger.B/M-get.disheveled_head

Intended reading: ‘I got here drunk/hungry/disheveled, etc.’

3.1. The depictive clause. The ungrammaticality of (19) shows that depictive semantics is not productively encoded in the language by means of its secondary predication construction. Instead, SI Otomi makes use of a juxtaposed clause to express such content. Examples in (20) illustrate such an expression:

- (20a) *xa=bi=máf-i* [*mí=húm='pu='na*]
 INT=3.PST=scream.F-F 3.IMP=sit.S=there=QUOT
 ‘They say he was screaming sitting there’ (lit., ‘He screamed a lot, he was sitting there, hearsay’).
- (20b) *bi=ón-i* [*mí=jóhyá*]
 3.PST=<NPS>exit.F-F 3.IMP=be.happy
 ‘She went out happy’ (lit., ‘She went out, she was happy’).

Examples in (20) show instances of the same type of depictive content as in (17). In (20a), the referent of the subject (the man) is sitting while he is screaming. Similarly, in (20b), the woman is happy at the very same time she is performing the action of exiting her house. However, the structure expressing this depictive content is an adjunct clause, not an adjunct predicate. I call such a clause the “depictive clause.” The three criteria presented in 2.2 may be used to show that it is a clause.

The structure is marked for TAM categories, as shown in (21a), in contrast to the ungrammaticality of (21b)—this is also shown in (19):

- (21a) *nú'mu* *mí=ndě*, *mí=<z>óhó*
 when 3.IMP=be.late 3.IMP=<NPS.ADV>arrive.here.F
xki=n-tĩ
 3.PPRF=M-get.drunk
 ‘When it was getting late, he arrived drunk’ (lit., ‘He arrived, he had got drunk’).
- (21b) **mí=<z>óhó* *n-tĩ*
 3.IMP=<NPS.ADV>arrive.here.F M-get.drunk
 Intended meaning: ‘He arrived drunk (when something happened)’.

Likewise, when the predicate in the depictive clause has a first- or a second-person plural subject, the clause is marked accordingly, as shown in (22) by the use of a plural number enclitic:

- (22) *yá dé 'rá^hto ndě ga=tsó=hé yá*
 P of six evening 1.IRR=arrive.here.S=1PL.EX P
xi=<z>d^h-k=hé
 (3.)PRF=<NPS>get.tired-1OBJ=1PL.EX

‘At six in the evening we arrived all tired’ (lit., ‘At six in the evening we arrived, we are all tired’).

Notice that, in (22), both the main clause and the depictive clause have a coreferential plural subject of first person, and each receives the relevant morphology for plural subject agreement. The second predicate in (22) is a patient verb that encodes its only argument as object (for more on this type of verb, see Palancar 2008).

Finally, a depictive clause is independently marked for negation. This may be seen in (23a), in contrast to (23b), which has a different meaning:

- (23a) *bi=<z>óhó hín-xki=n-tí*
 3.PST=<NPS>arrive.here.F NEG-3.PPRF=M-get.drunk
 ‘He didn’t arrive drunk’ (lit., ‘He arrived, he hadn’t got drunk’).

- (23b) *hím-bi=<z>óhó xki=n-tí*
 NEG-3.PST=<NPS>arrive.here.F 3.PPRF=M-get.drunk
 ‘He didn’t arrive; (because) he was drunk’.

When speakers of SI Otomi want to negate a given depictive meaning, they place the negation morpheme in the predicate of the depictive clause—as in (23a)—not in the predicate of the main one—as in (23b), where (only) the main event is negated (for more details on this structure, see Palancar [forthcoming]).

3.2. Other characteristics of the depictive clause. The depictive clause is commonly placed after the main clause. If this neutral word order is inverted, the delimitative particle *thǒ* is used on the depictive clause to mark that the preposed clause is in focus.¹³ Similarly, the neutral word order is often inverted when the depictive clause expresses a simultaneous action, rather than a state. Both cases are illustrated in (24):

¹³ Similar delimitative (or “restricted”) markers, which serve both the function of restrictive marking (‘just’) and focus marking, have also been identified in S-B&H (2004) and Himmelmann and Schultze-Berndt (2005) for other languages.

- (24a) *xki=n-tĩ=thǒ* *bi=<z>óhó*
 3.PPRF=M.get.drunk=DEL 3.PST=<NPS>arrive.here.F

‘It was drunk, he got here!’ (lit., ‘He’d gotten drunk all right, he arrived here’).

- (24b) *yá n-dí=ñó=hé* *(n-dí=)pá=hé=’pú*
 P IMP-1.PRS=talk=1PL.EX IMP-1.PRS=go.PL=1PL.EX=there

‘We were going along talking’/‘We were talking (as) we were going along’ (lit., ‘We were talking, we were going’).

In all the examples above, the depictive clause has a TAM which is temporally or aspectually coherent with the one in the main clause. However, more often than not the depictive clause in Otomi appears in a neutral TAM, which is elsewhere used to express present state of affairs.¹⁴ Various examples of this common marking are given in (25):

- (25a) *dá=n-txáhá* *xa=dí=jóhyá*
 1.PST=M-bathe.F INT=1.PRS=be.happy

‘I had a bath very happy’ (lit., ‘I bathed, I’m happy’).

- (25b) *bi=ón-i* *xa=∅=tú_nthǔ*
 3.PST=<NPS>exit.F-F INT=3.PRS=have_hunger.B

‘He went out hungry’ (lit., ‘He went out, he’s hungry’).

- (25c) *ma dómé* *bi=<z>óhó* *xĩ=n-tĩ*
 1POSS husband 3.PST=<NPS>arrive.here.F 3.PRF=M-get.drunk

‘My husband got here drunk’ (lit., ‘My husband got here, he’s drunk’).

All predicates in the main clause receive the relevant TAM appropriate for discourse purposes, which in the examples in (25) is the past tense. Notice the predicate of the depictive clause appears in the typical morphology which expresses present state of affairs: in both (25a) and (25b), this morphology is the so-called Present tense, while in (25c) it is the Perfect. The verb *ntĩ* is an inchoative (middle) verb, and Perfect is used with both inchoative and stative verbs to express current states of affairs.

As for syntax, depictive clauses bear one argument in coreference with the arguments of the main clause. In the examples above, this argument—the

¹⁴ Eva Schultze-Berndt, as a referee, has suggested to me that neutral TAM appears here to have a more general function of simultaneity (or “relative tense”), which probably developed from the interpretation of present states of affairs as being simultaneous or overlapping with speech time, resulting in an absolute tense interpretation.

controller—is the subject of the main clause, but example (26) shows that the controller can also be the object of the main clause:

- (26) *gá=h<y>and-g_i-a=gí_i* *xtá_i=n-txá_i ñó*
 2.PST=<NPS>see-1OBJ-B=1OBJ 1.PRF=M-get.disheveled_head
 ‘You saw me with untidy hair’ (lit., ‘You saw me, I had gotten uncombed’).

In this section, I have shown that SI Otomi has a rather marginal secondary predicate construction which is only used with predicates meaning ‘naked’. For the rest of the spectrum of depictive semantics, the language uses a juxtaposed adjunct clause, which I have labeled as the depictive clause. In the following section, I show that a similar type of juxtaposed clause is also used to express a given range of adverbial meanings.

4. Adverbial modification in SI Otomi. Adverbial modification may be intraclausally encoded by means of adverb phrases—headed by adverbs—as in (27a); prepositional phrases, as in (27b); and a number of noun phrases expressing temporal or locative meanings, shown in (27c):

- (27a) *nú muntxüi mí=m-pef-i*
 3 by.night 3.IMP=ANTI-work.F-F
 ‘She worked by night’.

- (27b) *nú'mu='yá n-ga=hú=hé ká tsibi*
 when=P IMP-1.IRR=sit.S=1PL.EX LOC~P hearth
mundě
 in.the.afternoon

‘When we would sit at the fireplace in the afternoons’.

- (27c) *'ně ar hyáx'á bi=náng-i bi=<m>á*
 and SG day.after 3.PST=rise.F-F 3.PST=<NPS>go
 ‘And the day after they rose and left’.

Alternatively, adverbial modification is also achieved by adjunct clauses, which may be headed by conjunctions. Conjunctions in Otomi may be native, as in (28a); borrowed from Spanish, as in (28b); or borrowed but adapted, as in (28c):

- (28a) *bi='<y>ot'-Ø-e [hángu gá='ěm-b-i]*
 3.PST=<NPS>do-3OBJ-F as/how 2.PST=say.S-3DAT-F
 ‘He did it as you told him’.

- (28b) *dá=hé^hk-∅=hé* *nǒ* *déhe,* [*pórke*
 1.PST=divide-3OBJ=1PL.EX DEF.SG water because
hín-té *mí=ja='pu*
 NEG-something 3.IMP=exist=there

‘We put the water away, because there wasn’t any there’.

- (28c) *jangě* *bi=pi-'-a='í*
 thus 3.PST=scare.S-2OBJ-B=2OBJ
 [*kó* *gí=nkréo-∅* *ar* *nhě^hmyó*]
 because 2.PRES=believe-3OBJ SG lie

‘He thus scared you because you believe that lie’.

SI Otomi does not have many native conjunctions. (28a) shows the native element *hángu* ‘as/how’. (28b)—with the conjunction *pórke* ‘because’ from Spanish *porque*—shows a typical case of a borrowing from Spanish. The causal conjunction *kó* in (28b) is also borrowed from Spanish *como* ‘as/like’ but has been fully adapted in the language both semantically and phonologically, to the extent that it is no longer recognizable as a borrowing.

4.1. Juxtaposed adverbial clauses. Crucially for the purpose of this work, Otomi also uses juxtaposed clauses to express adverbial content, which are similar to the depictive clause described in 3.1. Relevant examples are given in (29):

- (29a) [*mí=xúdi* *n'í^hthǒ*] (*mí=*)*pǎ-∅* *txí-ká^hfé*
 3.IMP=be.the.morning early 3.IMP=sell-3OBJ DIM-coffee

‘They used to sell coffee early in the morning’ (lit., ‘It was the morning early, they used to sell coffee’).

- (29b) *xa=mí=máf-i* *mí=nts'ed-i*
 INT=3.IMP=scream.F-F 3.IMP=be.strong.F-F

‘He was screaming loudly’ (lit., ‘He was screaming, it was strong’).

This type of adverbial clause expresses a variety of meanings, from temporal, as in (29a), to manner and extent, as in (29b). This type of clause is different from the depictive clause in that it does not share any arguments with the main clause. The subject argument of this adjunct clause is the event expressed in the main clause. Like the depictive clause, the adverbial clause commonly occurs in a neutral aspect, but not always. More instances of adverbial clauses are given in (30):

- (30a) *da=mǒ* *xí=ñ-hó-∅*
 3.IRR=talk (3.)PRF=ST-be.good.3OBJ

‘He’d talk **well**’ (lit., ‘He’d talk, it’s good’).

(30b) *ntónse dá='éhě mí=ndě de_gé^hnu*
 then 1.PST=come.F 3.IMP=be.late of_over.there

‘And then, I came back **late** from over there’ (lit., ‘Then I came, it was being late, from over there’).

(30c) *tó'be bi=sígi [bi=njapu] bi=n=ka^h-té*
 still 3.PST=go.on 3.PST=be.so 3.PST=NPS=deceive.S-ACT

‘He still went on thus deceiving people away’ (lit., ‘He still went on, it was like that, he deceived’).

(30a) is a common example of an adverbial clause. The clause *xi ñhó* ‘it’s well’ modifies the whole event *da mó* ‘He’d talk’ and expresses manner. This clause occurs in a neutral aspect, used elsewhere to express present states of affairs. In (30a), as in (25c) above, the verb *hó* ‘be good’ is a stative verb, and Perfect is used with such verbs to encode current states of affairs. In (30b) and (31c), the clauses *mí nde* ‘it was being late’ and *bi njapu* ‘it was like that’ are semantically coherent with the TAM of the main clause.

Like the depictive clause, the adverbial clause normally follows the main clause. There is one regular exception: clauses expressing temporal meaning commonly occur before the main clause, presumably for pragmatic reasons, as they set the temporal frame of the event. An example is given in (31) with the verb *ngüt’ä* ‘be early’, but (29a) may also serve as an illustration:

(31) *xi=ngüt’ä dá=tsón-i*
 3.PRF=be.early 1.PST=arrive.there.F-F

‘I arrived there early/soon’ (lit., ‘It’s early, I arrived there’).

In the next section, I look at this adverbial clause in further detail because its functional productivity is much lower than the depictive clause. I also propose that a number of former verbs in the adverbial clause have given rise to current adverbs in SI Otomi.

5. The functional decay of the adverbial clause. In contrast to the depictive clause, the adverbial clause is not a productive grammatical structure in SI Otomi, as is shown in the following subsections.

5.1. Lexical restrictions of the adverbial clause. Only a few verbs may head the verbal phrase of the adverbial clause. All the verbs in SI Otomi which, to my knowledge, can occur in an adverbial clause are given in table 1. Such verbs can only inflect in third person when they function as predicates of this clause. Notice that the first verbs listed in table 1 do not have any other meaning, while the last three verbs may occur as full-fledged verbs, inflected in all persons. When these verbs are so used, they have a slightly different meaning. For example, the stative verb *hó* ‘be well’ is used with the meaning ‘be good’ of a third person, as in (32a). The verb means

TABLE 1
VERBS IN AN ADVERBIAL CLAUSE

Verb	Meaning as Adverbial Modifier (Third-Person Only)	Semantics	Other Meaning (All Persons)
<i>nja_pu</i>	'be so/like that/that way'	Manner	—
<i>nja_nu</i>	'be so/like this/this way'	Manner	—
<i>ndě</i>	'be late, be the afternoon'	Temporal	—
<i>xüi</i>	'be the night'	Temporal	—
<i>xüdi</i>	'be the morning'	Temporal	—
<i>ngüt_á</i>	'be soon/quickly'	Temporal	—
<i>n'í^h_thö</i>	'be early/soon'	Temporal	—
<i>hó</i>	'be well'	Manner	'be good/well/fine'
<i>nts'edi</i>	'be strongly/firmly/fully/completely'	Extent	'be grave'
<i>wá_thö</i>	'be by foot'	Manner	'be barefoot' 'go about barefoot'

'be in a socially proper state' with the first and the second person, as shown in (32b):

- (32a) *nö= 'ó* *má=r* *zü'wé,*
 DEF.SG=DEM.SG.III 3.IMP.NP=SG worm/devil
hí-mí=hó-~~Ø~~= 'á
 NEG-(3.)IMP=be.good-3OBJ=3SG

'That one is the devil, he isn't good'.

- (32b) *xi=ñ-hó-g-a=gí*
 (3.)PRF-ST=be.proper-3OBJ-B=1OBJ

'I look fine/I'm properly dressed'.

5.2. Syntactic restrictions of the adverbial clause. In addition to lexical restrictions, the adverbial clause has other syntactic restrictions. The clause may accept negative polarity, but very rarely. Possible (elicited) examples are given in (33):

- (33a) *bi= 'wě^ht-~~Ø~~-i* *hí-már=ñ-hó-~~Ø~~*
 3.PST=sew-3OBJ-F NEG-(3.)IMP=ST-be.well-3OBJ

'She didn't sew it well' (lit., 'She sewed it, it wasn't well').

- (33b) *n-dí=pá=hé* *hí-már=wá_thö-~~Ø~~*
 IMP-1.PRS=go.PL=1PL.EX NEG-(3.)IMP=be.by.foot_DEL-3OBJ

'We didn't go by foot' (lit., 'We were going, it wasn't by foot').

Examples like those in (33), bearing negation in the adverbial clause, are rare in SI Otomi and do not occur in natural discourse, although they were

judged to be grammatical. Not all the verbs in table 1 can take negation, or they may not receive it in all circumstances. The verb *nts'edi* 'be strong' cannot receive negation, as shown in (34a). Similarly, the verb *hó* 'be well' can be negated in the Imperfect, as in (33a), but not in the Present, as shown by the impossibility of (34b):

- (34a) **xa=mí-máf-i* *hí-mí=nts'ed-i*
 INT=3.IMP=scream.F-F NEG-3.IMP=be.strong.F-F

Intended reading: 'He wasn't screaming loud'.

- (34b) **bi=hó^hk-∅-i* *híngi ∅=hó-∅*
 3.PST=make-3OBJ-F NEG (3.)PRS=be.well-3OBJ

Intended reading: 'She didn't make it well'.

When speakers wish to negate the content of the adverbial clause, they place negation on the predicate of the main clause. An example is given in (35):

- (35) *nó=r* *'béhñó* *hím-bi='wé^ht-∅-i*
 DEF.SG=SG woman NEG-3.PST=sew.3OBJ-F

xi=ñ-hó-∅

(3.)PRF=ST-be.well-3OBJ

'The woman didn't sew it well'.

In (35), negation occurs in the main clause and not in the adverbial clause. The negation placed in the main predicate has semantic scope over the predicate of the adverbial clause. In (35), what is being negated is not the sewing of the woman as such, but her ability to perform it in an appropriate fashion. This phenomenon is reminiscent of the negation of secondary predicates in (19a), repeated here for convenience as (36a). Adverbial phrases in Otomi show a similar behavior, as shown in (36b):

- (36a) *hín-n-dí=má* *(n)gǒxjí=thǒ*
 NEG-IMP-1.PRS=go naked=DEL

'I wasn't going around naked' (lit., 'I was going around, but not naked').

- (36b) *hím-bi='wé^ht-∅-a=muntxüi*
 NEG-3.PST=sew-3OBJ-B=by.night

'She didn't sew it by night' (lit., 'She sewed it, but not by night').

I take this phenomenon to be another piece of evidence that the adverbial clause is no longer an independent construct like the depictive clause but instead is lexicalized. In the next section, I propose that the adverbial clauses were at one time more independent, but then they became integrated into the main clause and the verb form was treated as a lexicalized unit, and was thus reanalyzed as an adverb.

TABLE 2
 LEXICALIZED ADVERBS FROM FORMER VERBS

Adverb		Verb	
<i>nts'edi</i>	'strongly/firmly/fully'	<i>nts'edi</i>	'be completely'
<i>nja_pu</i>	'that way'	<i>nja_pu</i>	'be that way'
<i>nja_nu</i>	'this way'	<i>nja_nu</i>	'be this way'
<i>ngǎt_á</i>	'quickly (in time)'	<i>ngǎt_á</i>	'be soon'
<i>xiidi</i>	'this morning'	<i>xiidi</i>	'be the morning'
<i>nsoni</i>	'quickly/fast (speed rate)'	<i>(n)xoni</i>	'be in a hurry'
<i>njǎnte</i>	'slowly'	<i>njǎnte=thǒ</i>	'be slow'
<i>nzá_tho</i>	'nicely'	<i>nzá_tho</i>	'be beautiful'
<i>nts'o_tho</i>	'badly'	<i>nts'o_tho</i>	'be ugly'
<i>gá'tho</i>	'fully/completely'	<i>gá'tho</i>	'be all'
<i>hǒnt=á</i>	'just/only'	<i>hǒnt=á</i>	'be just one'

5.3. The emergence of new adverbs from the adverbial clause. The adverbial clause is currently highly restricted as an adverbial expression in SI Otomi. However, I posit that this structure had a more productive profile in former stages of the language. There is some indirect evidence for this claim. A number of current adverbs in the language appear once to have been verbs functioning as predicates of an adverbial clause. Such predicates became frozen syntactically and were reanalyzed as adverbial phrases functioning as adverbial modifiers within the clausal boundaries of the main predicate. Because of this reanalysis, the verbal lexemes involved became adverbs. Consider the adverbs in table 2, which I take to have followed this development.

All of the adverbs in table 2 have a lexically related verb, which points to the diachronic relation between verbs and adverbs. The adverbs in table 2 lexicalized from a verb inflected in the Present tense in third-person subject, realized by a zero morpheme. Nowadays, only the first five adverbs in table 2 have a verbal correlate that can still function as the predicate of an adverbial clause. Under certain circumstances, there is no formal distinction between the adverb and the verb. Consider, for instance, (37):

- (37) *yá* *bi=t<h>ng-Ø-a=nú* *'rǎ* *txí-jǒ'i* *nú*
 P 3.PST=<IMPER>find.B-3OBJ-B=DEF IND.PL DIM-person REL
mí=t-'ag-Ø-i *nts'edi/Ø=nts'ed-i*
 (3.)IMP=IMPER-bury.F-3OBJ-F strongly/3.PRS=be.strong.F-F

'They found some poor people who had been **fully** buried'.

In (37), both the adverb and the verb express adverbial semantics of extent. One may analyze this example as presenting the adverb *nts'edi* 'strong/fully' or as having the adverbial clause *nts'edi* 'it's strong', which takes the event expressed by the main clause *mí t'agi* 'they had been buried' as its argument.

This ambiguity is only possible because the verb is inflected in Present tense with a zero morpheme. In other circumstances, there can be a formal distinction. Consider the examples in (38):

- (38a) *xa=mí=máf-i* *mí=nts'ed-i*
 INT=3.IMP=scream.F-F 3.IMP=be.strong.F-F
 'He was screaming loudly' (lit., 'He was screaming, it was strong').
- (38b) *xa=mí=máf-i* *řá-nts'edi*
 INT=3.IMP=scream.F-F INT-strongly
 'He was screaming loudly' (lit., 'He was screaming strongly').

(38a) shows a formal contrast: here, the lexeme receives an overt TAM marking and therefore must be analyzed as a verb. In (38b), the lexeme receives the intensive morpheme *řá-* 'very'. This prefix may occur with adverbs but not with active verbs (e.g., **mí=řá-ts'edi*, intended reading: 'it was very strong'); thus *nts'edi* in (38b) is an adverb. In other cases, the formal contrast is more straightforward. For example, in addition to the semantic difference, the inchoative verb *ngüt'á* 'be soon' always receives overt TAM morphology, as shown in (39a). This contrasts with the adverb *ngüt'á* 'quickly' in (39b), which never does:

- (39a) *xi=ngüt'á* *ga=pěng-i*
 3.PRF=be.early 1.IRR=turn.back-F
 'I'll come back soon' (lit., 'It's early, I'll come back').
- (39b) *ngüt'á* *da=n-ti* *nǒ=r* *jö'i*
 quickly 3.IRR=M-get.drunk DEF.SG=SG person
 'The man would quickly get drunk'.

In addition to having related verbs, all adverbs in table 2 show clear traces of having been former predicates. In 2.1, I pointed out that most Otomi verbs select an inflectional Free-form when occurring before another predicate—a circumstance which often coincides with a clausal boundary. I also pointed out that this Free-form is selected when the verb is followed by a number of adverbs. While there can be others, the adverbs in question correspond to the ones in table 2. The use of a Bound-form of the verb with these adverbs is ungrammatical, as illustrated in (40):

- (40a) *bi=h<y>ó^hk-Ø-i* *řá-nsoni=thǒ*
 3.PST=<NPS>make-3OBJ-F INT-quickly=DEL
 'He made it very quickly'.
- (40b) **bi=h<y>ó^hk-Ø-a=řá-nsoni=thǒ*
 3.PST=<NPS>make-3OBJ-B=INT-quickly=DEL
 Intended meaning: 'He made it very quickly'.

TABLE 3
LEXICALIZED ADVERBS IN OTHER ASPECTS

Adverb		Verb	
(x)tí-ndě	'this afternoon'	(x)tí=ndě	3.PRF.SUBJ=be.late
(x)tí-ntxüü	'tonight'	(x)tí=ntxüü	3.PRF.SUBJ=be.night
(x)tí-xúúdi	'tomorrow'	(x)tí=xúúdi	3.PRF.SUBJ=be.in.the.morning
nda-möni	'the day after tomorrow'	*n-da=möni	(IMP.3.IRR=be.the.third.day?)

(40b) shows that the verb cannot appear in its Bound-form when followed by the adverb *nsoni* 'quickly'. Only (40a), with a Free-form, is possible. This form is used elsewhere when the verb occurs at a clausal boundary or when it is followed by another predicate. I take this behavior as evidence that the adverbs in question are the lexicalized outcomes of former predicates in adverbial clauses.

A few temporal adverbs lexicalized from predicates inflected with other TAM but not the Present tense. The adverbs in question are given in table 3. The first three adverbs in table 3 have verbal correlates in SI Otomi: *ndě* 'be late', *xüüdi* 'be the morning', and *xüü* (NPS *n-txüü*) 'be night'. All can occur as predicates of an adverbial clause. I have reconstructed the last verb to which the remaining adverb is related. Like other temporal adverbs, the adverbs in table 3 also select a Bound-form of the predicate, as shown in (41). However, unlike other temporal adverbs, which were not predicates, the adverbs in table 3 can also select a Free-form, just like the ones in table 2. This is shown in (42):

(41a) *dá=hó^hk-Ø-a=mántxüü/mándě/máhó^h,mu*
 1.PST=make/prepare-3OBJ-B=last.night/yesterday/awhile.ago
 'I made it last night/yesterday/awhile ago'.

(41b) *ga=hó^hk-Ø-a=(x)tí-xúúdi/(x)tí-ndě*
 1.IRR=make/prepare-3OBJ-B=tomorrow/this afternoon
 'I'll make it tomorrow/this afternoon'.

(42a) *ga=hó^hk-Ø-i* (x)tí-xúúdi/(x)tí-ndě, etc.
 1.IRR=make/prepare-3OBJ-F tomorrow/this afternoon
 'I'll make it tomorrow/this afternoon'.

(42b) ??*dá=hó^hk-Ø-i* *mántxüü/mándě/máhó^h,mu*
 1.PST=make/prepare-3OBJ-F last.night/yesterday/awhile.ago
 Intended reading: 'I made it last night/yesterday/awhile ago'.

Adverbs in Otomi may be arranged in a continuum according to which inflectional form of the verb they trigger. Table 4 shows this continuum with a number of representative adverbs (see also table 6 in Palancar 2004).

TABLE 4
ADVERBS ARRANGED ACCORDING TO THE SELECTION OF INFLECTIONAL FORM IN THE VERB

Semantics	I		II		III		IV	
	F-Form / *B-Form	++F-Form / B-Form	++F-Form / B-Form	B-Form / F-Form	++B-Form / *(??)F-Form			
Manner	<i>ngüt'á</i> 'quickly'							
	<i>njánte</i> 'slowly'							
	<i>nsoni</i> 'quickly/fast'							
	<i>má-n(t)soni</i> 'in a hurry'		<i>nzátho</i> 'nicely'		<i>njanu</i> 'this way'			
			<i>nts'ótho</i> 'badly'		<i>nja(')pu</i> 'that way'			
Extent	<i>nts'edi</i> 'strongly, fully'							
	<i>gá'tho</i> 'fully/completely'							
	<i>hónt'á</i> 'just/only'							
Place	<i>má-nthĩ</i> 'outside'					<i>nthĩ</i> 'outside (schematic)'		
	<i>má-mbó</i> 'inside'					<i>mbó</i> 'inside (schematic)'		
			<i>nfrénte</i> 'opposite'					
			<i>n'a-ládo</i> 'beside'					
			<i>más (u)ngé^hnu</i> 'in front'					
Time					<i>xfi-xúdi</i> 'tomorrow'			
						<i>nu'wá</i> 'here'		
						<i>nunu</i> '(right) there'		
						<i>nu'pu</i> '(over) there'		
						<i>ya'pu</i> 'far'		
						<i>xúdi</i> 'this morning'		

	<i>má-nxáá</i>	'this morning'
	<i>mú-nxáá</i>	'in the mornings'
	<i>mú-xíá</i>	'in the mornings'
	<i>má-nxú</i>	'last night'
	<i>mú-nxú</i>	'by night'
	<i>má-ndě</i>	'yesterday'
	<i>mú-ndě</i>	'in the afternoons'
	<i>mú-mpá</i>	'by day'
	<i>má-m'el'o</i>	'a long time ago'
	<i>má-hó^{ph}, mú</i>	'awhile ago'
	<i>má-ngúndě</i>	'the day before yesterday'
	<i>má-márite</i> , etc.	'on Tuesday(s), etc.'
<hr/>		
<i>xíí-nxú</i>	'tonight'	
<i>xíí-ndě</i>	'this afternoon'	
<i>nda-móni</i>	'in three days'	

Table 4 presents four different stages in this continuum, arranged in four columns. Adverbs in boldface still have a verbal correlate in the lexicon. The adverbs in the first column are the type that select a Free-form in the verb but not a Bound-form. Here we see most of the adverbs in table 2, together with de-adverbial adverbs (derived from other adverbs with the prefix *ma-* or *mu-*). The second type of adverb commonly takes a Free-form, but the use of a Bound-form is not completely ungrammatical. Here we see, for example, the adverb *nzátho* ‘nicely’ from table 2. The other adverbs in this column are loanwords from Spanish. The third type is adverbs which normally take a Bound-form, but may also take a Free-form. Here we find two similar manner adverbs from table 2 and all the adverbs from table 3. Finally, the last type corresponds to adverbs which do not take a Free-form in natural circumstances. Nevertheless, the use of a Free-form with such adverbs would not be as ungrammatical as using a Bound-form with the first type of adverb. This has to do with the fact that all lexemes in table 4 are adverbs heading an adjunct phrase. Verbs in SI Otomi are more sensitive to arguments in the selection of a Bound-form than they are to adjuncts (see Palancar 2004 for more details). The grammatical distribution laid out in table 4 broadly indicates that manner and extent adverbs in SI Otomi descend from an adverbial clause whose function was mainly devoted to the expression of such categories. Notice that only a few temporal adverbs have this origin—and none of the place adverbs.

6. Conclusion: Comparing juxtaposed clauses. In this article, I have demonstrated that SI Otomi uses two types of juxtaposed adjunct clauses for a variety of semantic functions: i.e., to express depictive semantics and a miscellanea of other adverbial meanings. I want to conclude this article by proposing that these two types of clauses are indeed two different subtypes of one and the same schematic structure in Otomi. In other words, Otomi uses the general syntactic frame of a juxtaposed clause to express adjunct relations of two kinds: depictive or adverbial. I call this clause the “General Adjunct Clause,” because it is the structural equivalent in Otomi of the General Adjunct Construction found in Olutec (see **1** above). The main difference between Otomi and Olutec in this respect is that the Otomi structure is not a unit having an integrated position within a clause.

The General Adjunct Clause has the following characteristics: (a) it is juxtaposed to the main clause; (b) like other adjuncts in the language, it is commonly placed after the main clause; (c) it is an intransitive clause; (d) it may receive TAM coherent with the main clause, but it often appears in a neutral aspect; and (e) it shares an argument with the main clause, except with adverbial clauses, which take the whole event expressed in the main clause as their own argument.

Within this aspectual/temporal frame, the main clause and the juxtaposed adjunct clause form a larger unified structure for the presentation of information. This syntactic dependency increases with the possibility of sharing arguments between them, but it is the adverbial clause, which modifies the very internal core of the main clause, and not one of its arguments, that naturally displays a higher degree of integration with it (recall the semantic scope of negation shown in example 35 in 5.2). This degree of integration gave rise to the reanalysis of former verbal phrases as adverbial phrases, and the consequential lexicalization of the verbal lexemes into adverbs.

On the other hand, much of the semantic interpretation of this type of clause may not be due to the semantics of the construction per se (as it would be in languages with dedicated constructions for specific depictive or adverbial semantics), but rather to the pragmatic implications of the relation between the state of affairs encoded in the clauses involved. For example, the depictive reading is favored when neutral TAM is used as a way to mark event simultaneity. Similarly, sharing arguments in the depictive clause can be viewed as motivated by discourse pressure to keep topic continuity across participants involved in a series of events; one way to ensure this, if there is topic continuity, is to have the second clause constructed with a one-place predicate.

We have seen so far that this General Adjunct Clause is a juxtaposed clause that encodes adjunct modification at a higher level than the clause. However, juxtaposition as such is not a syntactic strategy circumscribed to adjuncts in Otomi, because the language may also use juxtaposed clauses as complements of a number of matrix verbs, as shown in (43):

- (43) *kási yá bi=<m>úd-i [bi=ne^hk-i]*
 almost P 3.PST=<NPS>start.F-F 3.PST=get.light-L
 [*xi=ñ-hó-θ=pya*]
 (3.)PRF=ST-be.good-3OBJ=now

‘It almost started getting light well now’.

(43) has two types of juxtaposed clauses: the complement clause *bi ne^hki* ‘it got light’, which is a clausal argument of the main predicate *fúdi* ‘start’, and the adverbial clause *xi ñhó(pya)* ‘it is well (now)’.

The study of complementation in Otomi deserves further study. Suffice it to say here that a juxtaposed complement clause behaves differently from a juxtaposed adjunct clause. For example, it may not be negated, as shown by the impossibility of (44*b*) when contrasted with (44*a*); it does not appear in neutral TAM, as indicated by the ungrammaticality of (45*b*), but may either copy the TAM of the main clause, as in (43), or appear in the Irrealis, as in

(45a). Also, a juxtaposed complement clause may freely be transitive, as shown in (45a), whereas a juxtaposed adjunct clause is always intransitive:

- (44a) *him-bi*=<*m*>*úd-i* [*bi*=*ne^hk-i*]
 NEG-3.PST=<NPS>start.F-F 3.PST=get.light-L
 ‘It didn’t start getting light’.

- (44b) **bi*=<*m*>*úd-i* [*him-bi*=*ne^hk-i*]
 3.PST=<NPS>start.F-F NEG-3.PST=get.light-L

Intended reading: either ‘It didn’t start getting light’ or ‘It started not getting light’.

- (45a) *kwándo* *dí*=*ne-∅* [*ga*=*hǒ^hk-∅-a=n’ǎ* *ma*
 when 1.PRS=want-3OBJ 1.IRR=make-3OBJ-B=one 1POSS
 tráhé=hé]
 costume=1PL.EX

‘When we want to make one of our costumes’.

- (45b) **dí*=*ne-∅* [*dí*=*hǒ^hk-∅=hé*]
 1.PRS=want-3OBJ 1PRS=make-3OBJ=1PL.EX

Intended meaning: ‘We want to make it’.

A summary of such formal contrasts is given in table 5.

Juxtaposition, as instantiated by both the General Adjunct Clause and the type of complement clause shown in (43) and (45), should be regarded, in turn, as a grammatical strategy related to a more general syntactic pattern of clausal sequence or clausal serialization, which is also observed in parataxis. Like juxtaposition, parataxis is a frequent discourse strategy in narrative texts and is employed when the speaker is encoding a series of events in sequence. This pattern is illustrated in (46):

- (46a) *bi*=<*b*>*ón-i*, *bi*=*n-txad-i*,
 3.PST=<NPS>exit-F 3.PST=ANTI-pray-F
 bi=*h<ñ>ǒx-∅-a=’nar* *pálángáná* *kó* *ya*
 3.PST=<NPS>carry.S-3OBJ-B=IND.SG washbasin with PL
 ndoní
 flower

‘They go out, pray, carry a washbasin with flowers’.

- (46b) *bi*=*m=pühn-i*, *hím-bi*=<*z*>*a-∅=r* *hñé^hthi*
 3.PST=M-forget-F NEG-3.PST=<NPS>eat-3OBJ=SG medicine

‘He forgot to take the medicine’ (lit., ‘He forgot, he didn’t eat the medicine’).

TABLE 5
COMPARISON OF JUXTAPOSED COMPLEMENT AND ADJUNCT CLAUSES

	General Adjunct Clause		Complement Clause of the Type in (43)
	Depictive Clause	Adverbial Clause	
(a) The clause is a juxtaposed clause.	Yes	Yes	Yes
(b) The clause shares arguments with the main clause.	Yes	No	Yes
(c) The clause copies the TAM of the main clause.	No	No	Yes
(d) The clause may be marked with neutral TAM.	Yes	Yes	No
(e) The clause is an intransitive clause.	Yes	Yes	No
(f) The clause receives negation.	Yes	(Yes/No)	No

In (46a), there are three clauses in a paratactic relation expressing a sequence of events: exiting, praying, and carrying. A similar phenomenon is shown in (46b), where the intransitive middle verb *mpühni* ‘forget’ expresses a cognitive event of forgetting, and the following clause the event that followed as a consequence of the forgetting. Similarly, interpreting whether a following clause is juxtaposed or paratactic appears to rely heavily on pragmatic implicature. The intricate relation between parataxis and juxtaposition in the encoding of a sequence of events, regardless of hierarchical syntactic relations, needs to be treated as an independent object of study.

The phenomenon I have studied in this paper and the theoretical perspective I have taken are novel in the literature of other Otomanguean languages and in the languages of Mesoamerica in general. Aissen and Zavala (forthcoming) offers an invaluable compilation of articles on the treatment of depictive semantics, mainly by secondary predication constructions, in various languages of the area. Less is known about the relation of such constructions to the structures encoding adverbial semantics. I hope that this study will encourage other linguists working on indigenous languages of the Americas to contribute their findings to this fascinating area of the grammar.

REFERENCES

- AISSEN, JUDITH. 2004. Notes from the workshop “Taller sobre Predicación.” CIESAS–Sureste, August 9–14, 2004, San Cristóbal de las Casas, Chiapas, Mexico.
- AISSEN, JUDITH, AND ROBERTO ZAVALA, eds. Forthcoming. *Estudios sobre la Predicación Secundaria en Mesoamérica*. Mexico, D.F.: Centro de investigaciones y Estudios en Antropología Social.
- CRYSTAL, DAVID. 1996. *Discover Grammar*. London: Longman.
- HEKKING, EWALD. 1995. El otomí de Santiago Mexquititlán: Desplazamiento lingüístico, préstamos y cambios gramaticales. *Studies in Language and Language Use* 17. Amsterdam: Institute for Functional Research into Language and Language Use.

- _____. 2002. Desplazamiento, pérdida y perspectivas para la revitalización del hñãñho. *Estudios de Cultura Otomame* 3:221–48. Universidad Nacional Autónoma de México.
- HEKKING, EWALD, AND SEVERIANO ANDRÉS DE JESÚS. 1984. *Gramática otomí*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- HERNÁNDEZ CRUZ, LUIS; MOISÉS VICTORIA TORQUEMADA; AND DONALDO SINCLAIR CRAWFORD, eds. 2004. *Diccionario del hñãñhu (otomí) del Valle de Mezquital, Estado de Hidalgo. Vocabularios Indígenas 45*. Mexico, D.F.: Instituto Lingüístico de Verano.
- HIMMELMANN, NIKOLAUS P., AND EVA SCHULTZE-BERNDT. 2005. Issues in the syntax and semantics of participant oriented adjuncts: An introduction. *Secondary Predication and Adverbial Modification: The Typology of Depictives*, ed. Nikolaus P. Himmelmann and Eva Schultze-Berndt, pp. 1–68. Oxford: Oxford University Press.
- LASTRA, YOLANDA. 1992. *El otomí de Toluca*. Mexico, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- _____. 1997. *El otomí de Ixtenco*. Mexico, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- PALANCAR, ENRIQUE L. 2004. Verbal morphology and prosody in Otomi. *IJAL* 70:251–78.
- _____. 2006a. Intransitivity and the origins of Middle voice in Otomi. *Linguistics* 44:613–43.
- _____. 2006b. Property concepts in Otomi: A language with no adjectives. *IJAL* 72:325–66.
- _____. 2008. The emergence of active/stative alignment in Otomi. *The Typology of Semantic Alignment*, ed. Mark Donohue and Søren Wichmann, pp. 357–79. Oxford: Oxford University Press.
- _____. Forthcoming. *Semántica depictiva sin predicación secundaria: El caso de la lengua otomí*. *Estudios sobre la Predicación Secundaria en Mesoamérica*, ed. Judith Aissen and Roberto Zavala.
- SCHULTZE-BERNDT, EVA, AND NIKOLAUS P. HIMMELMANN. 2004. Depictive secondary predicates in crosslinguistic perspective. *Linguistic Typology* 8:59–130.
- VOIGTLANDER, KATHERINE, AND ARTEMISA ECHEGOYEN. 1985. *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. Mexico, D.F.: Instituto Lingüístico de Verano. [First impression, without author names, 1979.]
- WALLIS, ETHEL EMILIA. 1956. *Diccionario Castellano–Otomí, Otomí–Castellano*. With the collaboration of Victor Manuel Arroyo. *Cuadernos del Valle de Mezquital* 1, no. 1. Itzmiquilpan, Hidalgo, Mexico: Patrimonio Indígena del Valle de Mezquital and Instituto Lingüístico de Verano.
- ZAVALA MALDONADO, ROBERTO. 2004. Adverbios y predicados secundarios depictivos en oluteco. *Memorias del VII Encuentro Internacional de Lingüística en el Noroeste*, ed. Isabel Barreras Aguilar and Mirna Castro Llamas, pp. 293–314. Hermosillo, Sonora, Mexico: Editorial UniSon.

VERBAL MORPHOLOGY AND PROSODY IN OTOMI¹

ENRIQUE L. PALANCAR

UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

Verbs in Otomi (Otopamean; Otomanguean) have two distinct inflectional forms whose selection is conditioned by an interesting combination of prosodic and syntactic factors. In this article, I describe these two inflectional forms and the factors that motivate their usage. The phenomenon studied here is not common cross-linguistically, given that it constitutes a rare case where different levels of linguistic organization are significantly interconnected at the time a clause is uttered.

[KEYWORDS: verbal morphology, prosody, Otomi, Otomanguean]

1. Introduction. Most verbs in Otomi have two distinct inflectional forms, a “free form” and a “bound form.” The selection of these two forms is conditioned by a combination of prosodic and syntactic factors. The goal of this article is to describe these two inflectional forms together with the factors that motivate their selection and conditioning. The present analysis, however, does not intend to provide a workable formal account of the phenomenon.

Otomi is a Mesoamerican language spoken in Central Mexico.² The language belongs to the Otopamean family, which is the northernmost branch of Otomanguean. Otomi has great dialectal diversification, but the phenomenon studied here appears to be dia-systemic, given that it occurs in some way or another in (almost) all dialects. The data in this article come from my research on the dialect spoken in San Ildefonso Tultepec, a village located in the south of the state of Querétaro, referred to here as SI Otomi.³

¹ I want to thank the anonymous referees for the very useful suggestions they made on an earlier version of this paper, and especially Albert Bickford, who helped immensely to improve the paper. I also want to thank Keren Rice for all her help. All errors and deficiencies remain my responsibility only.

² Using English vowels, the name *Otomi* can be pronounced in English as [oto'mi:]. The name corresponds to Spanish [oto'mi], spelled with an accent as *Otomí*. I use the term *Otomi*—without the accent—following Suárez (1983) as well as Leon and Swadesh (1949), Hess (1968), Voigtlander and Bartholomew (1972), Bartholomew (1973), and Bernard (1974); other authors, like Andrews (1993) and Lastra (1998), prefer *Otomí*, with the accent. The native noun for the language in San Ildefonso Tultepec Otomi is *hñóñhó /hñó'ñhó/*. This noun derives from the verb *ñóñhó /ñó'ñhó/*, which, according to Hekking (2002), means ‘speak well’.

³ IPA conventions have been used in the transcription of Otomi, with these exceptions: <a> represents central /a/; <ñ> palatal /ɲ/; <š> postpalatal /ʃ/; <č> affricate /tʃ/; and the subscript left hook <V> represents nasality.

In **2**, I introduce some relevant grammatical facts about the language and then (in **3**) I introduce both inflectional forms as they are used in SI Otomi. Section **4** provides a morphological characterization of the forms, while **5** is devoted to exploring in more detail the distribution of a bound form. Section **6** explains a number of idiosyncrasies in the use of the forms presented, while in **7**, I compare the situation of SI Otomi to other Otomi dialects.

2. Some facts about San Ildefonso Otomi verbal morphology.

Morphologically speaking, most native verbs in Otomi have a complex stem consisting of a purely morphological root and a suffixed stem formative.⁴ The pattern is illustrated in (1). For convenience, the morpheme boundary between the root and the stem formative is represented by a plus sign:⁵

(1a) $p\phi+d-$
ROOT+STEM FORMATIVE

‘know’

(1b) $?w\epsilon+d-$ (intr)

‘sew (activity)’

(1c) $?w\epsilon+t-$ (tr)

‘sew (accomplishment)’

(1a) illustrates the verbal stem $p\phi d-$ ‘know’. In this stem, the morph $p\phi+$ is the morphological root, while $+d$ is a stem formative suffixed to this root. The stem in (1b) shares the same stem formative but has a different root. In contrast, the stem in (1c) has the same root as in (1b) but has the stem formative $+t$. The verbal stems in (1) represent bipartite formations that have lexicalized synchronically in most cases. In this sense, stem formatives such as $+d$ and $+t$ in the stems in (1) are semantically opaque in Otomi; they reflect petrified derivational morphology. Similarly, roots in such formations are equally obscure semantically in most cases. Broadly speaking, the verbs in (1) reflect the same pattern as the one found in some English verbs borrowed from French, i.e., *pre+fer*, *pro+fer*, *in+fer*, etc. In this light, the lexeme of bipartite verbs is expressed by the two morphs together, rather than by just the root.

At times, clear morphological patterns involving the stem formatives are discernible. For example, verbs with the stem formative $+d$ often co-occur in lexical pairs with verbs with $+t$, as in (1b) and (1c). In such cases, the verb with $+d$ commonly denotes an activity, while the verb with $+t$ expresses the

⁴ In Palancar (forthcoming), this pattern is found in 83.5% of a sample of 310 verbs.

⁵ Otomi is a tonal language, but tone is represented here in function words only, where it has been identified to have a morphological function (see n. 11).

same action but directed to a human participant, i.e., $\acute{s}\text{ɔ}+d-$ ‘pray’ ~ $\acute{s}\text{ɔ}+t-$ ‘pray to someone’; $th\text{ɔ}+d-$ ‘answer’ ~ $th\text{ɔ}+t-$ ‘answer someone’; $k\acute{\alpha}+d-$ ‘wear something below waist’ ~ $k\acute{\alpha}+t-$ ‘dress someone below the waist’; $\text{?}\acute{\alpha}+d-$ ‘hear’ ~ $\text{?}\acute{\alpha}+t-$ ‘obey’; etc.⁶ The language has a substantial number of patterns like this (see Palancar [forthcoming] for SI Otomi and Voigtlander and Bartholomew 1972 for Sierra Otomi). Verbs having complex stems like in (1) are relevant because the forms treated in this article occur (mainly) with such verbs.

Additionally, a large number of Otomi verbs have at least two distinct stems. One stem is used for the third person, when the verb inflects in tenses other than the present and the imperfect, and another is used elsewhere.⁷ I call the first the “nonpresent stem” (NPS)⁸ and the latter the “base stem,” since it is a default stem. Consider the stems in (2) for the verb meaning ‘take somebody away’:

- (2a) Base stem *tsi+ts’-*
 (2b) Nonpresent stem *zi+ts’-*

The nonpresent stem is morphologically built on the base stem by means of a number of internal changes. Such changes may involve the voicing of an initial voiceless consonant (except for ejective series). The nonpresent stem *zi+ts’-* in (2b) is built on the base stem *tsits’-* in (2a) by means of consonant mutation involving voicing. Internal changes in nonpresent stems are represented by angle brackets, e.g., $\langle z \rangle i+ts’-$. In other cases, the nonpresent stem is formed by the application of a metathesized palatal prefix when the base stem has an initial glottal (e.g., $ho+n-$ ~ $h\langle j \rangle o+n-$ ‘look for’).

The language has zero object anaphora. Verbs inflect for both tense/aspect and the person of the subject by means of a complex paradigm of function words that always precede the verbal stem. Finally, SI Otomi has SVO as the pragmatically neutral order when a transitive clause has two overt NPs. Nonetheless, responding to pragmatic requirements, orders like OVS and VSO are also possible. When the transitive clause has a covert NP object, the order VS is rather common. As for in transitive clauses, VS is the norm.

3. Free forms and bound forms in Otomi. Verbs with complex stems have two relevant inflectional forms that can be called the “free form” and the “bound form.” We start with a free form (F-form). Broadly speaking, an F-form is used when a pause follows the verb. In this sense, the F-form of

⁶ Voigtlander (1998) is a useful attempt to reconstruct the meaning of these stem formatives. For some of them, traces of spatial meaning can still be observed.

⁷ The nonpresent stem is also used for the second person when the base stem of a verb begins with a glottal.

⁸ Kaufman (1988) refers to this morphology as the “potential/complete thematic state.”

a verb is a word form that can be uttered in isolation, and it is employed for this reason as the citation form of the lexeme. Examples in (3) illustrate F-forms. Following Katamba (1992), a prosodic pause is indicated by a double pipe (||):⁹

(3a) *ja bì theg-e* ||
CP 3.PST end-F

‘It’s over’. (Lit., ‘It ended’.)

(3b) *siempre mí pa Mʔondɔ mí m-pɛph-i* ||
always 3.IMP go Mexico 3.IMP DTR-work-F

‘They always went to Mexico to work’.

In (3a) and (3b), the F-form of the intransitive verbs *theg* ‘finish’ and *mpephi* ‘work’ precedes a prosodic pause. Nonetheless, an F-form cannot be characterized simply as a “pausal form” in prosodic terms because an F-form is also used when a verb precedes a clausal boundary, regardless of clausal prosody. (4) illustrates this point. Clausal boundaries are represented by square brackets:

(4) [*?in-dà ts’ɔn-i* [*porke ja bì <g>wadi*]] ||
NEG-3.FUT go.off-F because CP 3.PST NPS.finish-F
[*bì n-thin-i*]] ||
3.PST DTR-boil/ferment-F

‘It won’t go off because it fermented’. (Lit., ‘It finished fermenting’.)

(4) shows three F-forms. The verbs *ts’ɔni* ‘go off’ and *kwadi* ‘finish’ were phrased within the same intonational phrase; that is, no pause occurred between them, yet they occur in their respective F-forms. A pause is possible after *ts’ɔni* ‘go off’ just before the subordinate clause, but it is unnatural after *kwadi* ‘finish’ given that the two verbs *bì gwadi bì nthini* ‘it finished fermenting’ form a syntactic construct with semantic coherence. (3) and (4) suggest that an F-form is possible in SI Otomi in the following situations: (a) when the verb precedes a phrasal boundary realized by a prosodic pause; (b) when the verb precedes a clausal (syntactic) boundary; and (c) when both

⁹ Abbreviations are as follows: - morpheme boundary in source; < > nonconcatenative morpheme boundary; . nonmorpheme boundary in source; B bound form; CP complete particle; DEF definite; DEM demonstrative; DIM diminutive; DTR detransitivizer; F free form; FUT future; IMP imperfect; IND indefinite; INTER interrogative; INTR intransitive; LOC locative preposition; NEG negative; NPM nonpresent (stem) marker; NPS nonpresent stem; OBJ object; P.PERF pluperfect; PL plural; POSS possessive; PRES present; PST past; PURP purposive; QUOT quotative; SG singular; SI San Ildefonso Tultepec; SIM simultaneous; SUBJ subject; TR transitive.

clausal and prosodic boundaries coincide, as, for example, the last verb in (4), *nthini* ‘boil, ferment’, which occurs at both boundaries.

Apart from F-forms, verbs with complex stems also have a “bound form” (B-form). Broadly speaking, a B-form occurs in the context where an F-form is not required. In other words, a B-form is used when the verb does not immediately precede a clausal or phrasal boundary. (Things are not quite this simple, as later discussion will make clear, but this statement will suffice for now.) Additionally, a B-form is a cliticized form that cannot be uttered in isolation. Examples of B-forms are given in (5):

- (5a) [*?na=r* *či* *ču* *mí* *?ñe-?=joho* *ja* *?beto=?na*]
 IND=SG DIM lady 3.IMP have-B=two PL grand.child=QUOT

‘It’s said that an old lady had two grandchildren’.

- (5b) [*ntonse* *nu?ja* (*š*)*ta* <*g*>*wad-i* [*dà*]
 then now 3.P.PERF NPS.finish-F 3.FUT

n-thing-a=no=r *sej*]]

DTR-boil/ferment-B=DEF.SG=SG pulque

‘Then, when the *pulque* is fermented (. . .).’¹⁰ (Lit., ‘Has finished fermenting’.)

In (5a), the B-form of the transitive verb *?ñehē* ‘possess’ immediately precedes the NP object *joho ja ?beto* ‘two grandchildren’. The brackets representing clause boundaries indicate that the verb and the following NP belong to the same clause. In (5b), the B-form of the intransitive verb *nthini* ‘boil, ferment’ precedes the NP subject *nor sej* ‘the pulque’. In this sense, the examples show that the use of a B-form is not conditioned by the syntactic function of a given following constituent; rather, I claim it is used when the verb occurs intraclausally and no pause is placed after it. The following examples provide further evidence for this claim:

- (6a) [*bì* <*m*>*a*] [*ba* *k’q’s-a=ká=r*]
 3.PST NPS.go 3.PST.SIM go.to.see.someone-B=LOC.3POSS=SG

ngu=r *pale*]
 house=SG grandfather

‘They went to see her at their grandfather’s’.

- (6b) [*ntonse* *bì* *sigi* [*bì* *ñ?eng-a=man?agi=?na*]]
 then 3.PST go.on 3.PST play-B=once.again=QUOT

‘It’s said that he kept on playing once again’.

¹⁰ *Pulque* is the Spanish name given to the fermented alcoholic drink extracted from a *maguey* (a large cactus belonging to the species *amarilidacea Agave*).

In (6a), we have the B-form of the verb *k'ɔʔts'i* 'go to see someone' preceding a spatial PP headed by the preposition *ka*, a generic locative. Similarly, the B-form of the verb *ñʔeni* 'play' is used in (5b) before the temporal adverb *manʔagi* 'once again'. Since a B-form may also be used before a PP and an adverb, these examples show that the use of such a form is not motivated in principle by the syntactic function of the constituent following the verb.

F-forms and B-forms are not interchangeable. In other words, the use of one in a context where the other is expected (usually) results in ungrammaticality. This is shown by the ungrammaticality of (7):

- (7a) **[ʔin-dà ts'ɔng-a=[porke ja bi <g>wa-ʔ=[bi*
 NEG-3FUT go.off-B=because CP 3.PST NPS.finish-B=3.PST
n-thing-a]] ||
 DTR-boil/ferment-B

Intended meaning: 'It won't go off because it finished fermenting'.

- (7b) **[ntonse nuʔja (ʃ)ta <g>wad-i [dà*
 then now 3.P.PERF NPS.finish-F 3.FUT
n-thin-i no=r se]]
 DTR-boil/ferment-F DEF.SG=SG pulque

Intended meaning: 'Then, when the pulque is fermented (. . .)'.

In (7a), we have the ungrammatical rendering of (4) with B-forms, where only F-forms are possible given that the verbs precede a clausal boundary. The same applies to (7b), an impossible alternative to (5b). Here, the use of the F-form *nthin-i* is ungrammatical because the verb occurs intraclausally.

4. The morphology of the two forms. The morphology of an F-form and a B-form contrasts in an interesting way. A number of verbs inflect for both forms by means of a replacive morph. For such verbs, the default morphological exponent of an F-form is the suffix *-i* (e.g., *tse+k-i* 'hack'), while the default exponent of a B-form is the suffix *-a* (e.g., *tse+k-a* 'hack').¹¹ Additionally, an F-form may also be realized by the allomorph *-e*, which occurs when the stem has the vowel /ə/ or /e/ (at times also /ɛ/) and has the stem formatives *+t*, *+d*, *+k*, *+g*, or *+ts*' (e.g., *ʔə+t'-e* 'do'; *he+t-e* 'dress someone above waist'; *je+ts'-e* 'turn upside down'; etc.). A number of verbs that illustrate this pattern are given in table 1.

Verbs with the nasal stem formative *+n* fall into two morphological classes. One class receives a velar augment while the other may not. The former can

¹¹ Tone appears to play a significant role in the realization of the forms. The suffix *-a* realizing a B-form always receives low tone. F-forms realized by the suffix *-i* often receive high tone, especially preceding a prosodic boundary. This topic deserves further research.

TABLE 1

F-Form	B-Form	
<i>tsek-i</i>	<i>tsek-a=</i>	'hack'
<i>?ek-e</i>	<i>?ek-a=</i>	'comb'
<i>šond-i</i>	<i>šond-a=</i>	'breed'
<i>peng-i</i>	<i>peng-a=</i>	'return'
<i>šaš-i</i>	<i>šaš-a=</i>	'bother, disturb'
<i>nthot-i</i>	<i>nthot-a=</i>	'marry'
<i>het-e</i>	<i>het-a=</i>	'dress someone above waist'
<i>?at'-e</i>	<i>?at'-a=</i>	'do'
<i>kit'-i</i>	<i>kit'-a=</i>	'enter'
<i>tchm-i</i>	<i>tchm-a=</i>	'break (i.e., a volume)'

TABLE 2

Class	F-Form	B-Form	
I	<i>ñ?e+n-i</i>	<i>*ñ?e+n-a=</i>	<i>ñ?e+ng-a=</i> 'play'
II	<i>thi+n-i</i>	<i>thi+n-a=</i>	<i>thi+ng-a=</i> 'boil' (tr)

TABLE 3

Class	F-Form	B-Form	
I	<i>?ə+ts'-e</i>	<i>?ə+s-a=</i>	'detach'
II	<i>hə+ts'-i</i>	<i>hə+š-a=</i>	'carry (generic)'

be called class I and the latter class II. An example of each class is given in table 2. As shown in table 2, the stem of class I verbs receives a velar augment to construct the B-form. This is not a requirement for class II verbs, only a few of which may receive both morphologies.

Like the verbs in table 2, verbs with the stem formative *+ts'* also fall into two classes when they receive affixed material, as, for example, the exponent of a B-form. This is shown in table 3. Table 3 shows that verbs whose stem have the formative *+ts'* fall into two classes: verbs of class I mutate the ejective affricate of the stem formative into an /s/,¹² while those of class II mutate it into the postalveolar /ʃ/.

A number of verbs deviate from the pattern presented so far. Instead of an *-a*, verbs with the stem formatives *+i*, *+d*, *+g*, *+ph*, and *+h* receive a glottal stop as an exponent of a B-form. The stems of such verbs drop their formatives when the verb is inflected in B-form. The verbs in table 4 illustrate this

¹² Verbs of class I may at times surface as verbs in table 1, that is, without further stem adjustments.

TABLE 4

F-Form	B-Form	
<i>ne+i</i>	<i>ne-?</i> =	'dance'
<i>tsi+d-i</i>	<i>tsi-?</i> =	'hang' (intr)
<i>pe+d-e</i>	<i>pe-?</i> =	'tell'
<i>hc+g-i</i>	<i>hc-?</i> =	'let'
<i>the+g-e</i>	<i>the-?</i> =	'end'
<i>ma+ph-i</i>	<i>ma-?</i> =	'shout'
<i>ko+h-i</i>	<i>ko-?</i> =	'keep something'

TABLE 5

F-Form	B-Form	
<i>tsə+hə</i>	<i>tsə-?</i> =	'arrive (toward deictic center)'
<i>ki+hi</i>	<i>ki-?</i> =	'enter (deictic center inside)'
<i>tsi+hi</i>	<i>tsi-?</i> =	'bring (animate)'
<i>ša+ha</i>	<i>ša-?</i> =	'bathe someone'
<i>?nē+hē</i>	<i>?nē-?</i> =	'have'
<i>tə+hə</i>	<i>tə-?</i> =	'win'
<i>tʉ+hʉ</i>	<i>tʉ-?</i> =	'sing'

pattern. Additionally, a large group of verbs have the stem formative *+hV* in their F-form. The vowel of this formative copies the vowel in the morphological root. Like verbs in table 4, these verbs drop the formative entirely when they inflect in B-form (and they also receive a glottal stop suffix). Examples of this type of verb are given in table 5.

4.1. Restrictions. There are some morphophonological restrictions to the pattern presented so far. Verbs with the stem formatives *+hni* and *+šni* deviate from the pattern because they lack a morphologically distinct B-form. Examples of such verbs are *k'a+hni* 'shoot', *pe+hni* 'send', *pi+hni* 'peck, nibble', *tʉ+šni* 'get tangled', etc. The final /i/ in such verbs is treated as part of the stem, given that it does not have a contrastive morphological function. However, a verb ending in *+hni* occasionally surfaces as *+hna* in a context where a B-form would be expected, a fact which suggests that speakers have analogically reanalyzed such verbs as *+hn-i/a*, attending to the general pattern presented in table 1.

Other restrictions are lexical. Stative verbs, which qualify for the morphological rules above, do not inflect for the forms in contexts where they are otherwise expected.¹³ In other words, the rules that determine whether a verb

¹³ In addition to this feature, stative verbs show other inflectional idiosyncrasies that characterize them as a well-defined subclass (Palancar 2004).

appears in its F-form or B-form treat stative verbs differently (see 6.3 for more details). For example, such verbs do not take a B-form when they are part of the same clause as the following NP. For this reason, the final /i/ in these verbs has been treated as part of the stem. An example of the stative verb is given in (8):

- (8a) *ma ngu hingi Ø t'aši* ||
 1POSS house NEG 3.PRES be.white

‘My house isn’t white’.

- (8b) *hingi Ø *t'aš-a=/t'aši ma ngu*
 NEG 3.PRES be.white-B=/be.high 1POSS house

‘My house isn’t high’.

(8) shows the stative verb *t'aši* ‘be white’ in the natural context of an F-form, that is, preceding a clausal and/or prosodic boundary. However, (8b) shows that the potential morphological B-form of such a verb (i.e., **t'aš-a=*) is not used in the expected context. Other verbs of this (open) class include *ʔuši* ‘be salty’, *thengi* ‘be green’, *šini* ‘be thin’, *pots'i* ‘be worn out’, *šidi* ‘be wide’, *kīhi* ‘be flavorful’, *noho* ‘be fat’, etc.

In addition to the stative verbs, there are also a few isolated lexical exceptions that lack a B-form, such as *ništ'i* ‘run’ and *wqi* ‘rain’. Also, the verb *tɔ+i* ‘buy’ deviates from the rule as it is a verb that receives the B-form proper of verbs of class I in table 2, i.e., its B-form is *tɔ+ng-a=* and not **tɔ-ʔ=*, as would be expected by analogy to *ne+i* ‘dance’ in table 4.

4.2. Verbs with object markers. The morphological contrast concerning F-forms and B-forms presented so far involves complex stems that occur in isolation, that is, denuded of suffixal material. However, verbs can also show the contrast with object and dative suffixes, as shown in (9):

- (9a) *ta tišudi kha gà pɔd-i tema no*
 up.to tomorrow be.LOC 1.FUT know-F what 3.DEF.SG
dà ʔbeph-k-i ||
 3.FUT order-1OBJ-F

‘I won’t know until tomorrow what he’ll order me’.

- (9b) *bì ʔbeph-k-a=manʔagi=ʔna*
 3.PST order-1OBJ-B=once.again=QUOT

‘He ordered me (it) again’.

The examples in (9) show the suffix *-k*, an alternant of the morpheme for first-person object/dative. In (9a), the complex stem verb *ʔbephi* ‘order’ occurs at

a clausal and prosodic boundary and is inflected in F-form by means of the vowel *-i*. In (9*b*), the verb occurs intraclausally—it is followed by the adverb *manʔagi* ‘once again’—and it is inflected in B-form by means of the vowel *-a*. This serves as additional evidence that the markers *-i* and *-a* are the default morphological exponents of the forms.

The pattern illustrated in (9) applies to all verbs in the language that can receive object/dative morphology (except for stative verbs).¹⁴ Such verbs also include monomorphemic stem verbs which do not show the contrast in other situations. This is illustrated in (10):

(10*a*) *pa* *dà* *hɔ* ||
 PURP 3.FUT bring
 ‘(For the car) to bring (it)’.

(10*b*) *pa* *dà* *hɔ* *ma* *thɔ*
 PURP 3.FUT bring 1POSS corn
 ‘(For the car) to bring my corn’.

(10*c*) *pa* *dà* *hɔn-k-a=ma* *thɔ*
 PURP 3.FUT bring-1OBJ-B=1POSS corn
 ‘(For the car) to bring me my corn’.

In all examples in (10), we have different occurrences of the monomorphemic verb *hɔ* ‘bring’. This verb, like others of its type, does not have the inflectional contrast between an F-form and a B-form. This may be seen in (10*a*) and (10*b*), which illustrate how the same form of the verb is used in the same grammatical context where other complex stem verbs would show the contrast. However, the contrast is realized when the verb receives object morphology, as shown in (10*c*), where the verb receives the object/dative marker for first-person *-k* and is inflected in B-form by means of the vowel *-a*, in the context where such a form is expected.

On the other hand, object markers may appear in morphological doublets.¹⁵ The doublets for the first person are either *-kagi* or *-gagi*. When such doublets occur, the contrast between the forms is neutralized, as shown in (11):

¹⁴ When a stative verb denotes a property concept predicated of a first or second person, the latter is encoded as object. As a general rule, stative verbs that receive object morphology do not show the contrast other verbs do, i.e., (*ñ*)*hɛti* ‘be noisy’ : *ši ñhɛt-ki* ‘I’m noisy’ ~ *ši ñhɛt-ki manʔagi* ‘I’m noisy again’ but not **ši ñhɛt-k-a=manʔagi* ‘I’m noisy again’. However, this phenomenon is still poorly understood and requires further research.

¹⁵ In a morphological doublet, a given morph expressing person is repeated with a certain vocalic adjustment (i.e., *-g* > *-ga-gi*). This doublet carries out the same morphological function as the simple morph.

- (11) *no=r mole ja bì ?ra-kagi nu ma*
 DEF.SG=SG granny CP 3.PST give.to.1/2-1OBJ DEF 1POSS
bokho=?na
 money=QUOT

‘He says that the old lady gave me my money’.

(11) shows how the verb with a doublet object marker does not reflect the contrast. A hypothetical B-form like **?ra-kag-a* does not exist.¹⁶

5. A closer look at bound forms. A B-form of a verb is used when the verb does not precede a clausal or a prosodic boundary. Phonologically, a B-form cliticizes to the first free word of the syntactic phrase that follows the verb. This is illustrated in (12):

- (12) *nu?mi gá šok-a=[nu ma gɔʃthi]*
 when 2.PST open-B=DEF 1POSS door

‘When you opened my door’.

(12) shows that the B-form of the verb *šoki* ‘open (i.e., a door)’ cliticizes to the definite morpheme *nu*, which is the first free morphological unit of the NP *nu ma gɔʃthi* ‘my door’.

Most commonly, a B-form is also used as a host for the enclitic *=r /r/*, an allomorph of the singular morpheme {*ar*}. When this happens, B-forms marked with *-?* drop the glottal in contact with the flap because of phonotactic constraints:

- (13) *ja bì-n the=[r kwento-r pada ko=r šaha]*
 CP 3.PST-NPM end.B=SG story=SG vulture with=SG turtle

‘The story of the vulture and the turtle ended (here)’.

(13) shows that a B-form is also found when the verb receives the enclitic singular *=r*. This enclitic is always associated with an NP, and as we will see in the next section, NPs constitute the most natural context for a B-form to occur.

On the other hand, a B-form is sensitive to syntax given that it CANNOT occur when there is a clausal boundary, which is evidence that a B-form is not motivated by mere phonological sandhi. This feature is illustrated in (14):

- (14a) *pwes dà ?(j)>i-?=ha ja šo?thai*
 SO 3.FUT NPS.get.in-B=LOC PL cannon

‘And he would get in the cannons’.

¹⁶ A doublet for first person such as *-kagi* can be further analyzed as *-ka-gi*. It is interesting that the same vocalism that applies to F-forms and B-forms is also used here.

(14b) *hingi* ∅ *pɔd-i* / **pɔ-ʔ=* *ha* *dà* *khap-i*
 NEG 3.PRES know-F/know-B what 3.FUT do-F

‘He doesn’t know what to do’.

(14a) shows the natural occurrence of the B-form of the verb *kihi* ‘get in (deictic point inside)’¹⁷ before a PP headed by the preposition *ha*, a free alternant of the general locative preposition *ka*, seen in (6a) above. As expected, this B-form is cliticized to the preposition, since the latter is the first word of its phrase. (14b) provides a similar phonological context to the one in (14a). However, no B-form is possible here since the verb precedes a syntactic boundary before a subordinated clause. In (14b), only the F-form of the verb *pɔdi* ‘know’ is possible, even though the complementizer *ha* ‘what/if’ in (14b) is homophonous to the preposition in (14a).

5.1. Bound forms and prosody. A B-form of a verb is a dependent verbal form conditioned by prosody. As a general principle, a B-form is used when the verb is prosodically phrased together with a following constituent within a clause. However, Otomi shows certain idiosyncrasies regarding the interface between prosodic parsing and the type of syntactic phrase with which the verb is phrased together. Verbs in Otomi are ALWAYS phrased together with a following NP (providing it pertains to the same clause). This phenomenon is illustrated in (15). Following the convention used in Dresher (1994), natural prosodic phrases are indicated by parentheses:

(15) (*no=r* *ndə*) (*bi* *ʔɛʔm-a*=[*no* *ma*
 DEF.SG=SG man 3.PST drive.down.cattle-B=DEF.SG 1POSS
 mɛhti])
 cattle

‘The man drove my cattle down (the plain)’.

(15) shows that the inflected verb *bi ʔɛʔmi* ‘s/he drove cattle down’ is phrased together with the NP *no ma mɛhti* ‘my cattle’, and in such a context, it naturally occurs in B-form. In such cases, the head of the following NP carries the main stress of the intonational unit, while the verb in B-form receives a secondary stress.¹⁸

In current approaches to prosody, such as Selkirk (1984), Nespor and Vogel (1986), and Dresher (1994), prosodic structure is understood to be hierarchical. At the bottom of the prosodic hierarchy is the phonological word

¹⁷ The nonpresent stem *ʔjɪhi* (14a) of the verb *kihi* ‘get in (deictic point inside)’ is irregular; one would expect **<g>ihi* (see 2).

¹⁸ A similar phenomenon to the one treated here was first observed by Soustelle, who described it in passing for the Otomi of San José del Sitio, Mexico State (see 7 and Soustelle [1937] 1993).

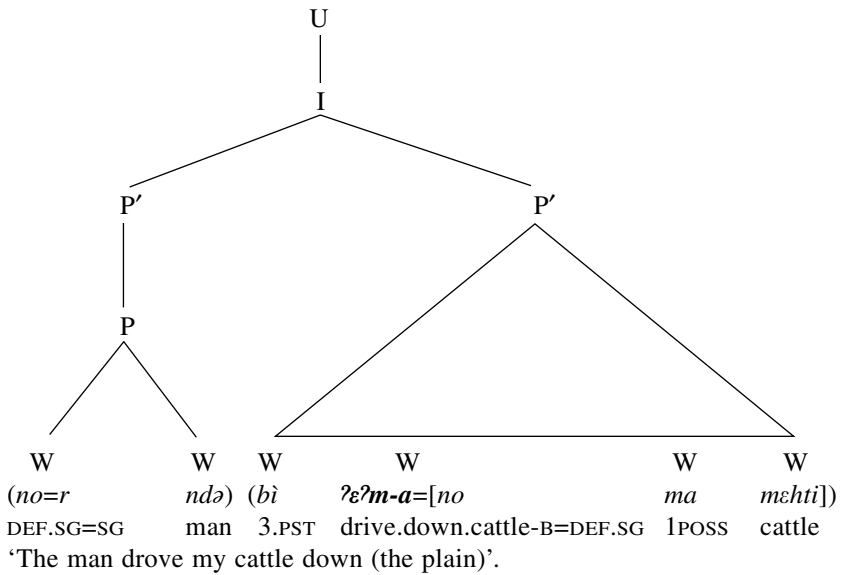


FIG. 1

(W), which is followed by the phonological phrase (P). The conjunction of phonological phrases within the same intonational unit creates the intonational phrase (I). At the top of the hierarchy, there is the utterance (U). Even though I do not try to give a formal account of the distribution of the forms here, the use of tree diagrams to represent relevant prosodic phrasings may be convenient to clarify the arguments. With this goal in mind, consider the tree diagram in figure 1.

All levels of the prosodic hierarchy have been represented for the clause in figure 1. What is interesting in the representation is that both the verb and the following NP are treated within the same phonological phrase (P'). Generally, a phonological phrase is defined here according to a phonetic parameter of whether a pause is naturally possible at the utterance of the structure.¹⁹ The utterance of the NP subject *nor ndə* 'the man' forms another phonological phrase. These two phrases together form the intonational phrase represented by the I-node, which in itself forms the utterance at a higher level.

The phenomenon illustrated in figure 1 occurs in all situations regardless of the type of verb used (i.e., transitive or intransitive) or the syntactic

¹⁹ Phrase-level stress also plays a role in defining the scope of a phonological phrase, but I leave this information unspecified here since it requires further research.

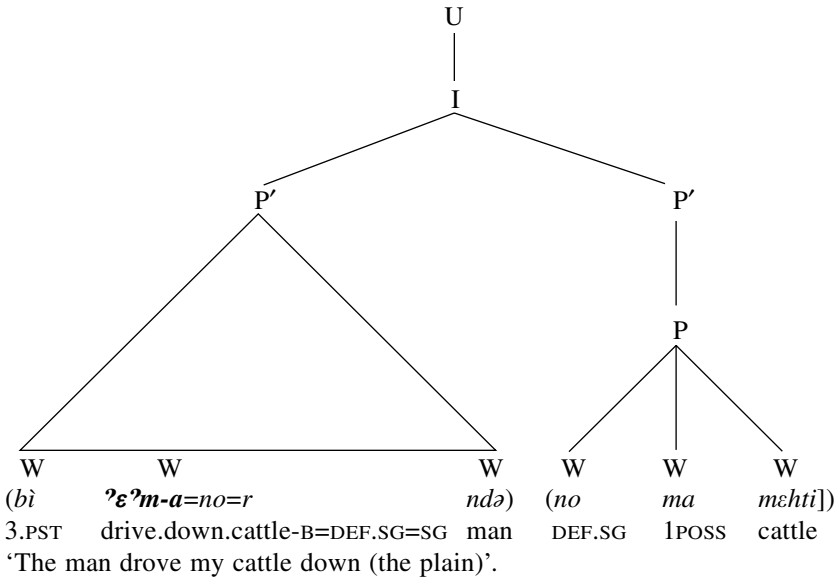


FIG. 2

function of the NP within the clause (i.e., subject, object, locative, etc.). As further evidence of this point, the example in figure 1, with the order SVO, is rendered in figure 2 with subject inversion (VSO). As in figure 1, the verb and the following NP are treated in figure 2 within the same phonological phrase. Here, the NP following the verb functions as subject, which proves that the prosodic parsing is not sensitive to syntactic function but only to the type of syntactic phrase that follows the verb. The utterance of the NP object *no ma mehti* 'my cattle' falls within a different phonological phrase.

Furthermore, no pause is possible between the verb and the NP when phrasing a clause like the one in figure 1. This is illustrated by the ungrammaticality of figure 3, which shows an ungrammatical prosodic alternative to figure 1. Figure 3 shows a verb that is phrased independently from a following NP. Since there is a pause after the verb, the verb must occur in F-form. However, such a prosodic parsing is utterly ungrammatical in natural discourse.

Figure 3 shows that no prosodic boundary is acceptable between a verb and a following NP. Nevertheless, a pause is possible within the NP itself, when speakers allow themselves time to think of a possible head for that phrase. This is illustrated in (16), taken from an oral text:

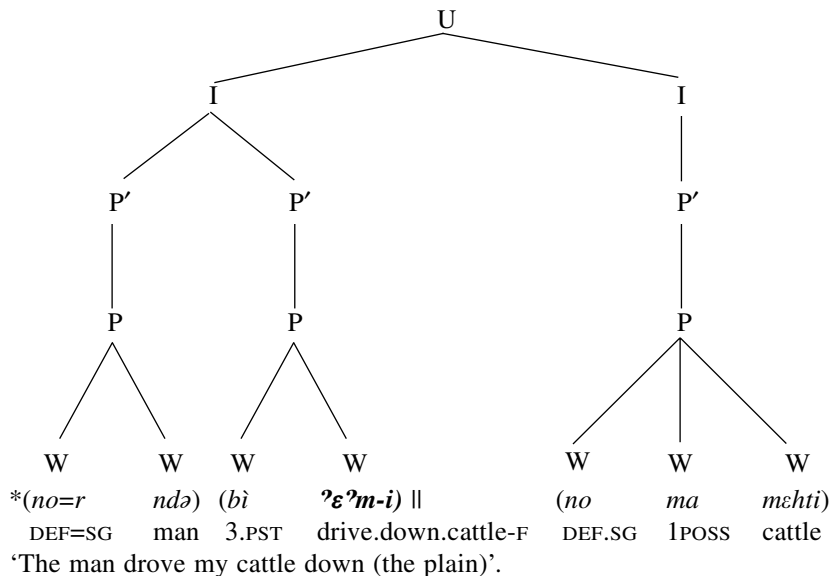


FIG. 3

(16) (dà <d>ɔng-a=[?na=r]) (/?na=r) (káro)
 3.FUT NPS.buy-B=IND=SG IND=SG car
 'He'd buy a . . . a . . . car'.

(16) shows a possible prosodic parsing of a verb and a following NP. The stem of the verb plus the first element of the following NP forms a phonological word on its own. They can be uttered in isolation and the verb in B-form receives the main stress of the sequence, i.e., /'dɔngə=/?nɑr/.

Examples like (16) strongly suggest that the complex phonological phrase formed by the verb and the following NP—as in figure 1—has the lower internal prosodic structure represented in figure 4. In figure 4, the phrase *bì ?ɛ?ma=no ma mɛhti* '(he) drove my cattle down' from figure 1 is divided into two phonological phrases; the first one comprises two words, the functional word expressing person and tense/aspect and the B-form of the verb together with the first unit of the following NP. The structure in figure 4 enables the spontaneous occurrence of the type of parsing shown in (16).

5.1.1. Exceptions to the pattern with noun phrases. The above examples illustrate that a B-form of the verb must be used if the verb is followed by an NP within the clause. The examples additionally show that in such cases the verb is prosodically phrased together with the NP. Nevertheless,

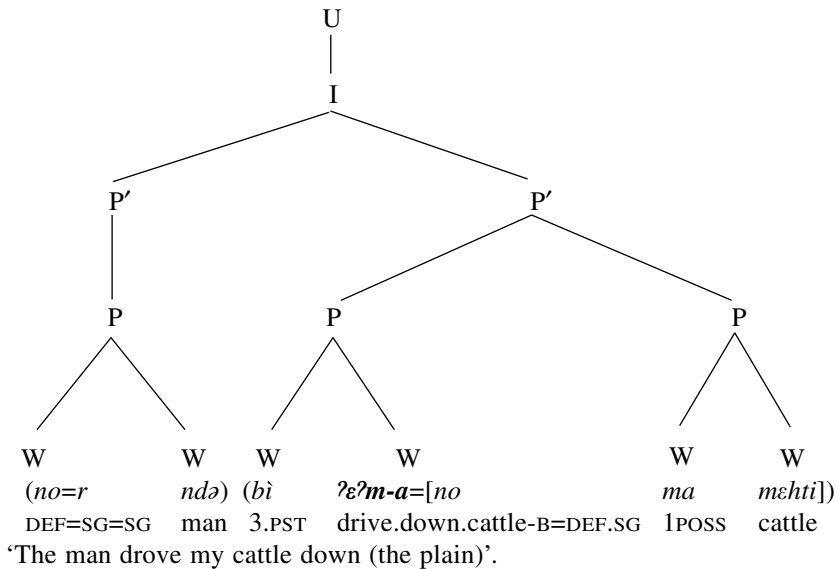


FIG. 4

there are possible discourse contexts which may alter these conditions. Consider for this purpose (17), the only example in my extensive data which apparently contravenes the generalization:

- (17) (*en otomí es*) ("pa dà zɨnt'-i || (ar bida")
in Otomi is PURP 3.FUT sound-F SG violin

'In Otomi it's: *pa dà zɨnt'i ar bida* (to play the violin).'

(17) is a metalinguistic example produced by my main consultant during a working session to correct a recorded textual sequence that she considered improper in Otomi. The prosodic parsing between the verb and the NP in (17) is very clear. According to the dominant pattern shown in figure 1, we expect such a phrasing to be ungrammatical. However, the example is an instance of metalinguistic use, and it is well known that such a use of language employs prosody in special ways for its own communicative purposes.

Apart from the exception in (17), the imperative appears to set a semantic restriction to the pattern. Verbs in the imperative may appear in F-form when the following NP has an animate referent. This is illustrated in (18):

- (18a) (*tsi-?=[ma ?behño]!*)
bring.animate-B=1POSS woman

'Bring my wife!'

- (18b) (*tsih-i*) ([*ma* *?behño*]!)
 bring.animate-F 1POSS woman
 'Bring my wife!'

In (18), the verb *tsihi* 'bring an animate' is used in the imperative. Both its B-form *tsi?e* (18a) and its F-form *tsihi* (18b) are possible here. However, such alternation is not possible when the clause is either a statement or a question, as shown in (19):

- (19) *(*ha* *gì* *tsih-i*) ([*ma* *?behño*]?)
 INTER 2.FUT bring.animate-F 1POSS woman
 'Are you going to bring my wife?'

The pattern illustrated in (19) is only possible with animate referents. Verbs in the imperative occur in B-form with a following inanimate NP. This restriction appears in (20):

- (20) (*k'í?š-a*/**k'í?ts'-i*) [*nu=jí* *ja* *?bɔirza*]!
 uproot-B=/uproot-F DEF=DEM.PL PL bush
 'Uproot those bushes!'

(20) shows how the use of the verb *k'í?ts'i* 'uproot' in the imperative still occurs in B-form before an NP expressing an inanimate referent. The use of the F-form of the verb in this context renders an ungrammatical result. This pattern contrasts with (18b), where the use of such a form is possible.

Finally, we have seen that some verbs lack a morphologically distinct B-form in contexts where other verbs have it. Among such verbs we have stative verbs, as in the examples in (8); monomorphemic verbs, as in (10a) and (10b); and loanwords. Such verbs are phrased in prosodic terms as if they occurred in F-form for all contexts. In other words, when a verb of this type is (for example) followed by an NP, both the verb and the NP form different phonological phrases because a pause may be placed freely after the verb, as shown in (21a). Most commonly, however, such a pause is not necessary, as indicated in (21b). This pattern contrasts fully with the ungrammaticality of figure 3.²⁰ Examples in (21) are repetitions of (8b):

- (21a) (*hingi* \emptyset *t'aši*) || (*ma* *ngu*)
 NEG 3.PRES be.white 1POSS house
 'My house isn't white'.
 (21b) (*hingi* \emptyset *t'aši*) (*ma* *ngu*)
 NEG 3.PRES be.white 1POSS house
 'My house isn't white'.

²⁰ The same could be said of (18) if the verb were to occur in F-form.

5.1.2. Bound forms and prepositional phrases. We have seen in the previous section that verbs in Otomi are phrased together with NPs and that they occur in B-form when this is the case. However, verbs do not have the same prosodic constraints when followed by other types of constituents. For example, verbs in Otomi are normally phrased together with most PPs and with adverbs that express location in space or time. Verbs followed by a PP headed by the generic preposition {*ka*} are naturally phrased together with the PP. This is illustrated in (22):

- (22) (gà *kĩʔt'-a=ka* ʔna=r *kubeta*) || (gà
 1.FUT put.mass.in-B=LOC IND=SG bucket 1.FUT

hqš-a=ka=r *molino*)
 carry-B=LOC=SG mill

'I would put it (the boiled corn) into a bucket and take it to the mill'.

Unlike what happens with NPs, as shown in the impossibility of figure 3 above, a pause may occur between the verb and the PP in (22). The use of such a pause will naturally trigger the use of an F-form. Such a prosodic parsing is shown in (23):

- (23) (gà *kĩʔt'-i*) || (*ka* ʔna=r *kubeta*) || (gà
 1.FUT put.mass.in-F LOC IND=SG bucket 1.FUT

hqts'-i) || (*ka=r* *molino*)
 carry-F LOC=SG mill

'I would put it into a bucket and take it to the mill'.

(23) shows a possible prosodic alternative to (22), where the verb is produced in isolation from the PP. Nevertheless, such a prosodic parsing is very infrequent in natural texts, even though it represents a possible structure. Moreover, the use of an F-form is unnatural if no prosodic pause is used between the verb and the PP. This is illustrated in (24):

- (24) (gà *kĩʔt'-i*) (*ka* ʔna=r *kubeta*) || (gà
 1.FUT put.mass.in-F LOC IND=SG bucket 1.FUT

hqts'-i) (*ka=r* *molino*)
 carry-F LOC=SG mill

Intended meaning: 'I would put it into a bucket and take it to the mill'.

(24) shows a rather unnatural prosodic rendering of an F-form followed by a PP headed by *ka*. I have found variation across speakers regarding the extent to which they accept sequences like (24) as appropriate. Some regard them as ungrammatical, others do not. However, in general, speakers do not

produce such sentences in a natural context. Otomi does not have a great number of prepositions, but the same applies for those PPs headed by prepositions that express a spatial relation, such as *maño* ‘on top of’ and *məte* ‘behind’, or others based on *ka*, such as (*š*)*ta ka* ‘up to’ and *de hɔi ka* ‘under’, etc.

Interestingly, a B-form is unnatural before a PP headed by the preposition *ko*, which may express instrument, (inanimate) comitative, and cause. Examples in (25) illustrate this point:

- (25a) (*komo ški <g>qt’i*) || ([*ko=r nu*] || (*ko*
 as 3.P-PERF NPS.paint-F with=SG DEF with
pintura)
 paint

‘As she had painted it (the fleece), with the . . . with paint’.

- (25b) (*ja bi <d>ehm-i*) ([*ko ?na=r do*])
 CP 3.PST NPS.break.of.volume-F with IND=SG stone

‘He broke it with a stone’.

Both examples in (25) illustrate the natural use of the F-form when the verb is followed by a PP headed by the preposition *ko* ‘with’. (25a) occurs frequently and illustrates how a prosodic pause can indeed be placed between the verb and the *ko*-PP. Similarly, (25b) shows how the verb *tehm* ‘break (i.e., of a volume)’ occurs before the *ko*-PP, with no prosodic pause between them.

Interestingly, the use of a B-form in a situation like (25b) is as unnatural as the use of an F-form in (24) above. This is shown in (26):

- (26)?? (*ja bi <d>ehm-a=[ko ?na=r do*])
 CP 3.PST NPS.break.of.volume-B=with IND=SG stone

‘He broke it with a stone’.

On the other hand, when the *ko*-PP is used to express cause, the use of a B-form is completely ungrammatical, as shown in the contrast between (27a) and (27b):

- (27a) (*bi ?<j>ot’-i*) ([*ko hjadi*])
 3.PST NPS.dry-F with sun

‘It dried in the sun’.

- (27b) *(*bi ?<j>ot’-a=[ko hjadi*])
 3.PST NPS.dry-B=with sun

‘It dried in the sun’.

TABLE 6

I		II		III	
++B-Form/F-Form		++F-Form/B-Form		F-Form/*B-Form	
<i>mbo</i>	'inside'	<i>nfrente</i>	'opposite'	<i>mambo</i>	'inside'
<i>nthi</i>	'outside'	<i>n'alado</i>	'beside'	<i>manthi</i>	'outside'
<i>nu'wa</i>	'here'	<i>mɔs-i-ngehni</i>	'in front'	<i>ši-ñho</i>	'well/good'
<i>nuni</i>	'over there'	<i>nzotho</i>	'beautiful(ly)'	<i>nts'edi</i>	'hard'
<i>ja'pi</i>	'far'				
<i>man'agi</i>	'once again'				
<i>mande</i>	'yesterday'				
<i>šudi</i>	'this morning'				
<i>nkhani</i>	'that way'				

5.1.3. Bound forms and adverbs. We have seen in the previous section that a semantic generalization may be made for the distribution of the forms with PPs. For adverbial phrases, such a generalization is difficult, and I prefer to view the phenomenon as lexically triggered. In general, adverbs differ from PPs in that adverbs that favor the use of a B-form will also accept an F-form, which suggests that F-forms are more acceptable with adverbs. As with PPs, pauses between the verb and the adverb are possible. In such cases, an F-form is used, as expected. Table 6 lists a number of adverbs, noting their preference for one of the forms.

The adverbs in table 6 are arranged in three columns according to the type of verbal form that is commonly used with them. Adverbs in column I normally occur with a B-form but an F-form is also possible. In contrast, adverbs in column II generally occur with an F-form, while a B-form is also acceptable. Finally, adverbs in column III occur only after a verb in F-form. A B-form is not possible with such adverbs. The adverbs *nfrente* 'opposite' and *n'alado* 'beside' in column II are loans from Spanish. Their preference for an F-form may be related to the fact that they are loanwords, like the preposition *ko* in (25) above. The adverb *mɔs-i-ngehni* 'in front' translates literally as 'more over there' and the preference for an F-form can be explained by the fact that the adverb *mɔs* 'more' involves a clausal boundary, as it is also commonly used with stative verbs denoting property concepts. The adverbs *mambo* 'inside' and *manthi* 'outside' in column III can only occur with an F-form. Morphologically speaking, these two adverbs include the free adverbs *mbo* 'inside' and *nthi* 'outside' in column I, plus the bound morph *ma-*, which appears to have functioned in historical times as a locative marker. However, their behavior regarding the forms is not motivated by the co-occurrence of *ma-*, given that the spatial preposition *maño* 'above' takes a B-form and it is based on *ma-* plus *ño*, a word meaning 'head'. Other adverbs denoting properties are placed in column III, such as *ši-ñho* 'well', *nts'edi* 'hard', etc., because they function as predicates; they are based on

stative verbs and thus set a clausal boundary before the verb. The lexeme *nzɔtho* ‘beautiful(ly)’ in column II is an apparent exception to this pattern.

6. Other uses of the forms. In this section, I explore other uses and idiosyncrasies of the forms in SI Otomi. In **6.1**, I describe how two completion verbs occur in B-form when they function as matrixes of a periphrasis. In **6.2**, I present an example of the use of an F-form before the quotative clitic. **6.3** introduces some examples of nominal compounds in the language which have stative verbs in B-form. Finally, in **6.4**, I briefly present two verbs that have a different meaning when they are used in either form.

6.1. Bound forms and the completion periphrasis. Completion verbs like *thege* (intr) ‘end’ and *tege* (tr) ‘extinguish something’ are often used in SI Otomi as the matrix verbs of a periphrastic construct that renders both the completion of an action and the total affectation of a thematic participant. I call such a construct a “completion periphrasis.” Under such circumstances, an F-form is expected, given that the periphrasis includes two verbs. However, the verbs may also occur in B-form. This phenomenon is shown in (28):

(28a) *nu=ja mbəhə=ʔi* ([*bì* *theg-e* [*bì* <*d*>*u*]])
 DEF=PL mestizo=3PL 3.PST end-F 3.PST NPS.die

‘Those mestizos people died out’.

(28b) ([*bì* *the-ʔ*= [*bì* <*d*>*u*]])
 3.PST end-B=3.PST NPS.die

‘Those mestizos died out’.

In (28a), we have the natural occurrence of the F-form of the intransitive verb *thege* ‘end’ before the inflected verb *bì du* ‘(s/he) died’. Prosodically, the matrix verb in F-form in this construction is always phrased together with the content verb. A B-form in this syntactic context is not expected, but it is possible, as shown in (28b). The use of a B-form in this context is common, and examples like these suggest that the inner syntactic boundaries within the periphrasis have become loose as the aspectual matrix is phrased together with the content verb. Nevertheless, the phenomenon is still poorly understood at the present time and deserves further investigation.

6.2. Free forms and the quotative enclitic. As with suffixed material in general, verbal stems in Otomi undergo readjustment when they host enclitics. This is illustrated in (29) with the verb *honi* ‘find’:

(29) *ge bì h<j>oñ=ʔa*
 REL 3.PST NPS.find=3SG(SUBJ)

‘He was the one who found it’.

In (29), the nonpresent stem of this verb is *hjon-*. In contact with the glottal of the clitic *=ʔa*, the nasal in the stem is palatalized. However, a verb is used in F-form when it hosts the quotative clitic *=ʔna*, which is commonly used when the predication was said by someone else and not by the speaker (i.e., quotative) or when the speaker does not commit to the truth of the predication. An example is given in (30):

- (30) (*ʔne* *bì* <*m*>*a=pja*) (*bà* ***hon-i***=*ʔna*)
 and 3.PST NPS.go=now 3.SIM look.for-F=QUOT

‘And it’s said that it (the tiger) went to look for it (the cricket)’.

The enclitic *=ʔna* in (30) is the grammaticalized outcome of the defective verb *ʔena* ‘say’. Since it is a clitic and has a glottal, we expect here the same phonologically adjusted stem as in (29) above, i.e., **hoñ=ʔna*. However, the verb must occur in F-form here, which suggests that the clitic preserves its verbal origin and places a syntactic boundary before the verb.

6.3. Bound forms and nominal compounds. We have already seen in 4.1 that stative verbs do not have a contrast in the grammatical context where it is expected. There is, however, a grammatical context where their B-form may be used. The vast majority of stative verbs in the language denote property concepts in the sense proposed by Dixon (1982). On the other hand, nominal compounding is a grammatical strategy in the language to attribute a property to the referent of a noun, and the language allows six stative verbs to occur as members of such nominal compounds: (*n*)*t’aši* ‘be white’, (*n*)*theni* ‘be red’, (*n*)*k’ast’i* ‘be yellow’, (*n*)*k’angi* ‘be blue/green’, (*n̄*)*ʔu* ‘be sweet’, and (*n*)*zehwa* ‘be silly (for an animal)’.²¹ The first four qualify for the morphological rules presented in 4 above, and they appear in their expected B-form when they occur in such nominal compounds, as illustrated in (31):

- (31) *dá* *tɔng-a=ma* ***k’ašt’-a-pahni***
 1.PST buy-B=1POSS be.yellow-B-sweater

‘I bought my yellow sweater’.

In (31), we have the compound *k’ašt’a-pahni* ‘yellow sweater’, where the verb expressing the color property (*n*)*k’ašt’i* ‘be yellow’ occurs in B-form. When such verbs are used in the compounds, they must occur in B-form, as shown by the ungrammaticality of (32), with an F-form:

- (32) ****k’ašt-i-pahni***
 be.yellow-F-sweater
 ‘yellow sweater’

²¹ The prefix in such verbs is a morphologically conditioned morpheme that is part of the inflection of all stative verbs.

In addition to stative verbs, the language has another large class of inchoative verbs. All members of this class can occur in such attributive nominal compounds, and when they do, they must occur in B-form. For example:

- (33) *ma nošk-a-ngə*
 1POSS harden (for meat)-B-meat
 ‘my hardened meat’

In (33), we have the B-form of the inchoative verb *noški* ‘harden (i.e., for meat)’. (31) and (32) above suggest that stative verbs take a B-form when they are part of the same NP as the following noun. The same applies to inchoative verbs, which like other verbs can also take a B-form when they precede NPs in a clause.

6.4. The forms and the lexicon. All verbs that inflect for the forms have the same meaning except (at least) two of them: *thogi* ‘pass by’ and *ʔopho* ‘write’. When *thogi* ‘pass by’ is used in B-forms, it evokes the motion of a moving figure who passes by a given location and has no contact with the observer, e.g., a person who walks by with no further interaction. On the other hand, when the verb is used in its F-form, the verb evokes contact and interaction, e.g., a person who comes by and stops to talk. Similarly, the intransitive verb *ʔopho* means ‘write’ in its F-form but ‘get married in a registry office’ in its B-form. The role played by the forms in such lexical processes is difficult to pinpoint at this stage and deserves further study.²²

7. The forms in other dialects. A number of previous studies have reported the existence of B-forms in other dialects of Otomi. Soustelle ([1937] 1993) is the earliest source attesting B-forms. The author mentions the phenomenon in passing for the dialect of San José del Sitio, Mexico State. He points out that a verb in this dialect at times forms a phonological word together with the following element, and speculates that such a phenomenon occurs in eloquent speech (Soustelle [1937] 1993:139 [my translation; Soustelle’s transcription has been adapted]): “The word that follows the verb fuses with the verb in the pronunciation. We have only one word from a phonetic point of view, as in *mizɔnkartsiwəni*, instead of *mizɔni kartsiwəni*.” Soustelle’s Otomi example is analyzed in (34):

- (34a) *mi-zɔn-kar-tsi-wəni*
 3.IMP-cry-DEF.SG-DIM-child
 ‘The small child was crying’.

²² The differences here are lexical; in other words, the verbs have a different meaning regardless of the syntactic context in which they occur.

(34b) *mi-zɔn-i* *kar-tsi-wɛni*
 3.IMP-cry-F DEF.SG-DIM-child

‘The small child was crying’.

Unfortunately, Soustelle did not investigate this phenomenon further, but the example he gives is reminiscent of a B-form in SI Otomi. In (34a) we have the form *zɔn-* for the verb *zɔni* ‘cry’ in a context where a B-form is expected. The B-form of such a verb would be *zɔnga=* in SI Otomi. B-forms are also attested in Bernal Pérez (1986), a valuable dictionary containing the inflectional paradigms of 505 verbs in Mezquital Otomi. B-forms in this dictionary are treated as “medial forms”; however, no further account is provided of their distribution. As expected, the texts from Mezquital Otomi in Hess (1968), Lanier (1974), and Salinas Pedraza (1984) reveal clear cases of B-forms. However, in such texts, B-forms are rather scattered and appear quite unsystematically, due mainly to the written character of the texts (see 7.1).²³

Apart from Mezquital Otomi, the existence of such forms has not been reported in Sierra Otomi by Voigtlander and Echegoyen ([1979] 1985), but their texts from this language do show a couple of instances which are reminiscent of Soustelle’s data given above. The contrast does not seem to exist in Ixtenco Otomi (Lastra 1997; 1998) given that this language is highly suffixal and B-forms are not expected in systems where verbal bases rarely occur stripped of all suffixal material (for an exception, see 4.2 above). Likewise, the contrast is not seen in the texts in Bartholomew (1973; 1984) or in Lastra (1989; 1992) for the Otomi spoken in the south of Mexico State; but Andrews (1993:1, n. 1) does report: “In this language, and all other Otomí dialects spoken in the State of Mexico, verb stems ending with certain syllables or *y* lose these endings in utterance-medial position. For instance *bi=thogi* ‘he passed by’ becomes *bi=thoh ni* ‘he passed by there’ when *ni* ‘there’ is added.” Andrews’s observations are very interesting and I take them to indicate that a similar phenomenon may be found in the Otomi dialects from the State of Mexico. However, I am not sure to what extent Andrews’s analysis is accurate, given that the adverb *ni* ‘there’ is probably an enclitic. If this is true, it is common for an Otomi stem to undergo phonological adjustment in contact with any type of suffixed material (see 29 above).

As for Santiago Mexquititlán Otomi, a neighbor of SI Otomi, Hekking and de Jesús (1984) do not report on the forms, but a closer look at the spoken texts in Hekking (1995) reveals that they are present in this dialect too, al-

²³ Ecker (1937) and Bernard (1974) account for a couple of B-forms in their samples as being verbs with a third-person object marker. However, intransitive verbs also inflect B-forms, making this analysis untenable. Similarly, the phenomenon was not noticed by Leon and Swadesh (1949).

though they may be at a different stage of development or use. Further research needs to be done on other dialects, in the light of the phenomenon described in this article, in order to understand to what extent this pattern is indeed a family feature, as I suspect it is.

7.1. Bound forms and current writing conventions. In the previous section, I pointed out that B-forms do not seem to occur in other dialects of Otomi, or they may not yet have been noted as a grammatical category. A number of such dialects have some degree of written tradition, which constitutes the bases for the writing standards of other dialects. This is certainly the case with the ethnographic document in Mezquital Otomi by Salinas Pedraza (1984) and religious translations by the Summer Institute of Linguistics. These texts have had some impact on both communities and various institutions in charge of native language programs. Salinas Pedraza (1984) and *Ra Maꞑat'øde de ra Marco* (1957) do not show B-forms, although they are written in a dialect where such forms exist.²⁴ B-forms are not used when writing Santiago Mexquititlán, which has traits of B-forms (see, e.g., Hekking and de Jesús 2002 or the texts in Villegas 2002). Thanks to the efforts of Ewald Hekking and his group, Santiago Mexquititlán Otomi has played an important role in the slowly emergent writing standard of Northwestern Otomi, the dialect closest to SI Otomi geographically. With a written tradition that does not include recording B-forms, speakers of dialects with B-forms have not really had a chance to express them in writing, since these forms are not yet properly categorized. For this reason, B-forms cannot be expected to occur in standard renderings of native texts, though they may occasionally appear as an isolated phenomenon (cf. Hekking and de Jesús 2002).

8. Conclusion. In this paper, I have presented data on the distribution and use of two inflectional forms of verbs in San Ildefonso Tultepec Otomi: “free forms” and “bound forms.” Broadly speaking, a free form (F-form) is used when the verb precedes a clausal boundary or a prosodic pause. Because an F-form can be uttered in isolation, it is also employed as the citation form of the verb. In contrast, a bound form (B-form) is used when the verb occurs intraclausally. A B-form cannot be uttered in isolation; it acts as a bound morph and is cliticized to the first word of the following phrase. It has

²⁴ Albert Bickford has raised the interesting point of why the written tradition for Mezquital Otomi does not represent B-forms. As he suggests, perhaps the reason is a conscious or unconscious decision on the part of the writers to write everything in the form that would appear in isolation, influenced perhaps by the tendency of newly literate readers and writers to pronounce each word in a text separately. Once the written tradition for Mezquital was established, it is understandable that other dialects would follow the same conventions and not represent B-forms either. I am grateful to him for this valuable observation.

- _____. 1984. Algunas diferencias entre la narración oral y la narración escrita, con referencia a una muestra en el otomí del Estado de México. *Summer Institute of Linguistics–Mexico Workpapers* 5:52–61.
- BERNAL PÉREZ, FELIPINO. 1986. *Hñāñhu: Recopilación de 505 verbos otomíes conjugados*. Cardonal, Hidalgo, Mexico: Centro Social de Cardonal.
- BERNARD, H. RUSSELL. 1974. Otomi tones in discourse. *IJAL* 40:141–50.
- DIXON, R. M. W. 1982. Where have all the adjectives gone? Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax, ed. R. M. W. Dixon, pp. 1–62. Berlin: Mouton.
- DRESHER, BEZAEEL ELAN. 1994. The prosodic basis of the Tiberian Hebrew system of accents. *Language* 70:1–52.
- ECKER, LAWRENCE. 1937. Los dos “Meteoros”: Yooho ya ny’oi. *Investigaciones Lingüísticas* 4, nos. 3/4:254–61.
- HEKKING, EWALD. 1995. El otomí de Santiago Mexquititlán: Desplazamiento lingüístico, préstamos y cambios gramaticales. Amsterdam: Institute for Functional Research in Language and Language Use.
- _____. 2002. Desplazamiento, pérdida y perspectivas para la revitalización del hñāñho. *Estudios de Cultura Otopame* 3:221–48.
- HEKKING, EWALD, AND SEVERIANO ANDRÉS DE JESÚS. 1984. *Gramática otomí*. Querétaro, México: Universidad Autónoma de Querétaro.
- _____, eds. 2002. *Ya ‘bede ar hñāñho Nsantumuriya: Cuentos en el otomí de Amealco*. Querétaro, México: Universidad Autónoma de Querétaro.
- HESS, H. HARWOOD. 1968. *The Syntactic Structure of Mezquital Otomi*. *Janua Linguarum, Series Practica*, no. 43. The Hague: Mouton.
- HOLLENBACH, BARBARA E. 1984. The phonology and morphology of tone and laryngeals in Copala Trique. Ph.D. dissertation, University of Arizona.
- KATAMBA, FRANCIS. 1992. *An Introduction to Phonology*. London: Longman.
- KAUFMAN, TERRENCE. 1988. Otomanguean tense/aspect/mood, voice, and nominalization markers. Paper presented at the Second Spring Workshop on Theory and Method in Linguistic Reconstruction, University of Pittsburgh.
- LANIER, NANCY, ed. 1974. *Ra ‘bede de hñāto ya zu’we: Cuentos de animales en otomí del Valle de Mezquital*. Mexico, D.F.: Instituto Lingüístico de Verano.
- LASTRA, YOLANDA. 1989. *Otomí de San Andrés Cuexcontitlán, Estado de México*. *Archivo de Lenguas Indígenas*, no. 13. México, D.F.: El Colegio de México.
- _____. 1992. *El otomí de Toluca*. México, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- _____. 1997. *El otomí de Ixtenco*. México, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- _____. 1998. *Ixtenco Otomí*. Munich: Lincom Europa.
- LEON, FRANCES, AND MORRIS SWADESH. 1949. Two views of Otomi prosody. *IJAL* 15:100–105.
- NESPOR, MARINA, AND IRENE VOGEL. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- PALANCAR, ENRIQUE L. 2004. Property concepts in Otomi. Paper presented at the 2003–4 Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas, Boston.
- _____. Forthcoming. *Transitividad en el otomí de San Ildefonso: una incursión tipológica en lenguas otomangués*. *Memorias del VII Encuentro Internacional de Lingüística del Noroeste*. Sonora: Universidad de Sonora.
- PYE, CLIFTON. 2001. La adquisición de la morfología verbal en español y maya k’iche’: Retos a la teoría de la “gramática universal.” La adquisición de la lengua materna: español, lenguas mayas, euskera, ed. De Lourdes de León Pasquel and Cecilia Rojas Nieto, pp. 51–74.

- México, D.F.: Universidad Autónoma de México and Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Ra Maçat'ode de ra Marco: El santo evangelio según San Marcos, en el idioma otomí de Mezquitil. 1957. México, D.F.: [no publisher].
- SALINAS PEDRAZA, JESÚS. 1984. *Etnografía del Otomí*. México, D.F.: Instituto Nacional Indigenista.
- SELKIRK, ELISABETH. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: The M.I.T. Press.
- SOUSTELLE, JACQUES. 1993. *La Familia Otomí Pame del México Central*. México, D.F.: Fondo para la Cultura y las Artes. [Spanish translation by Nilda Mercado of Jacques Soustelle, *La Famille Otomi-Pame du Mexique Central*. Paris: Institut D'Ethnologie, 1937.]
- SUÁREZ, JORGE A. 1983. *The Mesoamerican Indian Languages*. Cambridge: Cambridge University Press.
- VILLEGAS, MARÍA ELENA. 2002. *Las cláusulas de relativo en el español otomí de Santiago Mexquititlán*. M.A. thesis, Universidad Autónoma de Querétaro.
- VOIGTLANDER, KATHERINE. 1998. *La semántica de los sufijos formativos del tema verbal en yuhu (otomí de la Sierra)*. Paper presented at the II Coloquio Estatal sobre Otopames, Hidalgo, México.
- VOIGTLANDER, KATHERINE, AND DORIS BARTHOLOMEW. 1972. *Semiology and transitivity in Eastern Otomí verbs*. *Lingua* 29:38–53.
- VOIGTLANDER, KATHERINE, AND ARTEMISA ECHEGOYEN. 1985. *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. México, D.F.: Instituto Lingüístico de Verano. [First impression in 1979 without author names.]
- ZAVALA, ROBERTO. 1992. *El Kanjobal de San Miguel Acatán*. México, D.F.: Universidad Nacional Autónoma de México.

2. Travaux sur le lexique otomi

- TRAVAIL 1. Peralta, Valentin et Enrique L. Palancar. 2015. Actividades rutinarias en náhuatl y otomí. *Amerindia* 37:2, 259 – 284. 209
- TRAVAIL 2. Palancar, Enrique L. 2008. The emergence of active/stative alignment in Otomi. Mark Donohue et Søren Wichmann (eds.), *The Typology of Semantic Alignment*, pp. 357 – 79. Oxford: Oxford University Press. 235
- TRAVAIL 3. Palancar, Enrique L. 2007. Cutting and breaking in Otomi. *Cognitive Linguistics* 18:2, 307 – 17. 247
- TRAVAIL 4. Palancar, Enrique L. 2006. Property concepts in Otomi: A language with no adjectives. *International Journal of American Linguistics* 72:3, 325 – 66. 259
- TRAVAIL 5. Palancar, Enrique L. 2006. Middle voice and intransitivity in Otomi. *Linguistics* 43:1: 613 – 43. 301
- TRAVAIL 6. Palancar, Enrique L. 2004. Middle voice in Otomi. *International Journal of American Linguistics* 70:1, 52 – 85. 333

Actividades rutinarias en náhuatl y otomí

Valentín PERALTA-RAMÍREZ & Enrique L. PALANCAR

Escuela Nacional de Antropología e Historia & SeDyL/CELIA, CNRS
miahuatl@hotmail.com & palancar@vjf.cnrs.fr

1. Introducción a las actividades rutinarias

En este artículo estudiamos el modo cómo se codifica léxicamente el campo semántico de las actividades rutinarias en dos sistemas lingüísticos del área mesoamericana, como son el náhuatl (azteca, yutoazteca) y la familia otomí (otopame, otomangue). En el trabajo partimos de la hipótesis de que el campo semántico de las actividades rutinarias representa un área importante para la organización del léxico en las lenguas mesoamericanas que implica construcciones gramaticales específicas. El presente trabajo supone una primera aproximación al fenómeno y es por ello de carácter exploratorio y no tipológico. Nos concentramos solamente en náhuatl y en otomí por ser dos sistemas lingüísticos mesoamericanos que presentan características tipológicas y genéticas muy distintas, pero que se han venido hablando desde hace siglos en zonas geográficas colindantes, cuyas comunidades tienen afinidades culturales típicamente mesoamericanas que han permitido, sobre todo en tiempos históricos, un intenso contacto lingüístico. En el artículo mostramos que si bien los mecanismos empleados en náhuatl y otomí para el trato léxico-gramatical de las actividades rutinarias son en efecto muy diferentes, en ambos casos encontramos construcciones dedicadas a la codificación léxica de este área semántica determinada.

El mejor ejemplo de una “actividad rutinaria” es un tipo de actividad que el colectivo de la comunidad lingüística considera como de uso y costumbre.¹ En este sentido, las actividades rutinarias suelen designar acciones cotidianas que se llevan a cabo diariamente (p.ej. “lavar la ropa”, “echar tortillas”, “asear la casa”, “limpiar la milpa”, “trabajar en general”, etc.) pero también incluyen otras que están circunscritas a eventos culturales de índole más específica como son las celebraciones laicas o religiosas (p.ej. “rezar”, “hacer la matanza para la fiesta”, “emborracharse”, “bailar en un baile de la fiesta del pueblo”).

Dentro del rubro de las actividades rutinarias, incluimos también otras actividades que abarcan las etapas de la vida de un individuo, bien en su calidad de ser humano o como miembro de la sociedad (p.ej. “gatear”, “casarse”, “tener hijos”, etc.), así como situaciones de vida que conllevan estados emocionales de cierta duración que se espera que un individuo pueda experimentar en algún momento de su vida (p.ej. “estar de luto”, “envejecer y perder movilidad”, etc.), así como otras que se realizan atendiendo a convenciones sociales, tales como “peinarse”, “vestirse”, etc. Más lejos del centro de la categoría de actividad rutinaria se encuentran aquellas acciones que se realizan por necesidades corporales (p.ej. “orinar”, “estornudar”, etc.), así como otras que atienden a estados emocionales que caracterizan a un individuo del resto porque ese individuo las realiza de manera más usual que el resto (p.ej. “ser reflexivo”, “llorar mucho”, etc.).

En el caso prototípico, las actividades rutinarias refieren obviamente a acciones o experiencias protagonizadas por los seres humanos, pero se pueden extender al incluir comportamientos rutinarios o característicos de los animales también, sobre todo de los domésticos (p.ej. “inquietarse un caballo, asno o vaca”, “ladrar un perro”, “cacarear de las gallinas al poner huevos”, “lamerse un gato”, etc.), e incluso de ontologías inanimadas (p.ej. “el fermentar del pulque”, “el madurar de la fruta”, etc.).

Si bien tanto en náhuatl como en otomí las actividades rutinarias se codifican en verbos, estos mismos verbos suelen presentar características particulares bien porque reciben una flexión especial o porque resultan de mecanismos gramaticales de producción léxica exclusivamente dedicados a la codificación de actividades rutinarias. Por ejemplo, en las variantes del náhuatl las actividades rutinarias se codifican mediante dos construcciones de composición verbal. Se trata por un lado de la construcción de incorporación nominal del tipo en (1), y por el otro, de la construcción de predicación compleja

¹ Consideramos que la categoría de “actividad rutinaria” tiene estructura de prototipo, es decir, existen ejemplos típicos de la categoría y otros que son instancias más marginales.

desiderativa en (2). En el artículo mostramos que ambas construcciones contrastan con sus contrapartes analíticas que carecen de interpretación rutinaria. La semántica del verbo en (1) designa una actividad realizada de forma cotidiana como parte de las tareas diarias del hablante. El verbo (2) refiere a la etapa de los primeros pasos de un bebé.²

(1) Náhuatl de Amanalco

ni-taxkal-chi:wa
1SUJ-tortilla-hacer
'Hago tortillas.'

(2) Nawat de Pajapan

Ø-nehnemi-s-negi ho:n chogo-tzin
3SUJ-caminar-IRR-querer ese niño-DIM
'El niño ya quiere caminar.'

En contraste, en todas las variantes del otomí, existe un grupo de verbos intransitivos que recibe una flexión particular y que forman por lo tanto una clase flexiva independiente. La inmensa mayoría de los verbos de esta clase designan actividades protagonizadas por humanos en general, y de entre ellos muchos designan específicamente típicas actividades rutinarias, v.g. verbos del otomí de Tilapa como (INTR) 'barrer', ^h*peni* (INTR) 'lavar la ropa', ^h*tsu'ti* (INTR) 'fumar', *wédi* (INTR) 'coser', *húxi* (INTR) 'tocar un instrumento de viento', son típicos miembros de la segunda conjugación.

Por otro lado, existe asimismo en otomí una construcción de complementación especial que consta de un verbo matriz y una nominalización verbal. La construcción en cuestión se ejemplifica en (3), y en ella se requiere que el verbo nominalizado designe una actividad rutinaria. En (3), se tiene la nominalización *su* 'leñada' del verbo *xu* (INTR) 'leñar, cortar leña', que a su vez pertenece a la clase flexiva especial a la que nos referíamos antes.

(3) Otomí Oriental: Variante de Huehuetla

nge da=mba [ra su]
PROP 3.PRS.IRR=TN/ir SG leñada
'Para que fuera a leñar.' (Echegoyen & Voigtlander 2007: 379)

² Abreviaturas: 1 Primera persona, 2 Segunda persona, 3 Tercera persona, ABS: Absolutivo, ADV: Adverbial, AG: Agentivo, ANTER: Anterior, ANTIP: Antipasivo, CONT: Continuo, D: Tema dependiente, DEP: Dependiente, DET: Determinante, DIM: Diminutive, EXCL: Exclusivo, HAB: Habitual, IMPF: Imperfecto, INDF: Indefinido, INF: Infinitive, INHER: Inherente, INTR: Intransitive, IRR: Irrealis, NOMLZ: Nominalización, OBJ: Objeto, OP: Objeto primario, P: Preposición, PERF: Perfecto, PL: Plural, POS: Posesivo, PRS: Presente, PSD: Pasado, PSDO: Poseído, R: Realis, RED: Reduplicación, RFLX: Reflexivo, SG: Singular, SUJ: Sujeto, TA: Tema ajustado, TN: Tema nasal, VRBZR: Verbalizador.

Por otro lado, una actividad rutinaria lo suele ser por su semántica inherente, no porque aparezca en un marco gramatical determinado. Por ejemplo, es esperable que los lexemas que refieren actividades rutinarias se puedan y suelen flexionar en aspectos imperfectivos como el habitual o que aparezcan en construcciones que portan semántica habitual (Comrie 1976), como en (4a), pero no exclusivamente, como por ejemplo en (4b), en donde el mismo verbo intransitivo ^h*tsu*'*ti* que designa la actividad rutinaria de 'fumar', aparece en aspecto completivo, para indicar que la actividad se llevó a cabo.

(4) Otomí de Tilapa

a. *txú=mú=^htsu'ti*

1.HAB.R=IMPF.HAB=fumar

'Yo solía fumar.'

b. *txá=má=nde*

1.CONT.R=IMPF.CONT=querer[3OBJ]

'Yo quería fumar (y lo hice).'

tú=dú=^htsu'ti

1.PST.R=1=fumar

(Lit. 'Yo lo quería fumé'.)

Resulta por lo tanto pertinente distinguir entre “actividad rutinaria” y “aspecto habitual”, pues aunque van de la mano, son ontologías distintas. Las actividades rutinarias designan acciones que se llevan a cabo usualmente, es decir, de forma habitual, pero que están sancionadas por la comunidad como de uso y costumbre, al menos para el caso más típico. El carácter habitual de cada acción en cuestión depende a su vez del marco temporal que el saber cultural de la comunidad lingüística sanciona para cada actividad, de tal manera que hay actividades cuyo desempeño se juzga como cotidiano, otras como semanales, otras como anuales, etc. Por ejemplo, si alguien tiene el hábito de realizar la actividad rutinaria de “lavar los trastes”, uno infiere que lo hace todos los días, pero si el hábito es el de realizar la actividad rutinaria de “bailar”, uno puede asumir que esto tiene que suceder cuando hay fiesta, y las fiestas en los casos más típicos, no ocurren todos los días.

Por otro lado, si bien la función primordial de la flexión de aspecto habitual es presentar la acción de un verbo dado en una actividad habitual, nosotros distinguimos entre lexemas verbales con semántica inherente de actividad rutinaria (p.ej. “almorzar”, “segar”, “hacer tortillas”, “pescar”, etc.) de otros verbos con semántica más general que se emplean para eventos más puntuales (“comer una cosa (un taco, etc.)”, “cortar algo (una rama, etc.)”, “romper (un plato, etc.)”, etc.), cuyas acciones se pueden interpretar como rutinarias solamente si los verbos que las expresan se flexionan en aspecto habitual.

Por ejemplo, el presente simple en inglés es una construcción aspectual habitual, y como tal se emplea de la manera más natural y frecuente con verbos de actividad rutinaria, por ejemplo *she works* ‘ella trabaja’ o *she writes* ‘ella

escribe'. La misma construcción se puede emplear con otro tipo de verbos que no son de actividad rutinaria, por ejemplo *she breaks* 'ella rompe' o *she jumps* 'ella salta'. Este marco gramatical rinde la semántica de verbos como *break* 'romper' y *jump* 'saltar' como acciones habituales, y predispone a una interpretación de estas acciones como si fueran actividades sancionadas como rutinarias, cuando en realidad no lo son, al menos en el mismo grado que *work* 'trabajar' y *write* 'escribir', ya que para que estos enunciados sean gramaticales se necesita de información de contexto muy específica para entender bajo qué circunstancias y con qué propósito alguien puede romper o saltar de forma habitual, mientras que esta información cultural ya está presente en el marco conceptual para la interpretación de trabajar o de escribir.

Para el propósito de este artículo, estudiamos tanto construcciones que sirven para producir lexemas verbales de acción rutinaria (p.ej. el caso de la incorporación nominal para el verbo náhuatl intransitivo *taxkal-chi:wa* 'hacer tortillas' en (1)), como el trato gramatical que estos mismos lexemas pueden recibir (p.ej. la pertenencia a una clase flexiva especial, como el verbo otomí *^htsu'ti* (INTR) 'fumar' en (4)). El trato léxico y gramatical de este dominio conceptual nos parece interesante para ampliar nuestro conocimiento sobre el uso y función de ciertas estructuras lingüísticas que encontramos tanto en náhuatl como en otomí. Queda abierta la pregunta de similares comportamientos que se observan en otras familias de lenguas del área mesoamericana.

El artículo tiene la siguiente estructura. En la sección siguiente, presentamos las características de los mecanismos de generación de lexemas con semántica de actividad rutinaria en el náhuatl. Después, en la sección 3, estudiamos el modo en cómo verbos equivalentes se tratan gramaticalmente en otomí. El artículo concluye en la sección 4.

2. El trato de las actividades rutinarias en náhuatl

En esta sección mostramos que el campo semántico de las actividades rutinarias recibe una codificación especial en náhuatl y describimos de qué construcciones se trata. En las diferentes variantes del náhuatl, existen diferentes construcciones de composición verbal que se emplean casi exclusivamente para la expresión de actividades rutinarias. En este sentido, el náhuatl presenta un rico sistema de verbos complejos conformados por una raíz nominal más una verbal (N+V) o por dos raíces verbales (V+V). La primera estructura se conoce como "incorporación nominal" (IN) y tal como mostramos presenta varios subtipos.

Si bien una de las estructuras de IN del náhuatl clásico ha sido objeto de discusión desde los primeros estudios descriptivos de las lenguas indígenas

americanas (Kroeber 1909; Sapir 1911),³ en las dos últimas décadas, la discusión se ha centrado más bien en cómo se debe caracterizar, ya sea como un proceso esencialmente sintáctico (Sadock 1980; Baker 1988, etc.) o como léxico (Rosen 1989). Por otra parte, Mithun (1984) representa un trabajo seminal sobre la tipología de IN. Mithun divide este fenómeno en cuatro tipos: el Tipo I conocido como “IN por composición”, el Tipo II como “manipulación de caso”, el Tipo III que responde a “manipulación discursiva”, y el Tipo IV, conocido como “IN clasificatoria”. Peralta-Ramírez (2014) ha mostrado que el náhuatl de Amanalco tiene en efecto los cuatro tipos de IN de esta tipología. En este trabajo mostramos que los dos primeros, el de IN por composición y el de IN por manipulación de caso, se emplean como mecanismos gramaticales para la generación de léxico verbal cuya semántica denota actividades rutinarias exclusivamente.

En muchas ocasiones, las construcciones de IN contrastan con equivalentes estructurales analíticos, que no confieren semántica de actividad rutinaria, sino que refieren a eventos puntuales. Para este propósito compárese los ejemplos en (5). En (5a), se puede referir a un evento en el que el agente ha recogido granos de maíz que se encuentran tirados por el piso. En contraste, el ejemplo con IN en (5b), se emplea para designar la actividad en la que una persona realiza una actividad cultural específica, que se trata de elegir cuidadosamente los granos de buena calidad o de buen tamaño de otros que no lo son con vistas a la cosecha futura. De la misma forma, el verbo con IN en (6b) indica un tipo de actividad que un animal realiza con frecuencia, es decir, perfila al animal como el tipo de animales domésticos que tienen la maña de soltarse siempre de las cuerdas que los atan. La construcción analítica en (6a) designa un evento puntual protagonizado por un animal que se suelta por no estar bien amarrado. A no ser que se especifique, los ejemplos del náhuatl en este artículo proceden de la variante central hablada en Amanalco, Estado de México.

- (5)a. *Ø-ki-peʔpen in tlaoli*
 3SUJ-3OP-escoger.PERF DET maíz.ABS
 ‘Juntó el maíz.’
- b. *Ø-tlaol-peʔpen*
 3SUJ-maíz-escoger.PERF
 ‘Escogió maíz.’
- (6)a. *Ø-mo-kopi:na i-kech*
 3SUJ-RFLX-zafar 2POS-cuello
 ‘Se zafa de su cuello.’

³ Sobre el impacto del náhuatl clásico en la discusión sobre incorporación nominal véase Andrews (1975).

- b. *Ø-mo-kech-kopi:na*
 3SUJ-RFLX-cuello-zafar
 ‘Se zafa de su cuello.’

Tal como hemos apuntado, existen dos subtipos de IN empleados en el trato de las actividades rutinarias que presentamos en más detalle en las secciones siguientes. Se trata de (i) IN por composición, como los ejemplos en (5) y (6) y (ii) la IN por manipulación de caso. Ambas construcciones implican un reacomodo de los argumentos.

2.1. Incorporación nominal por composición

Como ya se había mencionado anteriormente, la IN por composición corresponde al Tipo I de la tipología de Mithun (1984). En este tipo, un verbo transitivo dado (V) forma un compuesto con su objeto nominal (N) dando por resultado un verbo intransitivo. La misma Mithun apunta que este tipo de IN expresa una actividad rutinaria. En el proceso de composición, se deriva una unidad léxico-conceptual compleja (N+V) en la que el N no refiere a ninguna entidad específica, es decir, carece de referencialidad debido a que pierde prominencia pragmática, semántica y sintáctica dentro del enunciado. Por ejemplo, en el náhuatl de la Huasteca, Merlan (1976: 187) muestra en los ejemplos en (7), que la construcción de IN en (7b) se emplea para designar la acción común, pragmáticamente no marcada, de cerrar una puerta. Es decir, para designar cualquier evento de la actividad rutinaria de “cerrar puertas” se emplea la IN. En esta construcción, el nominal aparece sin marcas de absoluto, las cuales se requieren por otro lado cuando el sustantivo no está incorporado, como en la construcción analítica en (7a). Esta construcción se emplea solamente cuando se quiere presentar el evento revestido de información pragmáticamente saliente, es decir, cuando por ejemplo el cerrar de la puerta se hizo de forma poco usual. Nótese, sin embargo, que en español, las dos lecturas tienen la misma codificación.

- (7) Náhuatl de la Huasteca
- a. *ne' Ø-ki-ca'-ki* *kallak-tli*
 él 3SUJ-3OP-cerrar-PERF puerta-ABS
 ‘Cerró la puerta.’
- b. *ne' Ø-kal-ca'-ki*
 él 3SUJ-puerta-cerrar-PERF
 ‘Cerró la puerta.’

La función de la IN por composición como mecanismo de codificación de acciones rutinarias aparece ya en Sapir (1911). Más recientemente, Mithun establece que en la base de la IN está la motivación de crear nuevo léxico.⁴

La composición se da por una razón. Una entidad, cualidad, o actividad se percibe lo suficientemente frecuente como para que se considere digna de recibir una denominación por sí misma. Por ejemplo, en inglés *bus money* o *lunch money* son compuestos nominales más probables que *sock money* o *screwdriver money*. [...] Lo mismo ocurre con los compuestos verbales, que son acuñados como nombres de actividades relevantes. Si alguien me pregunta dónde está mi hermano, yo le puedo responder, *he is out berry-picking* o *he is out mountain-climbing*, pero es probable que no responda *he is out ladder-climbing*, incluso si estuviera de hecho encaramado en una escalera [...]. *Ladder-climbing* ('escalada de escaleras') no es una actividad institucionalizada. Si yo en efecto lo dijera, la persona a quien se lo digo puede sospechar que *ladder-climbing* se refiere a una actividad reconocida en algún contexto como por ejemplo un nuevo deporte, o como una prueba para ingresar al cuerpo de bomberos. Los compuestos tienen un estatus léxico del que carecen sus contrapartes analíticas. (Mithun 1984: 848)

Debido a la existencia misma del compuesto que resulta de la IN está motivada por la necesidad de nombrar actividades rutinarias en la comunidad (sobre todo en los casos en los que una nueva actividad emerge y se establece), el resultado semántico puede perder propiedades de composicionalidad. Esto se puede observar en el ejemplo (8), en donde la estructura emergente refiere a la actividad de ser poeta o producir poesía, más que a la actividad concreta de dedicarse a la jardinería. Para esta última actividad, se podría emplear la construcción analítica, debido a que la construcción de composición ya tiene otra significación convencional.

- (8)a. *ni-xochi-teki*
 1SUJ-flor-cortar
 'Compongo poesía' (Lit. 'Corto flores.')
- b. *ni-k-teki* *xochi-tl*
 1SUJ-3OP-cortar flor-ABS
 'Corto flores.'

⁴ *Compounding is done for a reason. Some entity, quality, or activity is recognized sufficiently often to be considered name-worthy in its own right; thus Eng. bus money or lunch money are more likely nominal compounds than sock money or screwdriver money. [...] The same is true of verbal compounds, which are coined as names of recognizable activities. If you ask where my brother is, I might reply, he is out berry-picking or he is off mountain-climbing, but probably not he is out ladder-climbing, even if he is in fact climbing a ladder [...]. Ladder-climbing is not an institutionalized activity. If I did say it, you might suspect that ladder-climbing must refer to an activity recognized in some context-perhaps a new sport, or a test for joining the fire department. Compounds have a lexical status which their syntactic counterparts lack.*

En este sentido, muchos de los compuestos, al referir actividades rutinarias, reciben extensiones predicativas agentivas que se emplean para designar oficios. Compárense al efecto los ejemplos en (9). En (9a) se expresa la relación de posesión entre una mujer y su hijo. Nótese que el sustantivo *pil* ‘hijo’, que es un término de parentesco, aparece en una FN poseída. El predicado de IN en (7b) es intransitivo, y refiere a la actividad de cuidar niños. Es decir, (7b) no significa que uno tiene hijos, sino que los cuida, independientemente de que uno los haya de tener o no. Por último, en el ejemplo (9c), el verbo aparece con el sufijo agentivo *-ki*, que convierte la predicación en un oficio.

- (9)a. *Ø-ki-piya i-pil*
 3SUJ-3OP-tener 3SG.POS-hijo
 ‘Tiene un hijo.’
- b. *ti-pil-piya*
 2SUJ-hijo-tener
 ‘Tú cuidas niños’
- c. *Ø-pil-pix-ki*
 3SUJ-hijo-tener-AG
 ‘Es cuidadora de hijos.’

Se tienen en el náhuatl de Amanalco muchos verbos de oficio de este tipo, que como se observa están basados en construcciones de IN por composición. Algunos ejemplos se dan en (10).

- (10)a. *ni-kal-pix-ki*
 1SUJ-edificio-tener-AG
 ‘Soy cuidador de casas (por oficio remunerado).’
- b. *ni-kalaʔ-pix-ki*
 1SUJ-interior.de.la.casa-tener-AG
 ‘Soy cuidador de la casa (temporalmente, como favor).’
- c. *ti-xochi-teh-ke*
 1SUJ-flor-cortar-AG.PL
 ‘Somos cortadores de flores.’
- d. *ti-nanaka-te:moʔ-ke*
 1SUJ-hongo-buscar-AG.PL
 ‘Fuimos buscadores hongos.’

La construcción de IN por composición es la estructura favorita para codificar actividades rutinarias de carácter cultural e institucionalizado. La estructura de IN por manipulación, que presentamos en la sección siguiente, produce verbos que refieren actividades que suelen protagonizar los seres humanos. Las que derivan de verbos transitivos son de uso y costumbre, pero las que se generan de verbos intransitivos refieren a casos periféricos del concepto de actividad

rutinaria, ya que expresan acciones en donde un ser humano se ve afectado de manera más individualizada que de manera atemporal.

2.2. Incorporación por manipulación de caso

La construcción de IN por composición que vimos en la sección anterior disminuye la valencia del verbo, de tal forma que el verbo compuesto derivado es un verbo intransitivo (N+V). En la IN por manipulación de caso, que corresponde al Tipo II en la tipología de Mithun (1984), la construcción permite que un participante no central avance a la posición de argumento verbal que ha sido dejada vacante por el sustantivo incorporado (Mithun 1984: 856).

En particular, en náhuatl, cuando un verbo transitivo incorpora un sustantivo de parte del cuerpo, el poseedor de la parte del cuerpo pasa a asumir el rol de objeto.⁵ De igual forma, cuando un verbo intransitivo incorpora su sujeto, otro argumento puede avanzar a la posición de sujeto. En el primer caso, el verbo resultante refiere típicamente a una actividad rutinaria. Considérense como ilustración los ejemplos en (11) y (12). Los ejemplos en (a) refieren a eventos puntuales en donde un agente humano realiza una acción sobre una parte del cuerpo de otra persona, produciendo una afectación local de la persona. En estos casos, la FN del objeto aparece poseída. Los ejemplos con IN en (b), en cambio, codifican al poseedor como objeto del verbo y esta codificación permite que el poseedor se presente como altamente afectado por la acción, es decir, confieren una lectura de afectación holística. El sustantivo de la parte del cuerpo en la construcción de IN ha perdido referencialidad y solamente indica el ámbito corporal en donde se efectúa la acción sobre el cuerpo del paciente.

(11)a. *ni-k-paʔpa:k* *mo-ma:n*
1SUJ-3OP-RED/lavar.PERF 2SG.POS-mano
'Lavé tu mano.'

b. *ni-mits-ma:-paʔpa:k*
1SUJ-2OP-mano-RED/lavar.PERF
'Te lavé las manos.'

(12)a. *ti-k-wila:na* *no-tsonkal*
2SUJ-3OP-jalar 1SG.POS-cabello
'Tú jalas mi cabello.'

b. *ti-ne:ch-tsonkal-wila:n*
2SUJ-1OP-cabello-jalar.PERF
'Me jalaste de los cabellos.'

⁵ A este tipo de IN se la ha identificado dentro de la literatura lingüística como un subtipo de "ascensión de poseedor" o "posesión externa" (Aissen 1979; Mithun 1984; Payne & Barshi 1999).

Lo crucial para el propósito en este artículo radica en el hecho que los ejemplos (a) y (b) no sólo contrastan semánticamente en grados de afectación sino también en aspectualidad. Los eventos codificados en (a) refieren acciones puntuales, mientras que la semántica resultante en los ejemplos de IN es de actividad rutinaria. Más concretamente, en el caso de (11b) se trata de ayudarle a una persona a lavarse las manos, como a un niño o un enfermo. En el caso de (11b), se trata de una expresión convencional de castigo.

Para los verbos intransitivos, los únicos que admiten la construcción de IN en náhuatl son los inacusativos (para más detalles ver Peralta-Ramírez 2013). De estos se tienen a su vez dos tipos dependiendo de su origen: básicos y derivados. Ejemplos de ambos tipos se dan en (13-16). En estos ejemplos, la construcción analítica se emplea siempre para un evento puntual, que ocurre de manera esporádica. Nótese que la parte del cuerpo en estos casos es el sujeto inanimado del verbo. En contraste, los verbos resultantes de la construcción de IN tienen siempre un sujeto animado, prototípicamente humano, y designan estados o acciones que son esperables dentro de una condición crónica del individuo, como por ejemplo una enfermedad. Nótese, sin embargo, que la semántica de los verbos compuestos que resultan no es de actividad rutinaria *strictu sensu*, como en los casos anteriores. Esto se debe a que los verbos son inacusativos y por lo tanto no tienen un sujeto agentivo que realice una actividad *per se*. No obstante, comparten con los verbos de actividades rutinarias su aspecto inherente imperfectivo.

- (13)a. \emptyset -*wa:ki* *mo-kamak*
 3SUJ-secarse 2SG.POS-boca
 ‘Tu boca se seca.’ (p.ej. como consecuencia de no haber bebido)
- b. *ti-kama-wa:ki*
 2SUJ-boca-secarse
 ‘Se te seca la boca/Tienes la boca seca.’ (p.ej. por una enfermedad o por cansancio)
- (14)a. \emptyset -*ki:sa* *no-yes-yo*
 3SUJ-salir 2SG.POS-sangre-INHER
 ‘Sale mi sangre.’ (p.ej. me acaban de pinchar y la veo salir)
- b. *ni-yes-ki:sa*
 1SUJ-sangre-salir
 ‘Sangro.’ (p.ej. como síntoma inherente a una enfermedad dada)
- (15)a. \emptyset -*kopi:-ni* *no-tlan*
 3SUJ-arrancarse-NI 1SG.POS-diente
 ‘Mi diente se chispa.’ (p.ej. al morder algo duro)
- b. *ni-tlan-kopi:-ni*
 1SUJ-diente-arrancarse-NI
 ‘Tengo los dientes chispados’ (me he quedado chimuelo)

- (16)a. *Ø-posa:-wi no-ma:n*
 3SUJ-hincharse-WI 1SG.POS-mano
 ‘Mi mano se hincha.’ (p.ej. por un golpe o una picadura de un insecto)
- b. *ni-ma:-po?posa:-wi*
 1SUJ-mano-RED/hincharse-WI
 ‘Se me hinchan las manos.’ (p.ej. debido a una enfermedad crónica)

2.3. Desiderativos y predicación compleja

A parte de las dos construcciones de IN, para la creación de elementos léxicos que designan actividades rutinarias, en el náhuatl existe asimismo otra construcción de composición verbal pero esta vez formada por dos verbos: el verbo nuclear, que funciona como verbo matriz desiderativo, y otro verbo que funciona como su complemento y que aparece flexionado en modo irrealis. Un ejemplo de la construcción de predicación compleja se da en (17), en donde el complemento se indica entre corchetes.

- (17) *ni-k-[kwa:-s-]negi naca-tamal-i*
 1SUJ-3OP-comer-IRR-querer carne-tamal-ABS
 ‘Deseo comer un tamal de carne.’

El ejemplo de predicación compleja en (17) contrasta con la estructura en (18) en donde el complemento del verbo desiderativo es una cláusula con sintaxis menos reducida.

- (18) *ni-k-neki [ni-k-kwa:-s naka-tamal-i]*
 1SUJ-3OP-querer 1SUJ-3OP-comer-IRR carne-tamal-ABS
 ‘Quiero comer un tamal de carne.’

Nótese que la estructura de predicación compleja en (17) resulta de la integración léxica del predicado subordinado a la matriz. El modo irrealis con el que se marca el verbo integrado es una reliquia morfológica del modo que recibe el predicado en la cláusula de complemento en (18).⁶

Muchas de las instancias de la construcción de predicación compleja se emplean en náhuatl para expresar actividades rutinarias. Ejemplos de diferentes tipos aparecen en los ejemplos siguientes.

- (19) Alimentación
- a. *ti-[tla-kwa:-s-]neki-h*
 1SUJ-ANTIP-comer-IRR-querer-PL
 ‘Deseamos comer.’

⁶ Desde un punto de vista tipológico, los verbos desiderativos son junto a los manipulativos, modales y aspectuales, aquellos que suelen requerir de complementos con una sintaxis más reducida (Cristofaro 2003).

- b. *ti-[tlayi-s-]neki-ti*
2SUJ-emborrarse-IRR-querer-AND
'Deseas emborracharte.'
- (20) Etapas de vida
- a. *ti-mo-[na:miki-s-]neki*
2SUJ-RFLX-casarse-IRR-querer
'Deseas casarte.'
- b. *ti-mo-[cha:nti:-s-]neki*
2SUJ-RFLX-construir.casa-IRR-querer
'Deseas vivir aparte.'/'Deseas construir tu propia casa.'
- (21) Situaciones de vida
- a. \emptyset -[cho:ka-s-]neki
3SUJ-llorar-IRR-querer
'Desea llorar.' (p.ej. no tienes control del llanto)
- b. \emptyset -[ko:-kochi-s-]neki
3SUJ-RED-dormir-IRR-querer
'Está dormitando.' (p.ej. tiene mucho sueño y anda cabeceando)
- (22) Actividades cotidianas
- a. *ni-[tla-pa:ka-s-]neki*
1SUJ-ANTIP-lavar-IRR-querer
'Deseo ponerme a lavar la ropa.'
- b. *ni-[tekiti:-s-]neki*
1SUJ-trabajar-IRR-querer
'Deseo ponerme a trabajar.'

Además del verbo *neki* 'querer' se emplea también el verbo *miki* 'morir' con semántica desiderativa intensiva como en (23). La estructura con este segundo verbo matriz no tienen alternativa analítica.

- (23)a. *ni-[kochi-s-]miki*
1SUJ-dormir-IRR-morir
'Me quiero dormir.'
- b. *ni-[tla?-tlami:na-s-]miki*
1SUJ-ANTIP-orinar-IRR-morir
'Tengo ganas de orinar.'

En esta sección hemos mostrado que en náhuatl las construcciones de IN por composición y por manipulación de caso, así como la construcción de predicación compleja se emplean como estrategias prácticamente dedicadas a la expresión de actividades rutinarias. Esto se debe a que el complejo verbal derivado por estas construcciones forma una unidad más compleja de significación. En el caso de la IN, en el complejo N+V, el N pierde referencialidad, y sirve para clasificar el

ámbito de aplicación de la acción verbal. En el tipo de IN de composición, la acción del verbo se presenta de manera imperfectiva, lo cual supone un marco gramatical muy adecuado para expresar actividades que se llevan a cabo de manera usual. En la siguiente sección, exploramos el modo cómo se codifican las actividades rutinarias en otomí.

3. El trato de las actividades rutinarias en otomí

El término “otomí” se trata tradicionalmente como designatario de una lengua otomangue a la que se le reconocen diferentes variantes. Este trato resulta conveniente, no sólo porque es la postura tradicional, sino porque evita tomar compromisos sobre cómo caracterizar, para efectos estrictamente lingüísticos, la diversidad lingüística interna al sistema otomí (ver Lastra 2001, para un audaz estudio exploratorio de esta diversidad).⁷ Afortunadamente, nuestro conocimiento de esta diversidad como gran diasistema lingüístico ha ido en aumento en los últimos años a medida que hemos ido sabiendo más de la realidad de otras variantes a parte de las del Valle de Mezquital y del estado de Querétaro. Para romper con el impasse de esta tradición, la diversidad del otomí

⁷ La posición tradicional se justifica en dos puntos: inteligibilidad mutua e identificación etnolingüística. Para el primer punto, se suelen barajar porcentajes de inteligibilidad mutua altos entre las diferentes variantes, pero solamente de manera impresionista y por lo tanto imprecisa. Al respecto, el único estudio métrico de inteligibilidad existente entre las variantes del otomí aparece en Andrews *et al.* (1983[1978]), y en ese estudio los porcentajes altos se deben tomar como indicadores de variantes que en realidad son dialectos de una lengua común, de la que no hay un estándar ni forma consensuada de designarla. Existen otras variantes que en sí mismas representan lenguas independientes, que suelen corresponder de manera natural a grados de inteligibilidad bajos. Por otro lado, en la posición tradicional se aboga que los hablantes de otomí se reconocen entre sí como compartiendo la identidad étnica de ser otomíes. Por lo tanto, en esta tradición, el término “otomí” se emplea primero para designar una identidad étnica y después se entiende asociado a ella para designar de forma vaga al sistema lingüístico que sirve de exponente cultural fundamental en la caracterización y formación de esta identidad. Esta posición no es de gran ayuda para entender la diversidad lingüística del otomí. En este sentido, “otomí” sería *grosso modo* equivalente a decir que los hablantes de francés, español, portugués e italiano, hablan todos “latino” o “romance”, ya que existe una cierta identidad cultural -no étnica, en este caso- de ser “latino” frente a ser “anglo-sajón” o “germánico”. Esta forma de concebir la diversidad lingüística romance, nos obligaría a decir que en Francia se habla el “latino de París”, mientras que en España el “latino de Madrid, de Asturias, de Sevilla, etc.” y en Latinoamérica, por ejemplo, docenas de latinos, todos aquellos de los que se pueda registrar cierta variación regional, comenzando por el “latino de Ciudad de México”, frente a los latinos de Tlaxcala, de Tuxtla Gutiérrez, de Buenos Aires, de Córdoba, etc. Una posición de este tipo, no se comprometería ni con las diferencias entre el “latino de París” y el “latino de la Ciudad de México”, ni con las semejanzas entre el “latino de la Ciudad de México” y el “latino de Madrid”, las trata a todas por igual. Qué el lector juzgue por sí mismo si este hipotético escenario serviría de ayudar al lingüista que estudia lenguas romances a comenzar a hacer intentos para entender la diversidad lingüística de esta familia y traslade el símil al otomí.

se puede caracterizar alternativamente como representando una pequeña familia de lenguas cuyo número está todavía por determinarse consensualmente.⁸

Para el propósito de este trabajo, empleamos datos provenientes de tres lenguas otomíes distintas: (i) el otomí de Tilapa, (ii) el otomí septentrional, representado aquí por las variantes del Valle del Mezquital en Hidalgo y de San Idefonso Tultepec en Querétaro, que resultan muy diferentes entre sí, pero que se deben interpretar como variantes o dialectos de una misma lengua desde un punto de vista lingüístico, no sociolingüístico, y (iii) el otomí oriental, representado por la variante del municipio de Huehuetla, hablada en San Antonio el Grande y San Gregorio.

3.1. La clase flexiva de verbos de actividad

En todas las lenguas otomíes, existe una clase de verbos intransitivos que recibe un trato flexivo distintivo. Salvo algunas excepciones, los verbos de esta clase tienen un sujeto animado, típicamente humano, y casi en su totalidad refieren actividades humanas. Dentro de esta clase se encuentran por default los verbos que refieren típicas actividades rutinarias, como los que aparecen en (24a). Sin embargo, muchos otros miembros de la clase indican otro tipo de actividades que realizan usualmente los seres humanos bien en su calidad de seres humanos o en su calidad de miembros de una sociedad. Ejemplos aparecen en (24b). Los ejemplos en (24) provienen del otomí septentrional de la variante del Valle del Mezquital.

(24) Otomí septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004)

a.	<i>hēt'i</i>	'hilar'
	<i>'āfi</i>	'raspar maíz'
	<i>xófo</i>	'cosechar'
	<i>xadi</i>	'estudiar'
	<i>'énts'á-pa</i>	'dar felicitaciones, celebrar'
	<i>'ák-hñüni</i>	'servir la comida'
	<i>fá-'yo</i>	'pastorear'
	<i>'ú'-thyä</i>	'hacer carbón'
	<i>'ü-ngü</i>	'construir casa'
	<i>tán-thä</i>	'comprar maíz'

⁸ La posición oficial al respecto, apoya el reconocimiento de esta diversidad. En el catálogo del Instituto Nacional de Lenguas Indígenas (CLIN 2008) se reconocen nueve variantes diferentes, a las que se da el estatus de lenguas para fines públicos. Sin embargo, la delimitación lingüística de esta división no está basada en datos dialectológicos, sino en impresiones de especialistas y representantes de ciertas comunidades, por lo que esta división no es indicativa del grado de diversidad lingüístico del sistema.

b.	<i>hi</i>	‘bañarse’
	<i>’é-ñä</i>	‘peinarse’
	<i>tsí-’yui</i>	‘fumar’
	<i>tühü</i>	‘cantar’
	<i>dähñëi</i>	‘flojear’
	<i>k’üntsi</i>	‘renguear, cojear’
	<i>gë</i>	‘gruñir’
	<i>gä’tsi</i>	‘suspirar’
	<i>húxi</i>	‘chiflar’

En el otomí de Tilapa, los verbos de esta clase flexiva reciben material afijal diferente del que se emplea en la flexión de otros verbos. Esto se puede ver comparando algunas de las formas del verbo *xófo* ‘cosechar, pizar maíz’ en (25), que es un miembro de esta clase flexiva especial, con las formas flexivas de un verbo de la clase flexiva mayor, como por ejemplo el verbo intransitivo *Vü* ‘morir’ en (26). En ambos casos se emplean los mismos valores de persona y TAM. La flexión de los dos verbos es la misma para las celdas en (a-b), pero distinta para (c-e).

(25) Otomí de Tilapa

a.	{1, PRS.CONT, R}	<i>txá=xófo</i>	‘estoy cosechando’
b.	{1, IMPF.CONT, R}	<i>txámá=xófo</i>	‘estaba cosechando’
c.	{1, IMPF.ADV, R}	<i>támádi=xófo</i>	‘estaba cosechando con algo’
d.	{1, PSD, R}	<i>túdu=xófo</i>	‘coseché’
e.	{1, PRS, IRR}	<i>gutú=xófo</i>	‘cosecharé’

(26)a.	{1, PRS.CONT, R}	<i>txá=^htü</i>	‘me estoy muriendo’
b.	{1, IMPF.CONT, R}	<i>txámá=^htü</i>	‘me estaba muriendo’
c.	{1, IMPF.ADV, R}	<i>támátí=^htü</i>	‘me estaba muriendo de algo’
d.	{1, PSD, R}	<i>tú=^htü</i>	‘me morí’
e.	{1, PRS, IRR}	<i>gü=^htü</i>	‘me voy a morir’

De la misma forma, en el otomí oriental, los verbos de esta clase flexiva también reciben un trato distinto. Esto se puede ver tanto en los afijos del presente de irrealis, como en que el verbo tiene un tema nasal en todos los tiempos del paradigma menos en el presente y el imperfecto de realis. El caso se muestra con los verbos cognados *xófo* ‘cosechar, pizar maíz’ y ^h*tü* ‘morir’.

(27) Otomí Oriental: Huehuetla

a.	{1, PRS, R}	<i>dí=xófo</i>	‘estoy cosechando’
b.	{1, IMPF, R}	<i>dmí=xófo</i>	‘estaba cosechando’
c.	{1, PSD, R}	<i>dá=nxófo</i>	‘coseché’
d.	{1, PRS, IRR}	<i>da=nxófo</i>	‘cosecharé’

(28)a.	{1, PRS, R}	<i>dí=tü</i>	‘me estoy muriendo’
b.	{1, IMPF, R}	<i>dmí=tü</i>	‘me estaba muriendo’
c.	{1, PSD, R}	<i>dá=tü</i>	‘me morí’
d.	{1, PRS, IRR}	<i>ga=tü</i>	‘me voy a morir’

En el otomí septentrional, los verbos de esta clase se distinguen solamente porque tienen un tema nasal. No hay diferencia en los afijos empleados.

(29) Otomí Septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004)

a.	{1, PRS, R}	<i>dí=xófo</i>	‘estoy cosechando’
b.	{1, IMPF, R}	<i>ndí=xófo</i>	‘estaba cosechando’
c.	{1, PSD, R}	<i>dá=ñxófo</i>	‘coseché’
d.	{1, PRS, IRR}	<i>ga=ñxófo</i>	‘cosecharé’

(30)a.	{1, PRS, R}	<i>dí=tü</i>	‘me estoy muriendo’
b.	{1, IMPF, R}	<i>ndí=tü</i>	‘me estaba muriendo’
c.	{1, PSD, R}	<i>dá=tü</i>	‘me morí’
d.	{1, PRS, IRR}	<i>ga=tü</i>	‘me voy a morir’

La clase de estos verbos de actividad es cerrada, es decir, no admite más verbos. Los préstamos se flexionan de acuerdo a otras clases flexivas dependiendo de la lengua en cuestión. En Palancar (2012), la clase de los verbos de actividad representa casi el 12% de un corpus de 600 verbos. Para el otomí oriental, el diccionario de Echegoyen & Voigtlander (2007) contiene 416 entradas para esta clase frente a un total de 2000 verbos (20%). De estos 416 verbos, 124 verbos son verbos medios (v.g. están marcados con un prefijo nasal con función de voz media o antipasivo, ver Palancar 2006). Para el otomí septentrional, el diccionario de Hernández Cruz *et al.* (2004) contiene 178 entradas de un total de casi 2,100 verbos (8.5%). Los verbos medios en las diferentes variantes del otomí septentrional se flexionan como parte de la clase default.

En la clase flexiva de verbos de actividad, encontramos verbos básicos y verbos derivados. Dentro del corpus del otomí del Valle del Mezquital al que tenemos acceso para los verbos de esta clase, que aparece en el Apéndice 1, los verbos básicos de la clase suman 54 y representan el 30%. El resto de los verbos en la clase son derivados (124/70%). De los verbos derivados, 106 (85%) son compuestos, transparentes u opacos, y prácticamente todos denotan actividades rutinarias. Es decir, en otomí como en náhuatl, la IN por composición es uno de los mecanismos favoritos para la creación de léxico que refiere actividades rutinarias. Ejemplos ilustrativos de las características de estos compuestos aparecen en (31).

(31) Otomí septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004)

	V INTR.		V		N
a.	'á-the	'sacar agua'	'ái	(TR) 'sacar líquido'	+ déhé 'agua'
b.	'á-n-sēi	'sacar pulque'			+ sēi 'pulque'
c.	'ák-sēi	'servir el pulque'	'áki	(TR) 'sacar líquido con cuchara'	+ sēi 'pulque'
d.	'ák-hñūni	'servir la comida'			+ hñūni 'comida'
e.	'é-'wada	'quebrar maguey' ⁹	'egi	(TR) 'dar un tajo'	+ 'wada 'maguey'
f.	'é-tfi	'quebrar maguey'			+ t'afi 'aguamiel'
g.	pá-m-ngo	'vender carne'	pá	(TR) 'vender'	+ ngo 'carne'
h.	kún-the	'nadar'	kūni	(INTR) 'meterse'	+ déhé 'agua'

La composición verbal es un proceso de creación léxica muy antiguo en la familia otomí y de él han resultado los numerosos compuestos que se tienen hoy en día para poder nombrar un gran número de actividades rutinarias de manera muy específica, hasta el punto de que se pueden encontrar sinónimos, como por ejemplo (31b-c) o (31e-f). Sin embargo, es necesario subrayar que la regla de composición verbal no es productiva sincrónicamente. Esto tiene consecuencias para el léxico.

Los lexemas compuestos como los de (31) representan unidades léxicas convencionalizadas que se deben aprender como lexemas independientes en cada lengua y variante. Por otro lado, existen restricciones morfológicas. Por ejemplo, el lexema déhé 'agua' tiene dos temas: *déhé* y *-the*. El primero (*déhé*) se emplea como palabra libre en sintagmas nominales y en compuestos nominales (p.ej. *t'ax-déhe* {ser.blanco-agua} 'agua limpia'). El segundo es un morfo ligado que aparece en compuestos como por ejemplo en (31a) y (31h). Por otro lado, en otros muchos casos se tienen interfijos nasales que no responden a ninguna regla morfofonológica productiva, como por ejemplo en (31b) y (31g).¹⁰ Por último, el compuesto consiste normalmente de un verbo transitivo más un tema nominal que refiere al paciente, pero existen casos aislados como (31h) en donde el verbo base es intransitivo, y el nominal es una locación, o casos en donde el nominal no es precisamente el paciente del verbo transitivo, sino el producto culturalmente relevante de la actividad en cuestión, como en (31f).

⁹ "Perforándolo por un lado del cogollo para rasparlo" (Hernández Cruz *et al.* 2004).

¹⁰ Un dictaminador sugiere que esta nasal es productiva en el otomí colonial como exponente de una regla de nominalización. Lamentablemente, desconocemos a que morfema se está haciendo referencia, ya que los procesos de nominalización generalizados tanto en la lengua otomí antigua como en las lenguas modernas responde a procesos de conversión de verbos a partir del tema estativo que en efecto presenta un prefijo nasal similar al que se hace referencia, y que es cognado del marcador *na-* en mazahua. En los casos discutidos, el interfijo nasal en el compuesto '*á-n-sēi* 'sacar pulque' puede en efecto ser la reliquia de un prefijo estativo, ya que el sustantivo moderno *sēi* 'pulque' resultó en época histórica de la conversión del tema impersonal del verbo *xēi* 'perforar, excavar'. Para el caso de *pá-m-ngo* 'vender carne', desconocemos su origen.

Como en el náhuatl, el compuesto verbal describe la acción de manera holística e imperfectiva. El tema nominal carece de referencialidad. Esto se puede ver en el contraste entre (32a) y su contraparte analítica en (32b), en donde los magueyes se presentan como un conjunto de entidades no especificadas previamente en el discurso.

(32) Otomí septentrional: Valle del Mezquital

- a. $dí = \underline{e} - 'wada = he$ $ha\ ra\ wä'ta$
 1.PRS.R=dar.tajo.TA-maguey=PL.EXCL P SG magueyal
 'Estamos quebrando maguey en el magueyal.'
 (Adaptado de Hernández Cruz *et al.* 2004: 330)

- b. $mä = ga = \underline{e} = he$ $'ra\ ya\ 'wada$
 INM=1.PRS.IRR=dar.tajo.D[3OBJ]=PL.EXCL INDF.PL PL maguey
 $ga = 'at = he$ $ngü'ä\ s-tä = ñäts'i$
 1.PRS.IRR=raspar[3OBJ]=PL.EXCL antes ANTER-PRS.IRR=envejecerse[3OBJ]
 'Vamos a capar unos magueys para rasparlos antes que sazonen.'
 (Adaptado de Hernández Cruz *et al.* 2004: 58)

Ya hemos mencionado que los verbos en los que se basan los compuestos de esta clase son en su mayoría transitivos, y por lo tanto, reciben un trato flexivo diferente. Esto se puede ver si se comparan las mismas celdas del paradigma de dos verbos en cuestión. En (33) se contrastan algunas formas flexivas del verbo de actividad $\underline{e} - 'wada$ (intr) 'quebrar maguey' de (32a) con las del verbo \underline{egi} (tr) 'cortar, dar un tajo' de (32b).¹¹

(33) Otomí Septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004)

- | | | | |
|----|-----------------|-------------------------------|---------------------------|
| a. | { 1, PRS, R } | $dí = \underline{e} - 'wada$ | 'estoy quebrando maguey' |
| b. | { 1, IMPF, R } | $ndí = \underline{e} - 'wada$ | 'estaba quebrando maguey' |
| c. | { 1, PSD, R } | $dá = \underline{ñe} - 'wada$ | 'quebré maguey' |
| d. | { 1, PRS, IRR } | $ga = \underline{ñe} - 'wada$ | 'quebraré maguey' |
-
- | | | | |
|--------|-----------------|-------------------------|----------------------|
| (34)a. | { 1, PRS, R } | $dí = \underline{egi}$ | 'lo estoy cortando' |
| b. | { 1, IMPF, R } | $ndí = \underline{egi}$ | 'lo estaba cortando' |
| c. | { 1, PSD, R } | $dá = \underline{egi}$ | 'lo corté' |
| d. | { 1, PRS, IRR } | $ga = \underline{egi}$ | 'lo voy a cortar' |

Por otro lado, entre los verbos de la clase de actividad se encuentran asimismo muchos derivados mediante el sufijo *-te*. Estos denotan actividades propias de humanos o de animales, que si bien no son en sí actividades rutinarias en una comunidad, se presentan como comportamientos usuales o característicos de los individuos sobre los que se predicán. En general, todo verbo derivado con *-te* en

¹¹ El verbo \underline{egi} (tr) 'dar un tajo' aparece tanto en el compuesto de (32a) como en la cláusula de (32b) en un tema reducido (\underline{e} -) debido a condicionamientos prosódico-sintácticos.

una lengua otomí, suele pertenecer por default a la clase flexiva de verbos de actividad. Algunos ejemplos se dan en (35).

(35) Otomí septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004)

V INTR.		V TR.	
a. <i>hó-te</i>	‘matar’	‘ser un asesino’	<i>hó</i> ‘matar’
b. <i>fás-te</i>	‘prestar servicio’	‘ser ayudador’	<i>fáts’i</i> ‘ayudar’
c. <i>tsá-te</i>	‘morder’	‘ser mordedor’	<i>tsá</i> ‘morder’
d. <i>tä-te</i>	‘ganar’	‘ser un ganador’	<i>tähä</i> ‘ganar’
e. <i>xáx-te</i>	‘provocar’	‘ser un provocador’	<i>xáxi</i> ‘provocar’

3.2. La construcción de complementación por nominalización

Si bien la clase flexiva de verbos de actividad contiene en su mayoría verbos que tienen una semántica típica de actividades rutinarias en todas las lenguas otomíes independientemente de las características flexivas específicas de la clase en cuestión en cada lengua, existe en todas estas lenguas otra construcción que involucra verbos de actividades rutinarias exclusivamente. Se trata de una construcción de complementación que consiste de un verbo matriz y una nominalización verbal intransitiva que funciona como su complemento, por ejemplo en (36), (Hernández Cruz *et al.* 2004).

(36) Otomí septentrional: Valle del Mezquital (Hernández Cruz *et al.* 2004: 12)

mände bi=mfa’yo ra Xuwa
 ayer 3.PSD=pastorear SG Juan
nu mí=ma [ra ’beni] rá nänä
 cuando 3.IMPF.ADV=TN/ir SG lavada SG.3POS madre
 ‘Juan pastoreaba ayer mientras su madre se fue a lavar.’

En (36), el verbo de movimiento *pa* ‘ir’, funciona como un verbo matriz de movimiento propositivo, cuyo complemento eventivo aparece codificado en la FN *ar ’beni* ‘la lavada’, que contiene una nominalización del verbo *peni* ‘lavar’. En el otomí septentrional la construcción permite muy pocos verbos matrices. Por ejemplo, en la variante de San Ildefonso Tultepec se pueden emplear para este propósito solamente *pa* ‘ir a hacer algo’, *fáts’i* ‘ayudar a hacer algo’ y *jwádi* ‘terminar de hacer algo’. En otras lenguas, como el otomí de Tilapa, y sobre todo el otomí oriental, la construcción permite un gran número de verbos matrices (para más detalles ver Palancar, por aparecer), que incluyen desiderativos como en (37) o verbos de cognición como en (38).

(37) Otomí de Tilapa

txá=nde [ra t’ëni]
 1.CONT.R =querer[3OBJ] SG.NOMLZ juego
 ‘Quiero jugar.’

(38) Otomí Oriental: Huehuetla (Voigtlander & Echegoyen 1979: 232)

i=päh=[na nkuyä] yu ts'unt'u
 3.PRES.R=saber.D[3OBJ]=INF.INTR echar.clavados PL jóvenes
 'Los muchachos saben bucear.'

En todos los casos, la construcción se emplea solamente con nominalizaciones verbales que refieren actividades rutinarias típicas; es decir, aquellas cuya realización supone un marco de acción, temporal y espacial, esperable dentro del espacio social de la comunidad lingüística. En este sentido, un ejemplo como (39a), resulta agramatical debido a que la acción denotada en la nominalización (v.g. “el acto de saltar”) no es una actividad rutinaria. Para expresar el mismo contenido, se debe emplear en otomí una cláusula finita yuxtapuesta, que está subordinada al predicado matriz, como en (39b).

(39) Otomí septentrional: San Ildefonso Tultepec

- a. **nu m'í=ma [ra ntsagi] ka=r ma'ye*
 cuando 3.IMPF.ADV=TN/ir SG saltada P=SG barranca
 Lectura buscada: ‘Cuando se fue a saltar por la barranca.’
- b. *nu m'í=ma [nda=n-tsa]'=ka=r ma'ye]*
 cuando 3.IMPF.ADV=TN/ir 3.PRES.IRR.DEP=TN-saltar.D=P=SG barranca
 ‘Cuando se fue a saltar por la barranca.’

En (39), tanto el verbo matriz *pa* ‘ir’ como el verbo *tsagi* ‘saltar’ son verbos intransitivos que pertenecen a la clase flexiva de verbos de actividad. El hecho de que la nominalización de *tsagi* ‘saltar’ no se pueda emplear en el marco de la construcción en (39a) sugiere que no todos los miembros de esta clase se perciben como casos de actividades sancionadas por el colectivo como típicamente rutinarias, y que por ello no pueden emplearse en la construcción. El hecho de que la mayoría de los verbos que designan actividades rutinarias se encuentren en la clase es debido a que éstas se tratan como un subgrupo de actividades típicamente humanas.

4. Resumen de la propuesta

Las actividades rutinarias representan un campo semántico poco explorado en el estudio de la gramática de las lenguas mesoamericanas. Este artículo representa un primer paso en este intento. En el artículo, hemos mostrado que en la gramática de dos sistemas lingüísticos muy distintos del área mesoamericana, como son el náhuatl y el otomí, existen construcciones cuya función prototípica es la de generar o encapsular lexemas que denoten actividades rutinarias como verbos intransitivos. En este sentido, en las variantes del náhuatl, las construcciones verbales de incorporación nominal o verbal son mecanismos usuales para codificar acciones como actividades rutinarias. En las lenguas

otomíes, en cambio, el conjunto del léxico verbal que designa actividades rutinarias está limitado. Como en náhuatl, la mayoría de los lexemas que hoy en día expresan este tipo de actividades se produjeron en su día mediante composición, pero todos los lexemas que expresan este tipo de actividades, independientemente de que sean básicos o derivados, reciben un trato flexivo distintivo de otros verbos, perteneciendo a una clase que alberga verbos de actividad humana. El estudio es exploratorio desde un punto de vista tipológico. Queda por ver si en otras familias del área lingüística mesoamericana propuesta en Campbell *et al.* (1986) se encuentran construcciones similares tanto en su tipología como en su carácter funcional específico.

Referencias

- Andrews, Richard J. 1975. *Introduction to Classical Nahuatl* 1. Austin: University of Texas Press.
- Aissen, Judith. 1979. Possessor Ascension in Tzotzil. In *Papers in Mayan Linguistics*, Laura Martin (ed.), 89-108. Columbia: Lucas Brothers.
- Andrews, Henrietta, Bartholomew, Doris, Blight, Richard, Echegoyen, Artemisa, Lanier, Nancy, Sinclair, Donald, Voigtlander, Katherine, Eglan, Steven, Mellema, Paul & Persons, David. 1983 [1978]. Otomi. In *La inteligibilidad interdialectal en México: resultados de algunos sondeos*, Steven Eglan (ed.) in collaboration with Doris Bartholomew & Saúl Cruz Ramos, 46-48. México: Instituto Lingüístico de Verano.
- Baker, Mark C. 1988. *Incorporation: a Theory of Grammatical Function changing*. Chicago: University of Chicago Press.
- Campbell, Lyle, Kaufman, Terrence & Smith-Stark, Thomas. 1986. Mesoamerica as a linguistic area. *Language* 62: 530-70.
- CLIN (Catálogo de las Lenguas Indígenas Nacionales). 2008. Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. *Diario de la Nación* (14-01-2008). México.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Echegoyen, Artemisa & Voigtlander, Katherine. 2007. *Diccionario Yuhú: Otomí de la Sierra Madre Oriental (Estados de Hidalgo, Puebla y Veracruz, México)*. Versión Preliminar. Serie de Vocabularios y Diccionarios Indígenas. México: Instituto Lingüístico de Verano.
- Hernández Cruz, Luis, Victoria Torquemada, Moisés & Sinclair Crawford, Donaldo (eds). 2004. *Diccionario del hñähñu (otomí) del Valle de Mezquital*,

- Estado de Hidalgo* (Vocabularios Indígenas 45). México: Instituto Lingüístico de Verano.
- Kroeber, Alfred. L. 1909. Noun Incorporation in American Languages. *XVI. Inter-nationaler Amerikanisten-Kongress*: 569-576
- Merlan, Francesca. 1976. Noun incorporation and discourse reference in modern Náhuatl. *International Journal of American Linguistics* 42.3: 177-191.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60: 847-894.
- Lastra, Yolanda. 2001. *Unidad y Diversidad de la Lengua: Relatos otomies*. México: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Palancar, Enrique L. 2006. Middle Voice and Intransitivity in Otomi. *Linguistics* 43.1: 613-643.
- Palancar, Enrique L. 2012. The conjugations classes of Tilapa Otomi: An approach from canonical typology. *Linguistics* 50.4: 783-832.
- Palancar, Enrique L. Por aparecer. On the tightrope between infinitives and action nouns: The case of Otomi nominalizations. In *Finiteness and Nominalization*, Claudine Chamoreau & Zarina Estrada-Fernández (eds). Amsterdam: John Benjamins.
- Payne, Doris L. & Barshi, Immanuel. 1999. External Possession: What, Where, How, and Why. In *External Possession*, Doris L. Payne & Immanuel Barshi (eds), 3-29. Amsterdam: John Benjamins.
- Peralta-Ramírez, Valentín. 2013. Posesión externa en el náhuatl de Amanalco. In *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica*, Enrique L. Palancar & Roberto Zavala (eds), 103-132. México: CIESAS.
- Peralta-Ramírez, Valentín. 2014. La incorporación nominal en el náhuatl de San Jerónimo Amanalco, Texcoco Edo. de México. In *Lenguas, estructuras y hablantes. Estudios en Homenaje a Thomas C. Smith Stark*, Rebeca Barriga y Esther Herrera (eds), 737-758. México: El Colegio de México.
- Rosen, Sara T. 1989. Two types of noun incorporation: a lexical analysis. *Language* 65: 294-317.
- Sadock, Jerreold M. 1980. Noun incorporation in Greenlandic: a case of syntactic word formation. *Language* 56: 300-319.
- Sapir, Edward. 1911. The problem of noun-incorporation. *American Anthropologist* 13: 250-83. [Reimpresión 1990. In *The Collected Works of Edward Sapir*. American Indian Languages V, William Bright (ed.), 27-59. Berlín: Mouton-de Gruyter]
- Voigtlander, Katherine & Echevoyen, Artemisa. 1979. *Gramática del yuhú: Otomí de la Sierra Madre Oriental*. México: Instituto Lingüístico de Verano.

Apéndice 1. Verbos de actividad en el otomí del Valle del Mezquital

BÁSICOS			
<i>bi</i>	‘temblar temblar’	<i>’ǔ+s+t’i</i>	‘antar’
<i>bu</i>	‘zumbar’	<i>’ús+pi</i>	‘prender lumbre, hacer lumbre’
<i>dǎn+ts’i</i>	‘quejarse’	<i>’u+i</i>	‘urdir construir’
<i>fǒ+ge</i>	‘ladrar reventarse, azotarse’	<i>’yó</i>	‘andar, caminar valer, estar a’
<i>gǎ’+tsi</i>	‘suspirar’	<i>’ú</i>	‘doler’
<i>gá+ni</i>	‘tronar, sonar, roncar’	<i>kún+gi</i>	‘desmoronarse, deshacerse’
<i>gě</i>	‘gruñir’	<i>pa</i>	‘ir’
<i>gó+fo</i>	‘estar estreñado’	<i>sá+gi</i>	‘brincar, saltar’
<i>háha</i>	‘jadear, exhalar, dar boqueada’	<i>tǎ+ki</i>	‘revolcarse’
<i>há+ki</i>	‘quitarse’	<i>tǎ+ni</i>	‘revolcarse’
<i>hě+i</i>	‘recordar, pensar en alguien’	<i>té=tho</i>	‘vivir todavía’
<i>hehe</i>	‘toser’	<i>thé+de</i>	‘reírse’
<i>hě+t’i</i>	‘hilar’	<i>thí’+tsi</i>	‘estar abajo de’
<i>hé+ts’e</i>	‘estornudar’	<i>tihi</i>	‘correr’
<i>hi</i>	‘bañar’	<i>tón+ts’i</i>	‘rebosar, correrse’
<i>hya, hyáha</i>	‘lamar, vahear’	<i>té+xi</i>	‘tallar’
<i>hú+xi</i>	‘chiflar’	<i>ts’á+t’i</i>	‘atorarse’
<i>já</i>	‘pasar, acontecer’	<i>tsu</i>	‘asustarse’
<i>jo’o</i>	‘no estar no aparecer’	<i>tsú</i>	‘mamar’
<i>juá+ts’i</i>	‘recargarse, apoyarse’	<i>tsú+m’i</i>	‘pepenar chile’
<i>’á+fi</i>	‘raspar’	<i>tsí+m’i</i>	‘comer chile’
<i>’bē+ni</i>	‘estar acostado/tirado/tendido’	<i>tú+ni</i>	‘pelear’
<i>k’ě</i>	‘pujar, quejar’	<i>túhu</i>	‘cantar’
<i>k’o</i>	‘cacarear gruñir’	<i>xa+di</i>	‘estudiar’
<i>k’ua</i>	‘tronar hacer estampido sonar’	<i>xó+fo</i>	‘cosechar’
<i>k’ún+ts’i</i>	‘renguear, cojear’	<i>xu+ni</i>	‘pixcar’
<i>’ó=tho</i>	‘no hay’	<i>yú+ni</i>	‘oler’
<i>’ó+xi</i>	‘hospedarse, quedar la noche’		
<i>kó’+te</i>	‘eructar’		
Con raíz disilábica: Compuestos arcaicos			
<i>’bēbo</i>	‘estar abajo, estar debajo’	<i>hyóya</i>	‘encontrarse pobre’
<i>’ěhya</i>	‘robustecerse, recuperar la salud’	<i>k’uamba</i>	‘mentir’
<i>dá’thi</i>	‘estar enfermo, estar en cinta’	<i>séya</i>	‘resentir’
<i>děspi</i>	‘estar hecho brasa’	<i>tsáya</i>	‘descansar’
<i>dǎhñei</i>	‘flojear’	<i>yóbri</i>	‘arrepentirse’
<i>gógu</i>	‘molestar’		
DERIVADOS			
Derivados con -te			
<i>’áká-te</i>	‘acechar, espiar’	<i>hó-te</i>	‘matar embestir’
<i>’én-te</i>	‘aventar, tirar’	<i>tǎ-te</i>	‘ganar, vencer triunfar’
<i>’ěnts’á-te</i>	‘hablar indirectas’	<i>tsán-te</i>	‘blasfemar, maldecir’
<i>’e-te</i>	‘pensar vengarse’	<i>tsá-te</i>	‘morder’
<i>’ún-te</i>	‘lastimar afligir’	<i>tsú-te</i>	‘regañar’
<i>fás-te</i>	‘prestar servicio colaborar’	<i>tsǒn-te</i>	‘visitar’
<i>fétá-te</i>	‘remedar imitar burlarse’	<i>xás-te</i>	‘provocar’
<i>ha-te</i>	‘engañar’		

Compuestos

'á'-thä	'desvelar'	hü- 'ni	'hornear gallina'
'a'k-hñuni	'servir la comida'	hü- 'uada	'hornear pencas de maguey'
'ak-sei	'sacar pulque'	hü-hme	'hornear pan'
'á-mañ'u	'rogar por mal, desear un mal'	hü-jua	'hornear conejo'
'án-sēi	'sacar pulque, servir pulque'	hüm-ngq	'hornear carne'
'á-the	'sacar agua'	hü-nāni	'hornear cal'
'e- 'wada	'quebrar maguey, capar maguey'	hwí'-tŋi	'tlachiquear'
'et-fi	'capar maguey, quebrar maguey'	hyá'-nga	'estar ronco'
'en-do	'arrojar piedras'	k'ē-suni	'moler nixtamal'
'ē-nsagi	'brincar, echar brinco'	k'ū-ne	'tartamudear'
'ēnt'a-hme	'echar tortillas, tortillar'	k'ōm-nxāt'ä	'extraer carnosidad del nopal'
'ēnts'á-pa	'dar felicitaciones, celebrar'	kā-bqhai	'desempolvar'
'ä-x'yo	'trasquilar ganado lanar'	kā-do	'construir'
'ba-ne	'tartamudear'	kūn-the	'nadar'
'ék'-uāmba	'hacer mentiroso, contradecir'	kōngā-bāyo	'llevar puesto rebozo'
'é-ñä	'peinar'	pá'-nthä	'dormitar cabecear'
'ū-ngu	'construir casa'	pá-dethä	'vender maíz'
'etsua-rá-nsu	'alabar, elogiar, ensalzar'	pām-'i	'vender chiles'
'ok-jä	'cortar tuna'	pām-ngo	'vender carne'
'ot'á-'bai	'pararse, hacer parada, orinar'	pá-nthi	'vender lazos, vender mecates'
'ot'á-buhla	'burlarse'	pū-'hmi	'majar la palma'
'ot'á-ngo	'hacer fiesta'	pōn-ju	'sacar frijol'
'o-fádi	'estar en la cárcel'	tām-hai	'comprar tierras'
'ū'-thyä	'hacer carbón'	tán-thä	'comprar maíz'
'yo-mfo	'caminar con decencia'	thén-the	'acarrear agua'
'yo-mxúdi	'andar de madrugada'	thé-sēhe	'reírse solo'
'yo-gua	'andar a pie, caminar descalzo'	tix-fāni	'estar borracho'
dó-'y e	'renguear'	ts'ān-za	'lancear, formarse el tallo'
fa-'yo	'pastorear'	ts'ok-'bāha	'despreciar no agradecer'
fūx-huāhi	'barbechar'	tsá-'ni	'comer gallina'
gá'n-thä	'roncar'	tsēká-zä	'cortar árboles o leña'
go-ne	'azotarse'	tsēn-za	'cortar leña o árbol'
há-ne	'bostezar'	tsē-za	'cortar árboles'
hēká-ndäpo	'segar alfalfa segar hierba'	tsi 'm'i	'comer chile'
hén-the	'cortar surco, rayar surco'	tsi-'yui	'fumar'
hén-za	'cortar palos'	tsi-the	'beber agua'
hā-thä	'traer maíz'	tsú 'm'i	'pepenar chile'
hā-spi	'pedir lumbre'	tsú-dēnxi	'pepenar cebollas'
has-tai	'amanecer el día después del tianguis'	tsú-ju	'pepenar frijol'
hét'ä-nthähi	'hilar para hacer mecate'	tsú-thä	'pepenar maíz'
hén-thähi	'hilar'	tsú-jūmhai	'pepenar cacahuates'
hō'-thä	'fingir estar dormido'	tsut-'yui	'fumar cigarro'
hō-'yo	'matar ganado menor'	xān-the	'sudar'
hōnd-gu	'construir casa'	xū-'yē	'lavar las manos'
hō-ndri	'caer las flores'	xis-tehe	'bautizar'
hōngá-'yo	'buscar ganado menor'	xós-thūhu	'levantar la planta de maíz'
hōx-g'mi	'poner al fuego el agua de cal'	zōn-da	'lagrimear los ojos, llorar los ojos'
hōx-yä	'estar greñudo'		

Denominales	
<i>'bēgo</i>	‘estar de peón’
<i>zīhu</i>	‘enfurecerse’
Denumeral	
<i>yoto</i>	‘ser siete’

The emergence of active/stative alignment in Otomi

ENRIQUE L. PALANCAR

14.1 Introduction

In this chapter, I show that Otomi intransitive verbs display both an agent/patient and an active/stative semantic alignment patterns. Agent/patient alignments have been widely described in the literature (to mention just a few, Merlan 1985, Van Valin 1990, Mithun 1991, Dixon 1994). Less is known about the active/stative type of alignment in languages such as Guaraní, as proposed in Mithun (1991). The goal of this chapter is twofold: first, I characterize the two alignment splits in this language, and secondly I describe the active/stative alignment in more detail in order to propose how this alignment may have emerged diachronically.

Otomi is spoken in Central Mexico and it belongs to the Otopamean group of Otomanguan languages. The label 'Otomi' in fact refers to a number of closely related languages. The phenomenon I study here is found in all varieties. Every language, however, has its own distribution of the phenomenon I study here, and differences in the treatment and scope of semantic alignments both on the lexical level and on the structural level are expected. While I also use other sources, most of the data in this chapter come from the variant spoken in the village of San Ildefonso Tultepec, in the state of Querétaro, Mexico. I refer to this dialect as San Ildefonso Otomi.¹

This chapter is organized as follows. In the following section, I present a brief overview of the grammatical features of San Ildefonso Otomi which are relevant to understand the object of study. Next, in 14.3 and in 14.4, I describe the semantic alignments found in the language. In 14.5, I study a subclass of relevant verbs pertaining to the active/stative alignment system, which I call 'stative verbs'. In

¹ A growing number of speakers of Mezquital Otomi prefer to call their language 'Hñahñu', among other spellings. Speakers of San Ildefonso Otomi do not recognize the Mezquital word as a representative name for their own language, 'Hñóhñó' being their native denominative. 'Otomi', on the other hand, is a term which enjoys some history in linguistics and anthropology, and at present it is a denomination not regarded to my knowledge as a pejorative term by the majority of native indigenous people belonging to this ethnic group.

14.6, I introduce a special grammatical form of transitive verbs, labelled 'stative form', which I claim is related to stative verbs. In 14.7, I compare stative verbs and stative forms, and in 14.8 I present my proposal regarding the diachronic emergence of the active/stative alignment in Otomi. The chapter concludes in 14.9.

14.2 Overview of the language

San Ildefonso Otomi is a language with a basic SVO word order, although VS is also common in intransitive clauses. While noun phrases (NP) are not case-marked, the verb agrees with subject by means of proclitics which also encode Tense/Aspect/Mood (TAM).² Examples in (1) illustrate three different proclitics: *dá* in (1a) encodes 1st person and completive past; *dí* in (1b) encodes 1st person and present tense. Present tense is used to express habitual aspect. The zero morpheme in (1b), associated with the verb *hō*, encodes 3rd person and present tense.³

- (1) a. *dá* = *húd-i* pá *dá* = *k'ó'í-Ø-i*
 1.PST=sit.down.F-F to 1.PST=look-3OBJ-F
 'I sat down to look at it.'
- b. *ya* *nixi* = *r* *hyéthé* *dí* = *tsi-Ø* = *hé*
 PARTICLE not.even=SG mushroom 1.PRES=ingest-3OBJ=PLEX
ko ya 'ín-té \emptyset = *hō* *ya zá*
 because PARTICLE NEG-what 3.PRES=grow.plant PL tree
 'We don't even eat mushrooms now, because there are no trees.'

A paradigm of the most common verbal proclitics found in San Ildefonso Otomi is given in Table 14.1.⁴

First and 2nd person objects show suffixal agreement, as in (2). Third person object is encoded by a zero morpheme, as illustrated in (3):

- (2) *yá* *gá* = *tōn-k-i*, *pé* *ga* = *peñg-i*
 PARTICLE 2.PST=win.S-1OBJ-F but 1.IRR=return-F
ga = *tsix-i*
 1.IRR=take.animate.S-2OBJ-F
 'You've beaten me (this time), but I'll come back to get you.'

² In a number of dialects, the imperfect is a distinct morph (e.g. *dí-mé-nú* [1.PRES-IMPERF-see] 'I was seeing (it)', from Toluca Otomi; Lastra 1992: 23).

³ I have used the convention of Hekking (2002) for the transcription of Otomi, with two innovations: *r* for /r/ and *l'* for a salient aspiration existing in San Ildefonso Otomi. Deviations from the IPA convention are: *ʔ* /f/; *ʔ* /k^h/; *j* /k^h/; *y* /j/; *x* /ʃ/; *tx* /tʃ/; *ʃ* /ʃ/; *r* /r/; *a* /a/; *ɛ* /e/; *ɔ* /o/; *u* /i/. Umlaut indicates nasalization (i.e. *a* /ã/; *i* /ĩ/; etc.). Otomi has three tones: ascending *v*; high *v*; and low, which is not represented. Given that tone differences between San Ildefonso Otomi and Mezquital Otomi are minimal, I have consulted the excellent transcriptions in Wallis (1956) and Hernández Cruz et al. (2004) for uncertain cases in San Ildefonso Otomi. Conventions: (-) morpheme boundary in source; (<) non-concatenative morpheme; = clitic.

⁴ Paradigms can be very complex in other dialects (cf. Voigtlander and Echegoyen 1985).

TABLE 14.1. Verbal proclitics in San Ildefonso Otomi

Subject + TAM	1st	2nd	3rd
Present (habitual)	<i>dí</i>	<i>gí</i>	\emptyset / <i>bí</i> / <i>i</i>
Imperfect (past)	<i>n-dí</i>	<i>n-gí</i>	<i>mí</i>
Irrealis (future)	<i>ga</i>	<i>gí</i>	<i>dá</i>
Past (completive)	<i>dá</i>	<i>gá</i>	<i>bí</i>
Perfect	(<i>x</i>) <i>tá</i>	(<i>x</i>) <i>ká</i>	<i>xi</i>
Pluperfect			(<i>x</i>) <i>ki</i>

- (3) *yá* *hín-té* *dí* = *tō-Ø* = *ma* *bojó*
 PARTICLE NEG-what 1.PRES=win.B-3OBJ-B=IPOSS money
 'I don't earn my money.'

In Table 14.2, I present the object suffixes. Besides the use of affixes, a 1st and 2nd person object may optionally be cross-referenced in the verb by additional object enclitics, as illustrated in (4):

- (4) *yá* *gá* = *h* < *y* > *and* - *g* - *a* = *tō-Ø* = *gí*
 PARTICLE 2.PST=<NPS> see-1OBJ-B=DEL=1OBJ
 'You caught me red-handed.' ~ (Mex. Sp.) 'Me cachaste.'
 (lit. 'You just saw me.')

Doubling object agreement by means of both an affix and an enclitic is very common in San Ildefonso Otomi.⁵ In principle, the enclitic must always co-occur with the suffix, but there are exceptions. When a verb has a disyllabic root, i.e. it does not have a transparent stem formative, e.g. *künt'et* 'keep a jealous eye on someone', it cannot receive the object suffix, and object cross-reference is thus exclusively encoded by means of enclitics (e.g. *mí* = *künt'et* = *gi* [3.IMPERF=keep.a.jealous.eye.on=1OBJ] 'He was keeping a jealous eye on me' vs. **mí* = *künt'et* = *g-a* = *gí*). This distribution is, however, somehow reversed with stative verbs, which is a class of verbs highly relevant for this study (see 14.5). Not only can stative verbs receive the enclitic alone, regardless of the morphology of the root, but the use of the suffix alone is ungrammatical, unless in combination with the enclitic.⁶

⁵ In Palancar (2006), the combination of suffix plus enclitic was erroneously analysed as a case of doublet object morphology (e.g. *-gagí*). The present analysis is superior, as example (4) illustrates: here the delimitative clitic = *tō* occurs between the personal morphemes, implying that there is no morphological unity between *-g-a* and *-gí*. A similar phenomenon concerning double agreement marking has been described in Donohue (2003) for Skou, a non-Austronesian language spoken at the central north coast of New Guinea. In this language, verbs agree with subject by means of proclitics, but most commonly, consonantal prefixes are also used along with the proclitics.

⁶ The rationale behind this special distribution may lie in the fact that stative verbs do not have the inflectional contrast denoting free/bound from that other verbs have (see 14.5.3).

TABLE 14.2. Object markers in San Ildefonso Otomi

1st	-k/-g
2nd	-
3rd	-Ø

14.3 Semantic alignment in Otomi

Most intransitive verbs in San Ildefonso Otomi agree with S by means of the same proclitics used to encode A in transitive verbs, and differently from O, which shows suffixal agreement.⁷ This pattern, which is illustrated in (5) and (6), can be described as showing an accusative morphological alignment. Examples in (5) show the use of the proclitics *dí=* and *da=*. The former is a marker of present tense and first person S, the second of irrealis and third person S. Both markers are used to encode A in (6):

- (5) S a. *yá dí = zá = 'wá* *dí = 'bú = híú*
 PARTICLE 1.PRES=be.agreeable=here 1.PRES=live/be.S=PL.(INCL)
 'We are fine here, living together.'
 b. *Ø = hót = 'á* *go_gé da = <g>ó = 'á*
 3.SG=be.only=3SG REL.CLEFT(3) 3.IRR=<NPS>remain.S=3SG
 'That is the only thing that remains.'

- (6) A a. *nú = gá dí = pá-Ø* *nú = 'ú* *ya ts'oi*
 DEF=1SG 1.PRES=sell-3OBJ DEF=DEM.PL PL pot
 'I'm selling these pots.'
 b. *da = xít'-Ø-a = nó = r* *ts'oi*
 3.IRR=wash.dishes-3OBJ-B=DEF.SG=SG pot
 'She'll wash up the pot.'

However, a number of intransitive verbs in Otomi index S by means of the same morphology encoding O in transitive clauses. The phenomenon is illustrated in (7), which may be compared to the examples in (4) and the morphemes in Table 14.2 above:

- (7) a. *da = <z>ó^h-k-a = gí*
 (3.)IRR=<NPS>burn.S-1OBJ-B=1OBJ
 'I will/would get burned.'
 b. *hi-mí = pá-²-a = 'í*
 NEG-(3.)IMPERF=be.hot-2OBJ-B=2OBJ
 'You weren't hot.'

⁷ I use here the tripartite characterization of grammatical relations in Dixon (1994), which distinguishes three main categories: A, O, and S.

I call the intransitive verbs in (5) the A-verbs, while the verbs showing the object marking of S in (7) are called the O-verbs. O-verbs in (6) are inflected for TAM by means of proclitics, just like other verbs in the language. However, given that verbal proclitics in San Ildefonso Otomi cannot express TAM separately from grammatical person, O-verbs are always inflected by means of a neutral (or dummy) 3rd person proclitic. This may be seen in the glosses of (7), where the reference to this 3rd person has been placed in parentheses.

Given that the language has zero pronominal anaphora for a 3rd person object, as illustrated in (3) above, the semantic alignment shown in (7) is covert when the intransitive verb in question is inflected in 3rd person, as indicated in (8):

- (8) *pé 'ingí Ø = káhi-Ø,* *ká = Ø = n-jú-Ø*
 but NEG (3.)PRES=be.tasteful-3OBJ INTENS=(3.)PRES=st-be.bitter-3OBJ
 'But it isn't tasteful, it's quite bitter.'

14.4 Types of semantic alignment in Otomi

Otomi is an interesting language typologically because intransitive verbs in this language display the two types of semantic alignment proposed in Mithun (1991): agent/patient and active/stative. While A-verbs remain the same in the two alignments, the language has two types of O-verb. In this section I describe these two alignments separately, but I will focus later on the study and emergence of the active/stative alignment.

14.4.1 The agent/patient semantic alignment

A great number of intransitive verbs in Otomi display an agent/patient alignment. While the class of A-verbs here is enormous, the class of O-verbs is rather limited, containing a handful of verbs in each dialect. I call the O-verbs in this type of alignment 'patient verbs', as a way to distinguish them from the O-verbs in the active/stative alignment. For the variant of San Ildefonso Otomi, I have been able to identify only ten patient verbs so far, but there may well be more. Such verbs are given in (9):

- (9) Patient verbs in the agent/patient alignment
- | | | |
|----|----------------------------|--|
| a. | <i>tsá bi</i> | 'be/get tired' |
| b. | <i>pó bo</i> | 'be/get wet' |
| c. | <i>'yehmi</i> | 'be/get drenched' |
| d. | <i>tsó'ti</i> | 'be/get burnt' |
| e. | <i>mí'ti</i> | 'get one's hair stand on end' |
| f. | <i>tít'si</i> | 'be/get better (from an illness)' |
| g. | <i>tsá yé</i> | 'be/get worse (from an illness)' |
| h. | <i>sólo = sé/hónit'á</i> | 'be just/only' (i.e. 'it's just/only') |
| i. | <i>t'áxkt'⁸</i> | 'be/get white/pale' |
| j. | <i>há't'i</i> | 'be/get pale' |

⁸ Both verbs *t'áxkt'* 'be/get white' and *há't'i* 'be/get pale' are used of people when they get pale from an illness or from fright. The verb *t'áxkt'* is an inchoative verb related to the stative *t'áxkt'* 'be white'.

The verbs *tsábi* 'be/get tired' in (9a) and *póbo* 'be/get wet' in (9b) have the relics of allomorphs of an old object marker {-b}, whose reflexes in the modern languages have been reanalysed as dative markers (cf. Palancar 2004a). Historically, the class appears to have been larger. Urbano (1990[1605]: 43) mentions another four lexical items in passing: *piti* <*piti*> 'be/get afraid'; *kweri'i* <*kwenti*> 'swoon'; *hímui* <*hímucej*> 'be/get glad' (lit. 'sit-heart/belly'); and *hwáts'i* <*hwátsisi*> 'tremble'. All these verbs depict events that participants have but little control over. Lack of control is an important semantic feature in the emergence of similar splits in other languages (cf. e.g. Mithun 1991, Merlan 1985), and it could well have been the defining semantic feature of patient verbs in earlier stages of the language. Nowadays, many A-verbs depict uncontrolled events too. This may be seen in the non-exhaustive list in (10):

(10) A-verbs depicting uncontrolled events

- | | | |
|----|---------------|----------------------------------|
| a. | <i>səya</i> | 'be/get angry' |
| b. | <i>jóhyá</i> | 'be/get happy' |
| c. | <i>tánthü</i> | 'be/get hungry' |
| d. | <i>táthe</i> | 'be/get thirsty' |
| e. | <i>tági</i> | 'fall down to a different plane' |
| f. | <i>hwangi</i> | 'fall down on the same plane' |
| g. | <i>tsó</i> | 'fall into a hole' |

Similar agent/patient alignments have been described for other Otomanguan language groups, such as in Amuzgan and Chocho-Popolocan (for Amuzgo, see Smith-Stark and Tapia García 2002; for Chocho, Veerman-Leichsenring 2000; for Popolocan, Swanton 2005; and for Tlapanec, Wichmann 1996). It remains an open question whether similar semantic alignments exist in other families of this macro-phylum.

14.4.2. The active/stative alignment

Intransitive verbs in Otomi also display an active/stative alignment, similar to that found in other languages from South America such as Guaraní (Mithun 1991, Velazquez-Castillo, this volume) and Kurripako (Danielsen and Granadillo, this volume). The class of O-verbs in this active/stative alignment is much larger than the class of O-verbs in the agent/patient pattern we saw in the previous section. I will refer to the O-verbs in this alignment as 'stative verbs'. Stative verbs depict property concepts, which are the type of concept commonly encoded with adjectives in other languages (e.g. the property of being 'red', 'big', 'long'). A number of verbs of this class are given in (11) (see Appendix 14.1 for a non-exhaustive list of such verbs, and Palancar 2006 for more details):

(11) A few O-verbs of the active/stative alignment

- | | | |
|----|-------------|------------|
| a. | <i>pídi</i> | 'be thick' |
| b. | <i>k'á</i> | 'be wet' |

- | | | |
|----|-----------------|-------------|
| c. | <i>nóxi</i> | 'be stale' |
| d. | <i>bextüi</i> | 'be dark' |
| e. | <i>'ü</i> | 'be sweet' |
| f. | <i>hwé'is'i</i> | 'be golden' |

A number of A-verbs also express property concepts which are ascribable to human beings, even though a number of them may be extended to characterize animals too. A relevant list of such A-verbs in San Ildefonso Otomi is given in (12):⁹

- | | | | |
|----------------|------------------|--------------------|--------------|
| (12) DIMENSION | <i>dótiá</i> | 'be big' | (also Human) |
| | <i>ts'álo</i> | 'be small' | (also Human) |
| | <i>yá'pu</i> | 'be far away' | |
| TIME/AGE | <i>tsu</i> | 'be old (woman)' | (only Human) |
| | <i>x'ta</i> | 'be old (man)' | (only Human) |
| | <i>ts'unt'ü</i> | 'be young (man)' | (only Human) |
| | <i>nixántsi</i> | 'be young (woman)' | (only Human) |
| | <i>'rá'yo</i> | 'be new' | |
| PROPENSITY | <i>nzá_Tho</i> | 'be beautiful' | (only Human) |
| PHYSICAL | <i>tse = tho</i> | 'be strong' | (only Human) |

The last two verbs in (12) can also be inflected as stative verbs, but when they do so, they depict an inherent property. This is shown in (13a) if compared with (13b):¹⁰

- | | | | |
|---------|--------|---|--|
| (13) a. | A-verb | <i>dí = nzá_Tho</i> | |
| | | 1.PRS=be.beautiful | |
| | | 'I look fine (because something becomes me).' | |
| | | <i>dí = tse = tho</i> | |
| | | 1.PRS=be.strong=DEL | |
| | | 'I'm strong (I feel so).' | |
| b. | O-verb | <i>xí = nzá_Tho = gí</i> | |
| | | (3.)PERF=be.beautiful=IOB] | |
| | | 'I'm beautiful (by nature).' | |
| | | <i>xí = n-tse = gí</i> | |
| | | (3.)PERF=ST=be.strong=IOB] | |
| | | 'I'm strong (by nature).' | |

⁹ The verbs *xíta* 'be old (for a man)', *ts'unt'ü* 'be young (for a man)', and *nixántsi* 'be young (for a woman)', are derivations from the nouns *xíta* 'old man', *ts'unt'ü* 'young man', and *nixántsi* 'daughter'. The lexeme *tsu* receives low tone when it is a verb 'be old (for a woman)' but high tone when a noun 'old woman'.

¹⁰ Similarly, the verb *solo=se* [be.only=ALONE] 'be just/only' in (8g) may also inflect both as an A-verb and as an O-verb, but without an apparent semantic distinction, e.g. (patient verb) *ntf=solo=se-gí-gí* >=*gí* [(3.)IMPERF=bc.only=ALONE=IOB] vs. (A-verb) *n_dí=solo=se* [IMPERF_1.PRS=bc.only=ALONE] 'I was alone/it was only me'.

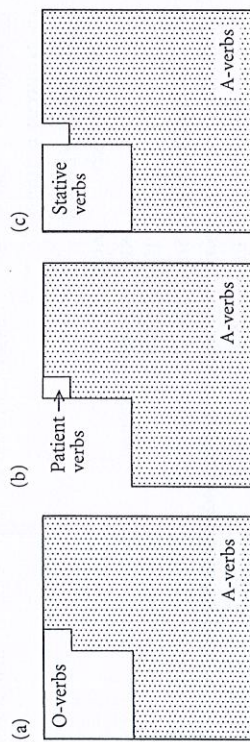


FIGURE 14.1. The split between O and A intransitive verbs in Otomi

14.4.3 Summary of the semantic alignment patterns

A summary of the semantic alignment patterns found in Otomi is given in the above figures. Figure 14.1(a) represents the body of intransitive verbs in the language, which has two main classes: the O-verbs and A-verbs. Figure 14.1(b) represents intransitive verbs that display an agent/patient alignment. Recall that the O-verbs in this alignment pattern are called ‘patient verbs’. Figure 14.1(c) shows the intransitive verbs of the language that display an active/stative semantic alignment. In this alignment pattern, O-verbs are called ‘stative verbs’ (O-verbs considered here do not include patient verbs).

At least one O-verb in San Ildefonso Otomi may be inflected both as a patient verb and as a stative verb, with a slightly different meaning. The verb is *háat* ‘be/get pale’, and is shown in (14):

- (14) a. Patient O-verb *xí=h<y>át-k-a=gí*
(3.)PERF=<NPS>be/get.pale.S-1OBJ-B=1OBJ
‘I’m pale (as a result of a long illness).’
- b. Stative O-verb *xí=ñ-hát-k-a=gí*
(3.)PERF=ST-be/get.pale.S-1OBJ-B=1OBJ
‘I’m pale (as a result of a fright).’

14.5 A closer view of stative verbs

Apart from the encoding of their semantic argument as O—the defining feature of the class—stative verbs have three further characteristics that differentiate them clearly from most A-verbs: (a) they use a special paradigm of proclitics; (b) they have a morphologically conditioned nasal prefix; and (c) they lack a morphosyntactic bound form. I study each of these features separately in the following subsections. For each case, I additionally present another realm of the grammar where each characteristic is also found, but without an apparent or obvious connection with stative verbs. In my opinion, the idiosyncrasies of stative verbs point to a complex historical relationship with those areas of grammar

where the characteristics are found; they reveal the origin of these verbs and, ultimately, provide hints concerning the emergence of the active/stative semantic alignment in Otomi.

14.5.1 The TAM inflection of stative verbs

Stative verbs inflect in the imperfect and the irrealis with a special set of verbal proclitics. Such proclitics can also be used with all the A-verbs that express property concepts (PC) in (12) above. The pattern is shown in the following examples. The proclitics *már* = in (14a) and *dar* = in (16a) express the imperfect and the irrealis for stative verbs. As (15b) and (16b) show, these proclitics may be used with A-verbs that express property concepts but not with others, as the impossibility of (15c) and (16c) indicates. Notice that the subject translation is different when a stative verb and an A-verb are used as the proclitics with A-verbs cross-reference subject:

- (15) Imperfect Stative verb a. *már/mí = ñ-héts’i = gí*
(3.)IMPERF.PC/3-IMPERF=ST-be.high-1OBJ
‘I was tall.’
- PC A-verb b. *már/mí = dótá*
3-IMPERF.PC/3-IMPERF = be.big
‘He was tall.’
- Other A-verb c. **már/mí = ’bái*
3-IMPERF.PC/3-IMPERF=be.standing
‘He was standing.’
- (16) Irrealis Stative verb a. *dar/da = ñ-héts’i = gí*
(3.)IRR.PC/3-IRR=ST-be.high=1OBJ
‘I will/would be tall.’
- PC A-verb b. *dar/da = dótá*
3-IRR.PC/3-IRR=be.big
‘He will/would be tall.’
- Other A-verb c. **dar/da = ’bái*
3-IMPERF.PC/3-IMPERF=be.standing
‘He will/would be tall.’

The morphemes *már* and *dar* emerged from the reanalysis of clitic morphology involved in the nominal predication construction. In this construction, as illustrated in (17), the predicative NP appears inflected for number and is preceded by a special paradigm of proclitics marking person and TAM, some of which are only used in nominal predication. In (17a) and (17b), the singular enclitic = *r* is hosted by the proclitics *má*= and *da*=. These proclitics encode imperfect and irrealis of a third person.¹¹ The morphological combinations *má = r* and *da = r* gave way to the

¹¹ Because they occur in a nominal predication construction, Ecker (1952) and Hess (1968) erroneously treat these markers as inflectional forms of an alleged copular verb.

synchronic proclitics *már* and *dar* used with stative verbs. Notice that in (17c), the number of the NP is plural, but the combination *má = ya* used in this predication was not reanalysed any further. This may be seen in (17d), where a proclitic such as *máya* does not exist, and the verb receives the proclitic *már=* with a plural subject:

- (17) Imperfect+_{SG}NP a. *má=[r bɛ]*NP
 3.IMPERF.NP=SG thief
 'S/he was a thief.'
 Irrealis+_{SG}NP b. *da=[r mǎh]ó*NP
 3.IRR=SG priest
 'He'll become a priest.'
 Imperfect+_{PL}NP c. *má=[ya bɛ]*NP
 3.IMPERF.NP=PL thief
 'They were thieves.'
 Stative verb d. *nu=ya ndó mármáya n-ts' o-Ø*
 DEF=PL man (3.)IMPERF.PC ST-be.evil-3OBJ
 'The men were evil.'

14.5.2 The nasal prefix of stative verbs

Stative verbs receive a nasal prefix *N-* I call the 'stative prefix'. The stative prefix is not used when the verbal stem has a nasal consonant in initial position (e.g. *mɛ* 'be hard', *nɔho* 'be fat'). The prefix has three allomorphs with regular phonological conditioning: [n-] <*n*> in (18a) is the default realization of the morpheme; [ɲ] <*ɲ*> in (17b) appears before a glottal consonant ([ʔ, h]); and [m-] <*m*> occurs before a labial consonant ([p, pʰ, b]), as in (18c):¹²

- (18) a. *bi = <z> i-Ø ya t'ɛi nú xi = n-k'áŋgi-Ø*
 3.PST=<NPS> ingest-3OBJ PL wheat REL (3.)PERF=ST-be.green-3OBJ
 'They ate green wheat.' (lit. 'They ate the wheat which is green.')
- b. *má = ɲá-ɲ-hɛts'i = ɲ*
 (3.)IMPERF.PC=INTENS-ST-be.high=2OBJ
 'You were very tall.'
- c. *xi = m-pá-g-a = ɠi*
 (3.)PERF=ST-be.hot-1OBJ-B=1OBJ
 'I'm hot.'

The prefix *N-* is not used in cases where the stative verb receives the same TAM morphology as A-verbs. This happens, for example, in the negative of both the present and the imperfect, where the use of the prefix is ungrammatical. As an illustration consider (19) and (20), which show two circumstances where the

12. When the nasal is attached to a stem with an initial glottal-labial consonant cluster (e.g. [ʔhmX] or [ʔhmX]), as in 'bɛxiɪ 'be dank' or 'hɪmɔdi 'be expensive', the labial is dropped, and the nasal assimilates the labial feature, as [m-ʔX] or [m-hX] (e.g. 'm-ɛxiɪ or 'm-hɔdi).

inflection of a stative verb coincides with that of an A-verb. In such cases, the stative prefix is not used:¹³

- (19) Negative present
 a. A-verb *hingi Ø=né-i*
 NEG 3.PRES=dance-F
 'He isn't dancing.'
 b. Stative verb *hingi Ø=(*n-)tsɛ-g-a=ɠi*
 NEG (3.)PRES=(ST-)be.cold-1OBJ-B=1OBJ
 'I'm not cold.'
- (20) Negative imperfect
 a. A-verb *hi-mj=né-i*
 NEG-3.IMPERF=dance-F
 'He wasn't dancing.'
 b. Stative verb *hi-mj=(*n-)tsɛ-g-a=ɠi*
 NEG-(3.)IMPERF=(ST-)be.cold-1OBJ-B=1OBJ
 'I wasn't cold.'

A few stative verbs denoting colours may appear as nouns designating the colour in question; as nouns, they always bear the stative prefix *N-*:

- (21) *n-t'áxi* 'white (as a colour)'
n-théni 'red'
n-k'áxt'i 'yellow'
n-k'áŋgi 'blue/green'

14.5.3 The lack of the free/bound form inflectional contrast

The vast majority of A-verbs in Otomi have stem formatives (e.g. *tsɛ+t'i* 'get cold'; *k'ó+ts'i* 'go to see someone'; *zó+n-i* 'weep' etc.). In Palancar (2004b) I claim that verbs with stem formatives have two inflected forms which I called a 'free form' and a 'bound form'. A free form is phonologically free and is used before a prosodic and/or a clausal boundary. An example is given in (22a). A bound form, in contrast, is phonologically bound and occurs intraclausally, as shown in (22b):

- (22) a. Free form [*da=tsɛt'-i*]
 3.IRR=get.cold-F
 'It'll get cold (i.e. something in a fridge).'
 b. Bound form [*da=tsɛt'-a=ma ngó*] [*Ø dá=tá-Ø-i*]
 3.IRR=get.cold-B=IPOSS meat REL 1.PST=buy-3OBJ-F
 'The meat I bought is going to get cold (i.e. in the fridge).'

13. In contrast, examples (14) and (15) above show two cases where a stative verb does not receive the same inflectional morphology of A-verbs. Notice that the stative prefix was used in such cases.

These forms are not interchangeable. For example, a free form may not be used in the context of a bound form as shown in (23):

- (23) * $[da=ts\acute{e}t-i \quad ma \quad ng\acute{o}]$
 3.IRR=get.cold-F 1POSS meat

Intended reading: 'My meat is going to get cold.'

On the other hand, stative verbs lack this inflectional contrast, despite the fact that a considerable number of them have stem formatives just like A-verbs (e.g. $ts'\acute{u}+f-i$ 'be thin (wood, thread)'; $h\acute{e}+ts-i$ 'be high'; $th\acute{e}+n-i$ 'be red'). Consequently, the /i/ in the citation form of such verbs should be considered part of the root. The pattern is illustrated by the following examples. In (24b) we may observe that the stative verb $ts'\acute{u}t-i$ 'be thin' does not undergo the same morphological adjustment as the verb $ts\acute{e}t-i$ 'get cold' in (22b). In fact, the use of the expected bound form of such a verb is impossible, as shown in (24c):

- (24) a. $n\acute{o}=r \quad z\acute{a} \quad h\acute{i}ngi \quad \emptyset=ts'\acute{u}t-i-\emptyset$
 DEF.SG=SG wood NEG (3.)PRES=be.thin-3OBJ
 'The tree isn't thin.'
 b. $h\acute{i}ngi \quad \emptyset=ts'\acute{u}t-i-\emptyset \quad n\acute{o}=r$
 NEG (3.)PRES=be.thin-3OBJ DEF.SG=SG wood
 'The tree isn't thin.'
 c. * $h\acute{i}ngi \quad \emptyset=ts'\acute{u}t-\emptyset-a=n\acute{o}=r$
 NEG (3.)PRES=be.thin-3OBJ-B=DEF.SG=SG wood
 Intended reading: 'The tree isn't thin.'

Beside a couple of irregular lexical items, stative verbs are the only verbs that lack the inflectional contrast of verbs with stem formatives. In this respect, stative verbs behave morphologically just like other nouns that also have stem formatives (e.g. 'bos+p-i' 'ash'; $t'\acute{a}+f-i$ 'candy'; $\acute{g}+n-i$ 'poultry', $g\acute{u}t+n-i$ 'incense'). Nouns never show the inflectional contrast I have presented in this section.

14.5.4 Summary concerning stative verbs

In the previous sections I have described a number of morphological idiosyncrasies of stative verbs; these verbs are the O-verbs of the active/stative semantic alignment of Otomi. A number of interesting features have been presented: (a) stative verbs receive the proclitics $m\acute{a}r=$ and $dar=$ for the imperfect and the irrealis, and such markers originally developed from the inflectional morphology involved in nominal predications; (b) stative verbs receive a nasal prefix, which is used in the cases where the inflectional paradigm diverts from A-verbs; and (c) unlike A-verbs, the stative ones lack the free/bound form inflectional contrast, and they resemble nouns in this respect, at least on morphological grounds.

I propose a hypothesis about how stative verbs may have developed diachronically. This development sheds important light on how the active/stative alignment emerged in the grammar of Otomi. Before the active/stative alignment pattern developed, Otomi had only an agent/patient alignment. I propose here that stative verbs were once former A-verbs that denoted property concepts.¹⁴ In my view of the process, speakers used A-verbs for the linguistic expression of property concepts, but at times they would also use a special 'stative form' of such verbs for the same purpose. Such a stative form is described in some detail in the following section. Stative verbs emerged as the outcome of a blending of inflectional features pertaining to the A-verbs involved and of features proper to stative forms. One of the main characteristics borrowed from the latter was the encoding of the semantic argument of the verb as an object. This change motivated the emergence of stative verbs as a subclass of intransitive verbs, which triggered the appearance of a new active/stative semantic alignment in the language.

The purpose of the remaining part of this chapter is to elaborate on this development step by step. In order to achieve such a goal in 14.8, I will have to describe first what a 'stative form' is in the following section, and then compare such forms with stative verbs in 14.7.

14.6 Stative forms of transitive verbs

In all dialects of Otomi, transitive verbs have a special form, which has been called 'participle' in the Otomian descriptive tradition (cf. Ecker 1952, Hess 1968, Hekking and Andrés de Jesús 1984, Voigtlander and Echegoyen 1985, Bartholomew 2004; etc.). I prefer to call such forms 'stative forms', more in tune with the description of the verbal system in other Otomanguan languages, such as Zapotec (cf. Pickett et al. 2001).

Stative forms express the state resulting from the action depicted by a transitive verb. Two examples are given in (25). These examples show the stative forms of the transitive verbs 'peel' and $h\acute{o} \acute{t}se$ 'cook (by boiling)' and portray the actions they depict as resulting states:

- (25) a. $n-t-\acute{t}se-i$ b. $n-t-h\acute{o} \acute{t}se-e$
 ST-IMPER-peel.F-F ST-IMPER.cook.by.boiling.F-F
 'Peeled.' 'Cooked (by boiling).'

Stative forms are morphologically built on impersonal stems plus the stative prefix $N-$, used with stative verbs (cf. 14.5.2). In other words, the stative forms

¹⁴ I claim that Otomi had an agent/patient alignment pattern before it developed the active/stative one. It remains an open question why property concepts were not at this stage expressed by the patient verbs in the original agent/patient alignment. It seems, though, that the semantics of such verbs was confined to the expression of uncontrolled affecting events, and the crucial point was that properties were not regarded as events.

in (25) are built on the impersonal stems *t-òts*- 'IMPER-peel' and *t-hò* 'ts'- 'IMPER-cook (by boiling)' plus stative morpheme *N-*. However, the morphological process by which stative forms are created is no longer productive synchronically. Consequently, many verbs lack a stative form. The phenomenon is perhaps better understood as lexically conditioned, and the verbal lexemes that have a stative form vary considerably from dialect to dialect.

Given that stative forms and impersonal forms are related, it is relevant first to describe what an impersonal form is. The impersonal form of a transitive verb is used as an exponent of 'impersonal voice'. Functionally speaking, impersonal voice has the same function as the prototypical passive voice proposed in Shibatani (1985): it serves as a de-agentivizer construction where the agent is not present, because it is either unknown or not relevant. An example is given in (26) with the impersonal form *thò* 'is e of the verb *hò* 'ts'e 'cook by boiling':

- (26) *bi=t-hò* 's-*ò*-*a*=*ma*
 (3.)PST=IMPER-cook.by.boiling.B-3OBJ-B=1POSS beans
 'My beans were boiled (by someone whom I don't know).' (lit. 'Boiled my beans.')

The event in (26) is presented as happening without a specified agent. The construction does not allow the overt expression of an agent, be it an indefinite subject, as in (27a) or an oblique participant, as in (27b):

- (27) a. **ná* *bi=t-hò* 's-*ò*-*a*=*ma*
 one 3.PST=IMPER-cook.by.boiling.B-3OBJ-B=1POSS beans
 Intended reading: 'Someone boiled my beans.'
 b. **bi=t-hò* 's-*ò*-*a*=*ma*
 3.PST=IMPER-cook.by.boiling.B-3OBJ-B=1POSS beans LOC 1POSS child
 Intended reading: 'My beans were boiled by my child.'

Crucially, however, the impersonal voice is a subjectless structure that does not promote O to S; the patient remains morphologically encoded as object. This feature will prove to be important for the emergence of stative verbs.

On the other hand, as TAM cannot be morphologically dissociated from grammatical person, the verb in the impersonal voice is inflected with a dummy 3rd person proclitic only to express TAM. As with stative verbs, reference to this dummy grammatical 3rd person has been placed in parentheses throughout, except in example (27).

The following examples show the contrast between the active rendering of the verb *handi* 'see' in (28a) and its use in the impersonal voice in (28b). Example (28a) is an instance of the transitive verb *handi* 'see', which has two arguments: first, a 3rd person who plays the role of an agent (or an experiencer) and is encoded as subject by means of the proclitic *bi*=; and second, a 1st person who

plays the role of a patient (or a stimulus) and is encoded as object by means of both the suffix *-g* and the clitic =*gi*. In (28b), in contrast, the verb appears in the impersonal voice, and has the agent argument cancelled, thus having only one argument left: the 1st person. Semantically, the agent is still present—the action is understood as instigated by some agency after all—but the construction just focuses on the resultant state in which the patient is found after the event has taken place. However, this patient is not grammatically treated or raised as subject, but remains an object, just as in (28a), while the construction becomes subjectless:

- (28) Active form a. *bi=h<y>and-g-a=gi*
 3.PST=<NPS>see-1OBJ-B=1OBJ
 'He saw me.'
 Impersonal form b. *bi=t-hand-g-a=gi*
 (3.)PST=IMPER-see-1OBJ-B=1OBJ
 'I was seen.' (lit. 'Saw me.')

On the other hand, speakers use perfect morphology with impersonal forms to encode present time reference, and the reading obtained is clearly one of resultant state:

- (29) a. *nò=r thuhme yá* *xí=t-hé^hk-ò-i*
 DEF=SG bread PARTICLE (3.)PERF=IMPER.cut-F-3OBJ
 Resultant: 'Someone has cut the bread/the bread's been cut.'
 (lit. 'The bread, has cut it:')
 b. *xí=t-hót-* 'a=*í*
 (3.)PERF=IMPER-comfort-2OBJ-B=2OBJ
 Resultant: 'You've been comforted.' (lit. 'Has comforted you.')

Predicates with stative forms are based on the impersonal voice. Stative forms cancel the presence of the agent completely, and present the participant as just being in a given state. When a verb happens to have both forms, we obtain subtle semantic contrasts. An example of one such contrast is given in (30):

- (30) Impersonal a. *hín-txi=k<h>ót-k-a=gi*
 NEG-(3.)PERF=<IMPER>dress.someone-1OBJ-B=1OBJ
 Resultant: 'I'm not (properly) dressed.'
 (e.g., a nurse has left me in such a state).
 (lit. 'Hasn't dressed me').
 Stative form b. *hín-txi=n-k<h>ót-k-a=gi*
 NEG-(3.)PERF=ST-<IMPER>dress.someone-1OBJ-B=1OBJ
 State: 'I'm not (properly) dressed.'
 (e.g., I think that gowns in this hospital are too short).
 (lit. 'Isn't dressed me').

To conclude this section, it is important to mention that a large number of stative forms may be converted into nouns. As nouns, such forms are commonly action

nouns, result nouns, or instrument nouns.¹⁵ A number of examples are given in (31).¹⁶

- (31) Action noun: Mezquital Otomi a. *n- <S> áha*
ST-<IMPER>bathe
'bathe, bathing'
- Mezquital Otomi b. *n-t- ún-the*
ST-IMPER-give.to.3-water.B
'watering, irrigation'
- Result noun: Mezquital Otomi c. *n-t- úf'i*
ST-IMPER-sprinkle
'manure'
- Instrument noun: San Ildefonso Otomi d. *n-k <h> ó'mi*
ST-<IMPER>cover
'lid, cover'

In summary, stative forms of transitive verbs are intransitive forms that express states. When used predicatively, their notional subject is still encoded as object. Besides, while morphologically based on impersonal stems, these forms receive the same prefix used with stative verbs. In the following section, I shall compare the stative forms with stative verbs in order to show how they can be related.

14.7 Comparison of stative verbs and stative forms

14.7.1 Similarities between stative verbs and stative forms

Both stative verbs and the stative forms of transitive verbs are intransitive verbal forms with only one argument. In both cases, this argument is encoded in the verb by means of object suffixes. Semantically, both forms express states. Morphologically, they use the perfect proclitic *xí=* to express a present state of affairs,¹⁷ and they both use the stative prefix *N-*. All these similarities may be observed in (32):

- (32) Stative form: a. *xí=n-t-hóf-k-i*
(3.)PERF=ST-IMPER-comfort-1OBJ-F
'I'm comforted.'
- Stative verb: b. *xí=ñ-hěts'i=gi*
(3.)PERF=ST-be.high=1OBJ
'I'm tall.'

Finally, although with lexical restrictions, both forms may convert into nouns and may be used as such. Examples are given in (33):

¹⁵ Most abstract nouns in Otomi were once derived from stative forms, but nowadays many of these nouns are no longer recognized by speakers as being related to their cognate verbal lexemes. Additionally, when a verb lacks a stative form, the impersonal form may be used for similar purposes.

¹⁶ The examples from Mezquital Otomi come from Bartholomew (2004).

¹⁷ The morphology is obligatory for stative verbs but optional for stative forms.

- (33) Noun from a stative form: *n-ts < > é-zá*
ST-<IMPER>hack-wood
'axe'
- Noun from a stative verb: *n-t' áxi*
ST-be.spicy
'chili'

Having presented the similarities in this section, I introduce the main differences in the following section.

14.7.2 Differences between stative verbs and stative forms

Stative verbs and stative forms show a number of differences. While stative forms are always related to synchronically attested transitive verbs, stative verbs are an independent class of verbs and do not derive from any other verbs, transitive or intransitive.

Similarly, in predication, stative forms use the same proclitics used with other A-verbs, while they cannot co-occur with the typical proclitics employed with stative verbs. This is shown in (34):

- (34) Imperfect
Stative form: a. *mí* már=n-jün-Ø-i*
(3.)IMPERF/(3.)IMPERF.PC=ST-grind.corn.F-3OBJ-F
'It (the corn) was ground.'
- Stative verb: b. **mí/már=m-pá-g-a=gi*
(3.)IMPERF/(3.)IMPERF.PC=ST-be.hot-1OBJ-B=1OBJ
'I was hot.'

On the other hand, stative forms have the inflectional contrast between a free and a bound form, while stative verbs do not. This is shown in the following examples. Examples in (35) show that stative forms have this inflectional contrast like other A-verbs, while examples in (36) illustrate that stative verbs lack such a contrast:

- (35) Stative form
a. *ma suni mí=n-jün-Ø-i||*
1POSS corn.dough (3.)IMPERF=ST-(IMPER.)grind.corn.F-3OBJ-F
'My corn-dough was ground.'
- b. *mí=n-jüŋ-Ø-a=ma*
(3.)IMPERF=ST-(IMPER.)grind.corn.B-3OBJ-B=1POSS corn.dough
'My corn-dough was ground.'

- (36) Stative verb
a. *ma zá már=ñ-hěts'i-Ø*
1POSS tree (3.)IMPERF.PC=ST-be.high-3OBJ
'My tree was tall.'

- b. *már=ñ-hěs' i-Ø* *ma zǎ*
 (3.)IMPERF.PC=ST-be.high-3OBJ 1POSS tree
 'My tree was tall.'
- c. **már=ñ-hěs' -Ø-a=ma* *zǎ*
 (3.)IMPERF.PC=ST-be.high-3OBJ-B=1POSS tree
 Intended reading: 'My tree was tall.'

Finally, on the morphological level, stative forms are always used with the nasal prefix in all grammatical contexts, unlike stative verbs, which do not use it in cases where the TAM inflection coincides with A-verbs. This is shown in (37). In (37a), the stative prefix is not used when a stative verb receives the same proclitics as other A-verbs. However, in similar circumstances, the prefix is obligatory with a stative form, as in (37b):

- (37) Stative verb: a. *hingi Ø=(^{*}m)-pá-g-a=gi*
 NEG (3.)PRES=(ST)-be.hot-1OBJ-B=1OBJ
 'I'm not hot.'
- Stative form: b. *tó'be hingi*
 still NEG
Ø=n-t-hó'is'-Ø-e
 (3.)PRES=ST-IMPERF-cook.by.boiling.F-3OBJ-F
 'It isn't cooked yet.'

14.7.3 Summary of the comparison

A summary of similarities and differences seen in 14.7.1 and 14.7.2 is presented in Table 14.3. Despite the differences, the morphological similarities between

TABLE 14.3. Comparison of stative verbs and stative forms

Features	Stative forms	Stative verbs
<i>Similarities</i>		
Intransitive	Yes	Yes
O encoding of semantic argument	Yes	Yes
Use of perfect conveys present	Yes	Yes
Semantic states	Yes	Yes
Convert into nouns (restricted)	Yes	Yes
<i>Differences</i>		
Derived forms	Yes	No
Stative prefix in all contexts	Yes	No
Inflectional contrast (bound/free)	Yes	No
Use of <i>mi</i> = for Imperfect (like A-verbs)	Yes	No
Use of <i>már</i> = for Imperfect	No	Yes (Yes only in negative)

stative verbs and stative forms are too great to be coincidental. In the following section, I propose that stative verbs were once former A-verbs whose inflection incorporated a number of features from stative forms, giving rise to a new class that would motivate the emergence of the active/stative semantic alignment of Otomi.

14.8 The emergence of the active/stative alignment in Otomi

In the previous sections, I have shown that Otomi intransitive verbs display both an agent/patient and an active/stative semantic alignment. I have studied the latter in detail, and have compared the O-verbs in this active/stative semantic alignment with a special stative form which transitive verbs have nowadays. In this section, I make a proposal about how the active/stative alignment emerged in Otomi. The dynamics in question are represented in Figure 14.2.

The extreme left of the figure represents a stage at which intransitive verbs already displayed an agent/patient semantic alignment—very much like the one found today, but perhaps having more patient verbs than now. At this stage, a number of 'agent' verbs denoted property concepts. To make the next stage possible, we have to imagine a situation in which intransitive verbs could have stative forms just like the transitive verbs of today—something which is not possible now. In this hypothetical situation, a number of A-verbs that denoted property concepts had such stative forms. These stative forms functioned in most cases as nouns portraying the properties in question, just as stative forms of transitive verbs may now convert into action nouns. The residue of this stage is the handful of stative

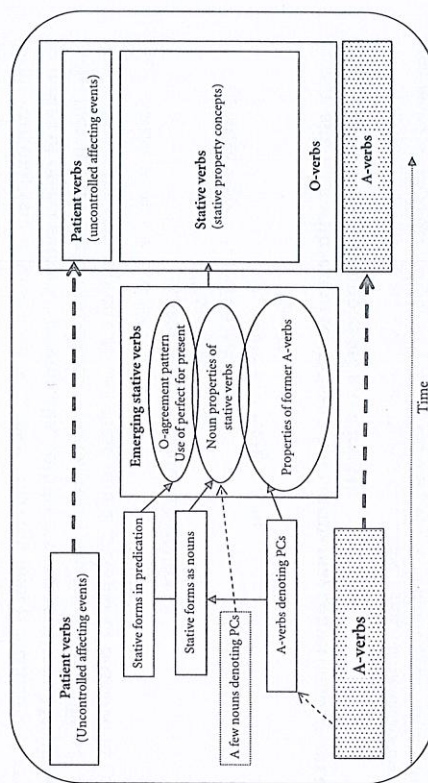


FIGURE 14.2. Development of the semantic alignment of intransitive verbs in (San Ildefonso) Otomi

verbs which by conversion can be used as nouns, as illustrated in (20) above for San Ildefonso Otomi, repeated here as (38):

- (38) *n-t'áxi* 'white (as a colour)'
n-théni 'red'
n-k'áxt'i 'yellow'
n-k'ángi 'blue/green'

At this point in this process, speakers began to use the stative forms of such A-verbs denoting property concepts as an alternative way to predicate the properties in question. Considering the A-verbs in (11) which still denote property concepts, it seems that stative forms were generally used when the property in question was an inherent property or was not typically ascribable to a human (or animate) being. This use must have become very common, to the extent that speakers reanalysed these stative forms as being a part of the inflectional paradigm of A-verbs denoting the same properties.

When the reanalysis was complete, a separate class of stative verbs had emerged in the system, displaying a unified inflectional paradigm, but with rather mixed origins. Stative verbs inherited from stative forms the fundamental feature of encoding their unique semantic argument as object. In this way, with the rise of stative verbs a new semantic alignment emerged in the language for intransitive verbs.

As already mentioned, stative verbs emerged as a rather mixed morphological class from the blending of different paradigms. On the one hand, stative verbs received from A-verbs the features they share now with them, as in (39a). On the other hand, the rest of their inflectional idiosyncrasies stem from their origin as stative forms. The list in (39) shows the inflectional spaces in the paradigm of stative verbs in San Ildefonso Otomi that stem from either former A-verbs or stative forms:

- | | | |
|---------|--|------------------------------|
| (39) a. | From A-verbs | PROCLITICS |
| | Imperfect (negative polarity) | <i>mí</i> |
| | Present (negative polarity) | \emptyset |
| | (Marginally) Irrealis | <i>da</i> |
| b. | From stative forms in predication | <i>xi=N-</i> |
| | Perfect for present (positive polarity) | $\emptyset=N-$ ¹⁸ |
| | Present (with intensifiers) | <i>már=N-</i> |
| c. | From stative forms as nouns in predication | <i>dar=N-</i> |
| | Imperfect (regardless of polarity) | |
| | Irrealis (regardless of polarity) | |

¹⁸ With the intensifiers *řá-* 'very' and *ká=* 'rather', the verbal proclitic used is zero, but the stative prefix *N-* must still be used (e.g. $\emptyset=řá-m-pá-á-a-ř$ [(3.)PRES=INTENS-ST-be.hot-2OBJ-P=2OBJ] 'you're very hot').

All Otomi languages have stative verbs, but they differ greatly as to what the outcome of the diachronic blending is, suggesting that the development in each Figure 14.2 started in the proto-language but was developed separately in each individual sub-group. In the cases where a given inflectional feature emerged from a stative form, we now see the use of the stative prefix *N-* proper of stative forms.

Moreover, the two last properties inherited from stative forms in (39c) reveal that stative forms were used as nouns in predication in the imperfect and the irrealis, regardless of the polarity of the verb. In such cases, stative verbs inherited the special proclitics used in the nominal predication construction. The relevant examples are repeated in (40):

- | | | |
|-----------|---|--|
| (40) | STATIVE VERBS | NOMINAL PREDICATION
(SINGULAR NP) |
| Imperfect | a. <i>már=ř-héts'i=gí</i>
(3.)IMPERF.PC=ST-be.high=1OBJ
'I was tall.' | b. <i>má=ř</i> <i>bé]NP</i>
3.IMPERF.NP=SG thief
'He was a thief.' |
| Irrealis | c. <i>dar=ř-héts'i=gí</i>
(3.)IRR.PC=ST-be.high=1OBJ
'I'd be tall.' | d. <i>dá=ř</i> <i>má'j]NP</i>
3.IRR=SG priest
'He'll become a priest.' |

From being nouns in predication, stative verbs also inherited the lack of the free/bound form inflectional contrast seen in 14.5.3, which other A-verbs have. They extended this feature to their whole paradigm.

14.9 Conclusion

In this chapter, I have proposed for the first time that Otomi has the two types of semantic alignment pattern described so far in the literature: an agent/patient and an active/stative. For this reason, the language is typologically very interesting, as it is not common to find both alignments in the same system. The goal of the chapter was to advance a proposal for the emergence of the active/stative system, which remains a type of alignment which appears to be less frequent cross-linguistically and is less understood in current linguistic theory. I have proposed that the O-verbs in this alignment emerged as reanalysed A-verbs through a process of blending the paradigm of A-verbs with the one of a special inflectional form I called the stative form, which is now mainly used to denote resulting states of transitive verbs and to convert verbs into nouns. I have shown in detail how this process may have come about in the grammar of San Ildefonso Otomi, while also predicting further differences in other Otomi languages. An agent/patient alignment has been found in a handful of Otomanguan languages; the extent to which this could be regarded a feature of the macro-phylum remains to be seen. Outside Otomi, however, nothing is known about the existence of an active/stative

Cutting and breaking verbs in Otomi: An example of lexical specification*

ENRIQUE L. PALANCAR

Abstract

In this article, I show that Otomi (Otopamean; Otomanguean) is a language with a high degree of lexical specification and a lack of generic verbs. This phenomenon is illustrated with the study of verbs denoting cutting and breaking events.

Keywords: cut and break; separation events; Otomi; Otomanguean; verb semantics; generic verbs.

1. Introduction

One organizational principle for the mental lexicon is to have taxonomic relations between words. For example, Lyons (1968) treats words like *duck* and *owl* as hyponyms of the more generic noun, *bird*, which functions as a superordinate category. This organization is not unique to nouns. English *cut* is commonly taken to be a superordinate generic verb with respect to verbs such as *slice*, *chop*, *hack*; etc. Hyponymy has traditionally been taken to be a common organizational pattern in the semantics of many languages, at least for Indo-European. However, more recent work suggests that this may be more true for some languages than others. For example, Majid et al. (this issue a) show that English, but not German, Dutch and Swedish, has generic cutting and breaking (C&B, hereafter) verbs. Other languages such as Tidore (van Staden, this issue) and Tzeltal (Brown, this issue) also seem to lack generic superordinate verbs. Otomi is an interesting language regarding semantic specification because it also (by and large) lacks generic verbs. In particular, I show that this language does not have verbs serving as a superordinate category for cutting and breaking events.

Otomi is a Mesoamerican language spoken in Central Mexico. The language belongs to the Otomamean family; the northernmost branch of Otomanguan. Most of the data in this article come from the variant of San Ildefonso Tultepec, in the state of Querétaro. But regardless of dialectal diversification, the phenomenon I study is a dia-systemic feature. The data I present here come from extended elicitation based on the videoclips by Bohmeyer et al. (2001), which were presented to two fluent native speakers of the language following the procedure as outlined in Majid et al. (this issue b). Additionally, I consulted natural text and dictionary sources (Wallis 1956; Hekking and Andrés de Jesús 1989; Hernández Cruz et al. 2004).¹

2. Cutting in Otomi

Cutting verbs in Otomi denote events involving material destruction brought about by sharp implements. Given that cutting events do not happen spontaneously, all cutting verbs in the language are transitive. Cutting verbs in Otomi are sensitive to either the manner of cutting or the type of patient involved, or both. Instrument is not lexicalized; that is, Otomi verbs do not mean ‘cut with scissors’ or ‘cut with knife’, etc. Nevertheless, in all actions designated by existing cutting verbs, there is an implement associated with the semantic prototype. In principle, speakers do not give information about cutting tools unless the situation has a non-prototypical instrument. The language has a considerable number of highly specific cutting verbs, but the number of lexemes is not close to the impressive lexicons found in other languages, such as for example Khmu and Thai (Premssirat 1987) or Tidore (van Staden, this issue).

It is convenient to start the discussion about cutting verbs with the verb that has the lowest semantic specificity. This verb is *hékí* and it typically depicts a single cutting action performed on the surface of an object with a sharp blade (i.e., knife, scissors, razor; etc). Given its low semantic specificity, the verb can be used in a variety of contexts. Consider the examples in (1):²

- (1) a. *bi=h<y>ǵ^hk-Ø-a=r* *mǔ.*
 3.PST=<NPS>cut-3OBJ-B=SG squash
 ‘He cut the squash.’
- b. *bi=h<y>ǵ^hk-Ø-a=r* *dú^htu* *kó=r* *t’ǒxte.*
 3.PST=<NPS>cut-3OBJ-B=SG cloth with=SG scissors
 ‘He cut the cloth with the scissors.’

Example (1a) depicts a situation where an agent slides the sharp blade of a knife into the surface of a squash (with unspecified pressure) leading to

one of two possible outcomes: (i) the squash is left with just an incision on its surface (as in clip 14: make single incision in melon with knife)³; or (ii) part of the squash is separated from the whole. Example (1b) shows that the verb is also used in the cutting of cloth as effected (typically) with scissors.

Furthermore, the verb is used for accidental cutting situations where the agent is already involved in a cutting action (i.e., as in clip 18: cut finger accidentally while cutting orange). This may be seen in (2a) below.⁴ Alternatively, the verb is also used with instrument subjects (medio-passive) to indicate that a given tool is good for cutting because its edge is sharp, as in (2b):

- (2) a. *dá=hě^hk-Ø-a=ma* 'yé.
 1.PST=cut-3OBJ-B=1POSS hand
 'I cut myself on the hand (while I was cutting something else).'
 (lit. 'I cut my hand.')
- b. *ma jwai Ø=hě^hk-Ø-i*
 1POSS knife 3.PRS=cut-3OBJ-F
xi=ñ-hó-Ø.
 (3.)PRF=ST-be.good-3OBJ
 'My knife cuts (it) well.'

Despite being a transitive clause, example (2a) does not render volition on the part of the agent; Otomi pragmatics establishes that people do not commonly inflict wounds on themselves on purpose.

I have used the English verb *cut* in the translations of *hěki* in examples (1) and (2). However, this translation should not be taken to mean that the Otomi verb has the same generic semantics as English *cut*, only that the semantics of the verb has very low specificity. The verb *hěki* cannot be used in other cutting contexts where the use of a more specific cutting verb is more appropriate.⁵ In other words, other cutting verbs in Otomi are *not* hyponyms of *hěki*; they all function at a similar semantic level. In Table 1, I give a list of other common cutting verbs in the language, (verbs marked with "M" come from Mezquital Otomi, a neighboring dialect of San Ildefonso Tultepec Otomi).⁶

All the verbs in Table 1 render very specific cutting actions for which *hěki* 'cut' cannot be used as a generic substitute. For brevity's sake, I will only comment here on the first four verbs. The verb *tséki* 'hack' depicts a cutting action performed with heavy blows on wood by using an axe or a machete. Deviations from this prototype exist as the verb may be used in other situations. For example, the verb can be used when the material is not hard, as with a squash, a melon or a carrot (e.g., clip 51: split melon in two with single knife blow, followed by pushing halves

Table 1. *Other verbs of cutting*

		Prototypical instrument
<i>tséki</i>	‘hack’	axe
<i>tsét’i</i>	‘fell a tree by hacking’	axe
<i>tsěni</i>	‘slice by hacking’	axe
<i>hěni</i>	‘slice, saw’	knife; machete; saw
<i>hěhmi</i> (M)	‘cut nopal or meat downwards’	knife, machete
<i>hě’mi</i> (M)	‘cut paper’	scissors
<i>hěts’i</i>	‘trim (hair, threads, etc)’	scissors
<i>hét’i</i>	‘cut furrow’	plow
<i>hěts’i</i>	‘jag’	knife; spike; chisel
<i>kýnt’i</i>	‘shred’	shredder
<i>fěts’i</i>	‘cut weed’	sickle
<i>pěhi</i>	‘cut <i>quelites</i> ⁷ and all types of pods or fruits that grow on creepers’	sickle, machete
<i>’óki</i>	‘cut nopal or nopal fruits’	knife; machete
<i>’ótsi</i>	‘cut wood-surface even’	plane
<i>’óxt’i</i>	‘cut beard, etc’ (possessor in dative)	scissors
<i>’óxi</i>	‘cut someone’s hair’ (possessor is object)	scissors
<i>xé’ts’i</i>	‘chop a herb into a broth’	knife
<i>xófo</i>	‘sow’	sickle; machete

apart by hand, clip 37: cut carrot in half lengthwise with single blow of axe), and also with a taut rope across a table (clip 13: saw stick propped between two tables in half). The verb may also be used when an atypical instrument is used: a knife (clip 4: chop cloth stretched between two tables with repeated intense knife blows) or a spike or chisel (as in clip 43: cut carrot crossways into two pieces with single blow of chisel and clip 53: break stick in two with single downward blow of chisel). All deviations must, however, preserve two essential features to qualify as an act of *tséki*: first, the action is performed with heavy blows using a sharp implement; and second, the agent does not act with precise control. The semantics of *tséki* reminds one of an instance of smashing destruction but performed with a sharp instrument.

The verbs *tsét’i* and *tsěni* in Table 2 are historically related to *tséki* ‘hack’.⁸ The verb *tsét’i* ‘fell a tree by hacking’ has a very specific object, while *tsěni* ‘slice by hacking’ implies cutting by heavy blows just like *tséki*, but with a more controlled result to obtain functional sections. Clip 54, cut carrot in half crosswise with single blow of axe, serves as a typical example.⁹ The verb is used when referring to the slicing of sugar cane or the chopping of a chicken’s leg or breast. The verb can also be translated as ‘chop’ or ‘mince’ when the action is performed on food or meat.

Finally, the verb *hěni* expresses separation carefully performed on an object several times so as to produce several pieces. An example is given in (3):

- (3) *bi=h<y>ěng-Ø-a=r* 'óni.
 3.PST=<NPS>slice.B-3OBJ-B=SG chicken
 'He sliced the chicken'.

Speakers may also use *hěni* to express separation performed in the geometrical middle of an object in order to produce two pieces. The phrase *madé* 'middle' is often used as a resultative in order to reinforce such a reading. This is shown in (4):

- (4) *bi=h<y>ěng-Ø-a=r* 'óni *madé*.
 3.PST=<NPS>slice.B-3OBJ-B=SG chicken middle
 'He sliced the chicken in two.'

The verb *hěni* is also used when cutting wood in a careful fashion so as to obtain a desired result, as in clip 15, saw stick propped between two tables in half. For such situations, the verb translates in English as 'saw'. The wood cutting scenario is prominent in its semantics given that the derived noun *nthěni* is used for a 'saw'; the typical instrument used for this purpose.

3. Breaking in Otomi

Otomi also has a well-defined group of breaking verbs. These verbs denote separation involving the (total) destruction of an object, achieved without the use of a bladed implement. As breaking events may happen either spontaneously or be caused by the agency of a (volitional) person, the language has both transitive and intransitive breaking verbs. Consider verbs of breaking presented in Table 2:

Table 2. *Verbs of breaking*

	transitive	intransitive	meaning
a.	<i>xéki/xíki</i> (M)	<i>xégi</i>	'tear (into pieces)'
b.	<i>wáki</i>	<i>wági</i>	'snap in two'
c.	<i>tǔni</i>	<i>tǔgi</i> (M)	'shred into pieces' (i.e., a dress, a book, etc)
d.	<i>pǔni</i>	—	'squash, crush into pieces' (i.e., a vase, a fruit, etc)
e.	<i>tehmi</i>		'smash into peaces' (i.e., pottery, etc)

The verbs in (a) in Table 2 depict the material destruction of things made of flexible material, typically fabric (i.e., pieces of clothing, napkins, etc),

wicker (i.e., baskets, etc.), but also individual pieces of such material, i.e., ropes or cloth. One speaker (Speaker 2, Table 3) was very sensitive to this semantic component, as she used the verb *xɛki* ‘tear’ any time a piece of fabric was affected in the clips, regardless of the manner of separation involved. Similarly, the language has a number of verbs that are sensitive to this semantic component: *kóki* ‘get flexible material down’; *kó’mi* ‘get flexible material down by stretching’; *kóhō* ‘carry flexible material on hands’; *kóts’i* ‘for flexible material to be located on top of something’; *kót’i* ‘put flexible material inside’; etc. All of these verbs involve the same petrified root. The verbs in (a) are typically used to render a type of breaking action achieved by the hand, but the transitive verb is also used for breaking scenes effected with tools acting on taut cloth or ropes (i.e., clip 45: poke hole in cloth stretched between two tables with a twig and clip 50: chop rope stretched between two tables in two with repeated blows of hammer). Even when such instruments are used, the object acted on frays. But if one uses a sharp instrument, the event is no longer framed as a breaking event.

The intransitive verb in (a) is commonly used with the clitic =*sɛ* ‘on its own’ to render situations where the speaker finds the soft material broken without witnessing the breaking. An example is given in (5):

- (5) *bi=xɛ=sɛ=r* *nthót’i.*
 3.PST=break.of.fabric.s=on.its.own=SG rope
 ‘The rope was (already) torn (when I found it).’

The suffix *-sɛ* in (5) serves as an evidential marker indicating that the speaker does not have direct knowledge of what exactly happened to the material, i.e., somebody could have torn the rope or the rope may have broken by itself.¹⁰

The verbs in (b) denote a breaking action affecting a long, rigid object (i.e., a wooden stick; a branch; a pen; etc.). A prototypical instance would be where a person uses both hands to snap a twig into two separate pieces (i.e., clip 19: snap twig with two hands). Deviations from this prototype do occur. A full detachment of the pieces is not necessary as evidenced by the fact that this verb is used for clip 25: snap twig with two hands, but it doesn’t come apart, where the twig remains attached by a few shreds. The intransitive verb is used for spontaneous events that happen due to climatic conditions, natural forces or other circumstances, but most commonly speakers use the suffix *-sɛ* ‘on its own’ when they don’t witness such events.

Finally, the verbs in (c) and (d) depict agentive actions that imply the complete destruction of the object. The last breaking verb in Table 3 is *tehmi*, which is a labile verb. The verb typically denotes a destructive sep-

Table 3. Extensional space of C&B verbs mapped onto the videoclips¹¹

Clip number	Clip description	Speaker 1	Speaker II	
27	cut hair (scissors)	' <i>õxi</i>		
56	cut cloth (scissors)	<i>hěki</i>		
24	cut rope (scissors)			
18	cut finger			
14	cut melon			
20	cut twig			
9	slice carrot			
26	cut carrot			
12	cut cloth			<i>xěki</i>
51	chop melon			<i>tsěki</i>
10	slice carrot		<i>hěni</i>	
28	cut fish			
15	saw stick			
8	<i>cut cloth by itself</i>	<i>xěgi</i>		
46	<i>cut rope by itself</i>			
49	cut rope	<i>tsěki</i>		
1	tear cloth			
36	tear cloth			
45	poke hole in cloth			
35	break yarn		<i>xěki</i>	
38	break yarn			
50	chop rope			
4	chop cloth			
34	chop cloth			
61	karate-chop rope			
43	chop carrot			
53	chop stick			
48	chop branch			
3	hack branch			
32	karate-chop carrot			
13	chop rope			
6	chop carrot (in fury)		<i>tsěni</i>	
37	chop carrot (lengthwise)		<i>tsěki</i>	
54	chop carrot (in half)	<i>tsěni</i>		
39	smash pot	<i>tehmi</i>		
40	smash plate			
31	smash stick	<i>pūini</i>		
21	smash carrot			
23	chop cloth		<i>xěki</i>	
16	<i>snap branch by itself</i>	<i>'wági</i>		
17	<i>snap carrot by itself</i>			
25	snap twig partial	<i>'wáki</i>		
19	snap twig			
5	snap stick			
57	snap carrot			
42	karate-chop stick			
2	chop rope			<i>xěki</i>

aration affecting surfaces or artifacts made of clay or glass, or any other hard breakable material (i.e., vase, plate, glass, CD, pen, etc). Clip 39 and clip 40 in Bohnemeyer et al. (2001) depict typical scenes as the clips portray people smashing pieces of pottery.¹² The meaning of the verb may be loosely translated with English ‘break’, and I suspect that both verbs share a common core, but the Otomi verb lacks the generic meaning that English *break* has. For example, *break* can be used for all the snapping clips in Bohnemeyer et al. (2001), while *tehmi* cannot. In other words, the verbs in (a–d) in Table 3 are not kinds-of *tehmi*.

4. Conclusion

Otomi is a language with verbs which have very specific semantics, but lacks verbs that are generic superordinates. I have used C&B verbs to illustrate this claim, and have proposed that a speaker of Otomi is required to use a specific C&B verb for a given situation if a suitable lexical option exists. Other semantic fields where a similar or greater degree of semantic specificity is observed are: having, grinding, falling and dropping, existence, location, extracting, tying, seeing, peeling, opening, etc. The claim may be further supported with a closer look at verbs from the semantic field of having, as in Table 4:

All the verbs in Table 4 are used to predicate possession. As indicated, the verbs have different meanings in other realms. The verb *tú* is used for possession of body parts; the verb *’ñéhē* is used for animates, whether inalienable (i.e., kinship terms) or alienable (i.e., animals). The verbs *hó*, *thē* and *tsíhi* also show specific semantics. Finally, the verb *pě’ts’i* is used with a wider range of objects than the rest of the class, and

Table 4. *Verbs of having*

Verb	Possession meaning	Other meaning
<i>tú</i>	‘have a body part’ (of humans, animals, and plants)	‘be attached’
<i>’ñéhē</i>	‘have an animate’ (i.e., dog, horse, mother, children, etc)	(this verb is an obscure derivative of <i>’éhē</i> ‘come’)
<i>hó</i>	‘have money on self’	‘carry’ (less specific)
<i>thē</i>	‘have (inherited) land’ (i.e., cornfield, meadow, etc)	‘carry with hands extended (i.e., tray)’
<i>tsíhi</i>	‘have a vehicle’ (i.e., car, tractor, etc) ¹³	‘drive (a vehicle)’
<i>pě’ts’i</i>	‘have alienable purchased property’ (i.e., house, money, tools, clothes, etc)	‘keep, watch’

is mainly used to predicate the possession of anything inanimate that one has purchased.¹⁴ The verb *pé'ts'i* in Table 4 should be taken to be a functional equivalent of the pseudogeneric cutting verb *héki* in Section 2, but within the possession realm. Verbs like these have less semantic specificity than their counterparts, and so are found in a wider variety of contexts, thus making them appear as if they were generic. Overall, then, Otomi is a language that has verbs with a high degree of semantic specificity.

Received 12 October 2004

Universidad Autónoma de

Revision received 5 December 2006

Querétaro, Mexico

Notes

- * This article was financed by the CONACyT (project No 47475). I want to thank Asifa Majid for her editorial help in improving the coherence of this paper to fit the present issue and for her help in the overall organization of the extensional semantic space in Table 4. I am also grateful to the anonymous referees for their comments. My deep gratitude to Mrs. Anastacia Cruz Vázquez for sharing her ancient knowledge with me. I may be contacted at <epalancar@hotmail.com>.
1. I conducted a first elicitation in early 2002, but the exploration had some errors due to a number of misunderstandings on my part. To mend such errors, a second elicitation was carried out in 2005 with the same speakers. The present paper is mainly based on the data obtained in the second elicitation.
 2. In the transcription of Otomi, the deviations from the IPA are the following: <ŋ> /ŋ/; <ñ> /ɲ/; <x> /ʃ/; <f> /p^h/; <j> /k^h/; <y> /j/; <r> /r/; <''> nasality; <a> /a/; <á> /ɔ:/; <é> /ɛ:/; <í> /i:/; and <ó> /ɔ/. As for tone, <́> represents high tone and <̀> ascending tone; low tone is not represented. Abbreviations: (-) morpheme boundary; (< >) infix or consonant mutation; (.) cumulative morpheme or absence of morpheme boundary in gloss; (=) clitic; B—bound form; DAT—dative; F—free form; intr.—intransitive; M—Mezquital Otomi; NPS—non present stem; OBJ—object; PRF—perfect; s—suffixal form; ST—stative; POSS—Possessive; PRS—present (habitual); PST—past (completive); SG—singular.
 3. I give the short descriptions of the video clips by Bohnemeyer et al. (2001) as they appear in Majid et al. (this issue b).
 4. Speakers use the intransitive middle verb *ñhēni* for accidental cutting situations where a person is not already involved in a cutting action, as when one cuts one's hand on a sheet of paper. This verb derives from *hēni* 'slice/saw'. The verb *héki* lacks a middle form (see Palancar 2004 for more details).
 5. This implies that the speaker has to know of other lexical options, both available and suitable for a given cutting situation.
 6. A number of verbs are used for separations involving plants, i.e., *túki* 'separate fruits from a plant'; *tóki* 'separate flowers/leaves from a plant' and *thoki* 'separate corncobs from plant'. Such separations are commonly performed by hand, but a penknife can also be used.
 7. "Quelite" is a generic name for a variety of edible, wild plants such as the *Amaranthus hybridus* and the *Chenopodium album*.

8. Most verbal stems in Otomi consist of a morphological root and a stem formative. The verbs *tsé+k-i*, *tsé+t'-i* and *tsé+n-i* all share the same morphological root *tsé*, while *+k*, *+t'* and *+n* are stem formatives. Stem formatives are semantically opaque, but reflect lexicalized spatial relations (see Voigtlander 1998). The verbs reflect the pattern found in some English verbs borrowed from French, i.e., *pre+fer*, *in+fer*, *di+fer*, etc. Roots in such stems are equally obscure.
9. Despite the rage with which the separation was performed, which rendered no apparent functional sections, the verb was also used by one speaker for clip 6: chop multiple carrots crossways with big knife, intensity.
10. Speakers would use the transitive verb in the impersonal voice if they suspected that an unidentified person had torn the rope.
11. In Table 4, I give a representation of the extensional semantic space onto which representative C&B verbs can be mapped. I have organized the space with reference to the videoclips in Bohnemeyer et al. (2001), rendered with a brief description.
12. The verb is also used to express that a machine is out of order.
13. The verb depicts the possessor as having a vehicle at his disposal.
14. Speakers may also use *pě'ts'i* to predicate the possession of money, land and vehicles despite the existence of *hō*, *thē* and *tshī*; but when *pě'ts'i* is used, speakers emphasize that the type of possession involved is legal and not due to change over time.

References

- Bernal Pérez, Felipino
 2001 *Diccionario Hñähñu-Español Español-Hñähñu del Valle de Mezquital Hidalgo*. Cardonal/Hidalgo/Mexico: Hmunts'a Hēmi.
- Bohnmeyer, Jürgen, Melissa Bowerman, and Penelope Brown
 2001 Cut and break clips. In Levinson, Stephen C., and N. J. Enfield (eds.), *Field Manual 2001, Language and Cognition Group, Max Planck Institute for Psycholinguistics*. Nijmegen: MPI, 90–96.
- Brown, Penelope
 this issue 'She had just cut/broken off her head': Cutting and breaking verbs in Tzeltal. *Cognitive Linguistics* 18(2), 273–285.
- Hekking, Ewald, and Severiano Andrés de Jesús
 1989 *Diccionario Español-Otomí de Santiago Mexquititlán*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- Hernández Cruz, Luis, Moisés V. Torquemada, and Donaldo Sinclair Crawford (eds.)
 2004 *Diccionario del Hñähñu (Otomí) del Valle de Mezquital, Hidalgo*. Mexico, D.F.: Instituto Lingüístico de Verano.
- Lyons, John
 1968 *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Majid, Asifa, Melissa Bowerman, Miriam van Staden, and James S. Boster
 this issue a The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2), 133–152.
- Majid, Asifa, Marianne Gullberg, Miriam van Staden and Melissa Bowerman
 this issue b How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cognitive Linguistics* 18(2), 179–194.

- Palancar, Enrique L.
2004 Middle voice in Otomi. *International Journal of American Linguistics* 70 (1), 52–85.
- Premrsirat, Suwilai
1987 A Study of Thai and Khmu cutting words. *Papers in South-East Asian Linguistics No 10*. Canberra: Pacific Linguistics, A-75, 145–187.
- van Staden, Miriam
this issue 'Please open the fish': Verbs of separation in Tidore, a Papuan language of Eastern Indonesia. *Cognitive Linguistics* 18(2), 297–306.
- Voigtlander, Katherine
1998 La semántica de los sufijos formativos del tema verbal en yuhu (otomí de la Sierra). Paper presented at the II Coloquio Estatal sobre Otopames, Hidalgo.
- Wallis, Ethel Emilia
1956 *Diccionario Castellano-Otomi, Otomi-Castellano*. (Cuadernos del Valle de Mezquital, vol. 1.1.) Itzmiquilpan: Patrimonio Indígena del Valle de Mezquital and Instituto Lingüístico de Verano.

PROPERTY CONCEPTS IN OTOMI:
A LANGUAGE WITH NO ADJECTIVES¹

ENRIQUE L. PALANCAR

UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

In this article, I study the way property concepts such as ‘big’, ‘red’, ‘high’, etc., are lexically treated in Otomi, an Otomanguean language from Central Mexico. I show in detail that such concepts are not encoded as adjectives but rather as verbs and nouns. Lexemes denoting such concepts also occur in an interesting type of nominal compound. The present analysis is based on the dialect of San Ildefonso Tultepec Otomi, but the claims made can be extended to other Otomi variants.

[KEYWORDS: Adjectives, property concepts, stative verbs, Otomi, Otomanguean]

1. Introduction to Otomi and property concepts. Otomi is an Otomanguean, Otomanguean language spoken in Central Mexico. Given its great dialectal diversification, the language can be considered a complex linguistic dia-system consisting of a variety of “Otomi languages”—to use the term coined by Suárez (1983). The phenomenon I study here is a feature of Otomi because it is found in all varieties despite their dialectal differences. Most of the data in this article come from the dialect spoken in the village of San Ildefonso Tultepec, in the state of Querétaro, Mexico. The data from this dialect were collected from texts and field notes from my work on the dialect. This dialect is referred to as SI Otomi. I also present data from other sources, and mention those accordingly.

¹ The present study was financed by the Universidad Autónoma de Querétaro and by the CONACYT (project no. 47475). The analysis is the result of the superb Mexican workshop on property concepts led by Paulette Levy, in which other Mexicoamericanists actively participated, such as David Beck, Alejandra Capistrán, Valentín Peralta, Thomas Smith, Ausencia López, Rosa María Rojas, Verónica Vázquez, and Roberto Zavala. I want to thank all these colleagues, especially Paulette Levy, for their immense assistance and patience in helping me to understand some of the complexities of Otomi grammar. This article would never have been written without their many excellent comments on earlier drafts, including the paper I presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas, held in Boston in January 2004. I also want to thank my referees—one of whom was Scott Berthiaume—for all the invaluable comments they made to improve this paper. I am also grateful to Keren Rice for all the help provided with the final editing, and to the anonymous associate editor for his/her expeditious work. As always, my most sincere thanks to Mrs. Anastacia Cruz Vázquez for sharing her ancient wisdom with me. All errors and deficiencies remain my own responsibility.

Property Concepts (henceforth PC) are also known as “adjectival” concepts because they are commonly expressed by adjectives across languages. In a seminal article, Dixon (1982) proposes various semantic categories for PCs in a wide variety of languages. Such categories include properties regarding Dimension (i.e., ‘big’, ‘thin’, ‘long’, etc.); Age (i.e., ‘new’, ‘old’, etc.); Evaluation or Value (i.e., ‘good’, ‘bad’, ‘fair’, etc.); Color (i.e., ‘white’, ‘black’, etc.); Physical Properties (i.e., ‘soft’, ‘hard’, ‘heavy’, etc.); and properties associated with Human Propensity or character, such as ‘ugly’, ‘jealous’, ‘envious’, etc. Dixon proposes that languages vary with respect to the number of members they have in the adjectival class. The study covers a spectrum from Indo-European-like languages, which have an open and productive class, to languages like Igbo (Kwa; Niger-Congo) or Hausa (Chadic), with only eight and twelve core members, respectively.

More recently, Dixon (2004) has proposed that adjectives are a part of speech existing in all languages. Dixon readily warns that the criteria available to distinguish adjectives from nouns and verbs may be very subtle at times. In the core of his proposal, he predicts that lexemes encoding PCs in a given language are treated in a special fashion within the linguistic system. Such a special treatment could be realized by means of a distinct part of speech (i.e., by adjectives in the traditional sense) or by a given subclass of verbs or nouns (see Chafe 2004 for an apparent counterexample to this generalization in Iroquoian languages).

Dixon (2004) takes this typological insight to its furthest consequences, and thus calls “adjectives” any well-defined class containing PCs in a given language, regardless of whether the lexemes in question are treated as adjectives, nouns, or verbs by a more traditional approach. For example, a large number of lexical items expressing PCs in Korean have been regarded as verbs by Martin (1992), attending to syntactic criteria. More recently, Sohn (2004) presents enough evidence to prove that the lexemes in question form a distinct subclass of verbs if one applies morphological criteria, but prefers to call such a subclass “adjectives” instead of verbs, following Dixon (2004). In my opinion, this is a problematic stance because it implies granting to morphology the same definitional power as syntax in the characterization of a part of speech.

In contrast, Beck (1999; 2002) maintains that syntactic criteria are crucial to the definition of a part of speech and proposes a characterization of adjectives by using a fully revised and adapted revision of the markedness theory in Hengeveld (1992). In this article, I follow Beck’s proposal (also applied to Totonac in Beck 2000) because I think it is more convenient for the purposes of linguistic description. In this light, the word “adjective” is used here only in its traditional sense, that is, as an independent part of speech fully distinct from nouns and verbs.

1.1. Property concepts in Otomi. In this article, I study the lexical treatment of PCs in Otomi in a fairly extensive way and make the claim that Otomi is a language with no adjectives. Lexical items denoting PCs in Otomi can be divided into two main lexical groups: Group 1 and Group 2. All members of Group 1 are verbs, while all members of Group 2 are nouns. For this reason, the PC lexemes in the groups could certainly be called “PC verbs” and “PC nouns.”

Group 1—the PC verbs—form a large class that encodes prototypical PCs. A number of examples are given in (1):²

- (1) GROUP 1 PC LEXEMES: *dótá* ‘be big’, *mě* ‘be hard’, *tx’úlo* ‘be small’, *hěts’i* ‘be high/tall’, *pothi* ‘be black’, *t’áxi* ‘be white’, *xidi* ‘be wide’, *pǐdi* ‘be thick’, *mǎ* ‘be long’, *noho* ‘be fat’, *ódi* ‘be rough’, *hú* ‘be heavy’, etc.

In contrast, Group 2—the PC nouns—form a much smaller class with just a few members, and they denote properties ascribable to human beings. Some examples are given in (2):

- (2) GROUP 2 PC LEXEMES: *měfo* ‘a short person’, *bó^hka* ‘a cruel person’, *zone* ‘an evil or cruel person’, *ndúxte* ‘a naughty person’, *bíbo* ‘a clever person’, etc.

Additionally, speakers of Otomi commonly employ PC verbs and PC nouns in nominal compounds. Two examples of such compounds are presented in (3), where the PC verb *dótá* ‘be big’ and the PC noun *měfo* ‘a short person’ function as dependent members:

- (3a) *dótá-m’asmé*
be.big-napkin
‘tablecloth’

² In the transcription of Otomi, I have used the orthography suggested by Hekking (2002), with only two innovations: *ř* for a trill /r/ and ^h for a salient phonetic aspiration. Deviations from the IPA convention are: ’ /l/; f /p^h/; j /k^h/; y /j/; x /ʃ/; tx /tʃ/; ñ /ɲ/; r /r/; ř /r/; a /a/; ǫ /ɔ/; ǧ /ʒ/; ǫ /ɔ/; ǔ /i/. Umlaut indicates nasalization (i.e., *ä /ā/, ü /ū/,* etc.). Otomi has three tones: ascending *ǎ*; high *ǎ*; and low, which is not marked. Tone does not play a fundamental role in the organization of the lexicon or the morphology. In general, (monosyllabic) morphological roots receive an inherent tone, but this is commonly conditioned by prosody. Prosodic parameters also condition the tone of suffixes and other morphemes. Tone does not differ substantially across dialects. Given that tone is still poorly understood in SI Otomi, I have used the excellent transcriptions in Wallis (1956) and Hernández Cruz et al. (2004) from Mezquital Otomi as a way to confirm the correct forms for SI Otomi, which is its closest variant.

- (3b) *měfo-ndǒ*
 short.person-man
 ‘short man’

While many authors acknowledge the verbal character of lexemes in Group 1, these lexemes nevertheless have been labeled “adjective” (cf. Ecker 1952, Hess 1968, Voigtlander and Echevoyen 1985, Hekking and Andrés de Jesús 1984, Lastra 1992; 1997, Andrews 1993, Hekking 1995, and Bartholomew 2004). These authors motivate the treatment of these forms as adjectives on semantic grounds: the lexemes in question express concepts similar to the ones rendered by adjectives in Spanish or in English. I propose a novel position, that, despite their semantics, the PC lexemes in SI Otomi are verbs (Group 1) and nouns (Group 2).

In the next section, using markedness theory from Beck (1999; 2002), I show that lexemes belonging to Group 1 are indeed verbs. In 3, I also show that such verbs form a well-defined subclass devoted solely to the expression of PCs (i.e., the alleged “adjectival class” according to Dixon 2004). This subclass involves two further subclasses. Having studied PC verbs in detail, I briefly consider the lexemes of Group 2 in 4 and show that they are nouns. Both PC verbs and PC nouns show interesting inflectional overlaps, which are presented in 5. Finally, in 6, I describe in detail the type of nominal compounds illustrated in (3) above. I conclude with a short prospective summary of the proposal in 7.

2. “Without further measures”: A key concept to the lexical encoding of PCs. Beck (1999; 2002) presents criterial definitions of the major parts of speech (i.e., noun, verb, and adjective) based on a theory of syntactic markedness.³ Lexical classes are characterized in terms of both their semantic prototypes and their unmarked syntactic roles. In this theory, lexical elements may also appear in other extended syntactic roles beyond their prototype (i.e., verbs as heads of NPs, verbs as modifiers, nouns as syntactic predicates, etc.). As Beck (2002:25) points out, lexical items require “additional grammatical machinery” to appear in such extended cases, and for this reason they could be said to be “marked” cases. Hengeveld (1992) refers to this “additional grammatical machinery” as “further measures” (FM).

Borrowing Hengeveld’s concept, and modifying it slightly, Beck defines FM as “the morphological, syntactic, or semantic properties acquired by an element in a non-prototypical syntactic role” (1999:20). In other words, markedness theory claims that a given lexical element will act as an un-

³ Beck (1999; 2002) advances a sophisticated theory of markedness aimed at a typology of adjectival encoding. In his typology, Otomi is aligned with Cora (Uto-Aztecan).

marked element in its most natural or prototypical context but will require FM to function elsewhere. If “marked” is equated with FM, “unmarked” is synonymous to what can be called “without further measures” (WFM).

In this light, an adjective is defined as follows:

ADJECTIVE: A lexical item expressing a semantic predicate that can be WFM a syntactic dependent of a lexical item expressing its semantic argument (Beck 2002:84).

In other words, AN ADJECTIVE IS AN UNMARKED MODIFIER, an element that may function as a modifier WFM. To illustrate this, consider the examples in (4), from Beck (2002:25) (taken from Hengeveld 1992:58):

- (4a) *the **intelligent** detective*
 (4b) *the **singing** detective*
 (4c) *the detective **who is singing***

All the elements in boldface in (4) modify the noun *detective*, which expresses a semantic argument. The element *intelligent* in (4a), a semantic predicate depicting a PC, is the only element in (4) which functions as a modifier WFM. In contrast to adjectives such as *intelligent*, verbs in English cannot function as modifiers WFM. This is illustrated in (4b) and (4c) because the verb *sing* uses FM to function as a dependent in this noun phrase structure, i.e., in (4b) by means of suffix *-ing* (a morphological measure) and in (4c) by means of a relative clause (a syntactic measure).

The examples in (4) show the contrast involving the markedness of verbs in an attributive context with respect to adjectives. Relevant contrasts between adjectives and verbs may also occur in syntactic predication. Following Beck, a verb is defined as:

VERB: A lexical item expressing a semantic predicate which can WFM be syntactic head of a lexical item expressing its semantic argument (Beck 2002:78).

In this way, consider the following examples:

- (5a) *the detective **sings***
 (5b) *the detective **is intelligent***

The elements in (5) are both syntactic and semantic predicates. However, in (5a), the verb *sing* functions as a predicate WFM (the *-s* being inflection). This is the prototypical function of a verb. In contrast, *intelligent* in (5b) cannot function predicatively WFM; it requires a copula to do so (a syntactic measure).

Using contrastive structures similar to (4) and (5) could prove useful in solving the dilemma of whether a given PC element is an adjective or a verb.

Of course, the linguist needs to establish what criteria count as WFM for a particular language. Once the class to which a given lexical element is established by attending to syntactic criteria, it is then necessary to establish whether there are subclasses. This is achieved mainly by studying morphological deviations.

As for Otomi, except for a number of nouns, the vast majority of PCs are treated lexically as verbs; the crucial point is that there are no adjectives in Otomi. I treat these lexical items as verbs because, like verbs, they cannot work as syntactic modifiers WFM and because, like verbs, they serve as syntactic predicates WFM. I present evidence for this in the following two sections. Later, I introduce a detailed description of the classes found using morphological criteria.

2.1. Group 1 PC lexemes in attribution. Verbs in Otomi modify nouns with FM by means of a relative clause (a syntactic measure). Otomi relative clauses are postnominal and often lack an overt complementizer, as illustrated in (6a). Nevertheless, the definite determiner *nú/nǒ* is used occasionally as a (pronominal) relativizer, as shown in (6b) (parentheses indicate the scope of a syntactic phrase):⁴

- (6a) (*nú = ya nǒbió* \emptyset *mí = 'ñěhě- \emptyset)_{RC}*)_{NP} 'yá
 DEF=PL boyfriend REL 3.IMP=have.animate.F-3OBJ P
'í-mí = já^h-p-a = nkásó
 NEG-3.IMP=do-3DAT-B=notice

'The boyfriends she had took no notice of her'.

- (6b) (*yá da = <z>ǵng-a = (nǒ = r* *n'wó (nǒ*
 P 3.FUT=<NPS>arrive.there.B-B=DEF.SG=SG boa REL.SG
da = xot'- \emptyset -a = nǒ = r *bágá)*_{RC})_{NP}
 3.FUT=coil.around-3OBJ-B=DEF.SG=SG COW

'The boa that coils around the cow would arrive (there)'.

⁴The abbreviations used in this article are: - morpheme boundary; <> infix or consonant mutation; . cumulative morpheme; _ lexicalized morpheme boundary; = clitic; 1, 2, 3 grammatical persons; ANTI antipassive; B bound form; CONT continuative; DAT dative; DEF definite; DEL delimitative; DEM demonstrative; DIM diminutive; DU dual; EX exclusive; F free form; FOC focal; FUT future; FW function word; I, II types attending to spatial/discourse distance; IMP imperfect; IMPER imperative; IN inclusive; IND indefinite; INT intensive; LOC locative; M middle voice; NEG negative; NI nasal intransitive verb; NP nominal predication; NPS nonpresent stem; OBJ object; P particle; PC property concept; PERF perfect; PL plural; POSS possessive; PRES present; PST past; QUOT quotative; REL relative; S suffixal form; SG singular; ST stative.

In (6a), the relative clause *mí 'ñěhě* '(that) she had' modifies the head noun *nóbió* 'boyfriend', restricting its reference. In (6b), the RC *nǒ da xǒt'a nǒr bágá* 'that coils around the cow' shows the use of the definite determiner *nǒ* as a (pronominal) relativizer.

The PC lexemes of Group 1 modify nouns in attributive position by means of relative clauses. This is illustrated by the following:

- (7a) (*da = ' <ñ>ét' -∅ -a (= nǒ = r* *bǒjǒ* (∅)
 3.FUT=<NPS>place.volume-3OBJ-B=DEF.SG=SG iron REL
 ∅ = *dǒtá*)_{RC}NP
 3.PRES=be.big

'They'd set up the big iron (i.e., the crane)' (lit., '... the iron that is big').

- (7b) (*dá = hǒx-' -a (= 'na = r* *t'ó* (*nǒ*)
 1.PST=bring.S-2DAT/OBJ-B=IND.SG=SG stick REL.SG
xi = mě)_{RC}NP
 3.PERF=be.hard

'I brought you a hard stick' (lit., '... a stick that is hard').

Like verbs, lexemes such as *dǒtá* 'be big' in (7a) and *mě* 'be hard' in (7b) require a relative clause to modify a noun. In other words, such lexemes are not adjectives because they are not "unmarked" modifiers, as they cannot function as dependents in an attributive context WFM (i.e., without a syntactic measure such as a relative clause). Other combinations like the ones in (8) are impossible (AP indicates "adjectival phrase"):

- (8a) *('na = r *t'ó* (*mě*)_{AP})NP
 IND.SG=SG stick hard

Intended meaning: 'a hard stick'.

- (8b) *('na = r (*mě*)_{AP} *t'ó*)NP
 IND.SG=SG hard stick

Intended meaning: 'a hard stick'.

2.1.1. Degree of comparison. Like many other Mesoamerican languages, Otomi lacks a genuine comparison construction. The language expresses comparison by means of a relative clause and the adverb *más* (a borrowing from Spanish *más*). An example of a superlative structure in attribution is given in (9), where a focus relative clause is used:

- (9) *bi* = <*d*>*ang-∅-a* = (*nǒ* = *r* *ngǔ* (*gó* *más*
 3.PST=<NPS>buy.B-3OBJ-B=DEF.SG=SG house REL.FOC more
∅ = *dǒtá*_{RC})_{NP}
 3.PRES=be.big

‘He bought the biggest house’ (lit., ‘. . . the house that is the one more big’).

(9) shows that degree of comparison for a given PC lexeme of Group 1 is encoded by means of a relative clause when the lexeme functions attributively. This represents the same type of evidence as in (7) to show that lexemes of Group 1 cannot function as modifiers WFM.

2.2. Group 1 PC lexemes in predication. Otomi verbs inflect for Tense/Aspect/Mood and for subject agreement by means of proclitics, which I treat here as function words (FW).⁵ Two examples are given in (10) with intransitive verbs:

- (10a) *nǒ* = *r* *ndǒ* *mí* = '*bú*^h = '*pú*
 DEF.SG=SG man 3.IMP=live.S=DISTAL.II

‘The man lives over there’.

- (10b) *yá* *xi* = '*ñ*>*é*^h = *nú* *nó* = *r*
 P 3.PERF=<NPS>come.S=DISTAL.I DEF.SG.3POSS=SG
jödó
 brother.of.man

‘His brother is coming over there’.

The examples in (10) show the use of two proclitics or FWs for third person with the verbs '*bú*' 'live' and '*é*' 'come'. The FW *mí* in (10a) encodes imperfect (past), while *xi* in (10b) encodes perfect.⁶ In (11), I give a paradigm of the most common FWs found in SI Otomi. The FWs in (11) are used with both intransitive and transitive verbs:⁷

- | | | | | |
|------|------------------|-------------|-------------|---------------|
| (11) | | 1st | 2nd | 3rd |
| | Present | <i>dí</i> | <i>gí</i> | <i>∅/bí/í</i> |
| | Imperfect (Past) | <i>n-dí</i> | <i>n-gí</i> | <i>mí</i> |

⁵ This happens in SI Otomi, Mexquitilán Otomi, Mezquitil Otomi, and Sierra Otomi. In other dialects, such as in Mexico State Otomi and Ixtenco Otomi, the existing descriptive literature treats them as prefixes.

⁶ The perfect has various uses in SI Otomi; with motion verbs like the one in (10b), the perfect is used to encode a present state of affairs.

⁷ FW paradigms in other dialects can get much more complex as they may include a substantial number of verbal subclasses and other grammatical categories not taken into account here (cf. Sierra Otomi in Voigtlander and Echegoyen 1985).

Future	<i>ga</i>	<i>gi</i>	<i>da</i>
Imperfect Future	<i>n-ga</i>	<i>n-gi</i>	<i>n-da</i>
Past	<i>dá</i>	<i>gá</i>	<i>bi</i>
Perfect	<i>(x)tá</i>	<i>(x)ká</i>	<i>xi</i>
Pluperfect			<i>(x)ki</i>

Like any other verb in the language, the PC lexemes of Group 1 can function as intransitive predicates WFM. Consider the examples in (12):

(12a) *ma ngü mí = dótá*
 1POSS house 3.IMP=be.big
 ‘My house was big’.

(12b) *nú = nú = r zǎ xi = mě*
 DEF=DEM.II.SG=SG wood 3.PERF=be.hard
 ‘That stick is hard’.

The examples in (12) illustrate the predicative use of two PC lexemes in Group 1, *dótá* ‘be big’ and *me* ‘be hard’. Such lexemes do not require additional grammatical machinery to work as verbal predicates; that is, they function as syntactic predicates WFM just like the verbs ‘*bui*’ ‘live’ in (10a) or ‘*éhě*’ ‘come’ in (10b).

3. PC verbs. In the last section, we saw that PC lexemes of Group 1 are not “unmarked” modifiers; that is, they cannot modify a noun WFM. Similarly, such lexemes work as syntactic predicates WFM. These two facts support a full treatment of such lexemes as verbs. However, if one now attends to morphological criteria only, these verbs in SI Otomi form a subclass with respect to the great bulk of intransitive verbs in the language. Given that the most relevant semantic feature of this subclass is the expression of PCs, I have called the verbs of Group 1 “PC verbs.” For convenience, other intransitive verbs that do not belong to this subclass are called “active verbs.”⁸ For example, in contrast to active verbs, which can only use the FW *mí* for the imperfect of third person, PC verbs may inflect by means of the FW *már*. This is shown in the contrast in (13):

⁸ If one were to follow Dixon’s (2004) terminology, PC verbs in Otomi would be the “adjectives” of Otomi. In other words, they would realize a different part of speech although still being very close to verbs. Such a label has been used before under similar circumstances. For example, a number of scholars in Dixon and Aikhenvald (2004) study languages where PCs are also encoded as a subclass of verbs (i.e., Sohn 2004—Korean, Kruspe 2004—Semelay, Enfield 2004—Lao).

- (13a) *ma t'ó már = dótá/mě*
 1POSS stick 3.IMP.PC=be.big/be.hard
 'My stick was big/hard'.

- (13b) *mí = !*már = 'ótho té n-ga = tsi-∅ = hé*
 3.IMP=3.IMP.PC=be.nothing what IMP-1.FUT=ingest-3OBJ=1PL.EX
 'There was nothing we could eat'.

This is a relevant morphological criterion to distinguish a PC verb in SI Otomi.⁹ Quantification is yet another area where PC verbs stand out from active verbs. Active verbs receive the quantifier FW *xa*, as shown in (14), while PC verbs cannot:

- (14a) *xa = mí = níxt'í'ófó*
 INT=3.IMP=run/write.F
 'She used to run/write a lot'.

- (14b) **xa = már = dótá/mě*
 INT=3.IMP.PC=be.big/be.hard
 'Intended meaning: 'It was very big/hard'.

On the other hand, PC verbs receive the prefix *řá-* or the FW *ká* as intensifiers, while active verbs cannot. The contrast is shown in (15) and (16), where (15a) and (16a) have PC verbs, while (15b) and (16b) have active verbs:

- (15a) *∅ = řá-dótá/mě*
 3.PRES=INT.PC-be.big/be.hard
 'It's very big/hard'.

- (15b) **∅ = řá-'ófó*
 3.PRES=INT.PC-write.F
 Intended meaning: 'She writes a lot'.

- (16a) *ká = ∅ = dótá/mě*
 INT.PC=3.PRES=be.big/be.hard
 'It's rather big/hard'.

⁹ Additionally, if compared to active verbs, most PC verbs also show restrictions concerning the number of tenses and aspects in which they inflect. Nonetheless, this restriction is not a defining feature of the class but rather is largely motivated by the stative semantics the verbs convey.

TABLE 1
CLASS 1 PC VERBS

Dimension	<i>dóta</i>	'be big'	
	<i>tx'úlo</i>	'be small'	
	<i>yá'pu</i>	'be far away'	
Property	<i>nzá_Tho</i>	'be beautiful'	
	<i>tse^h_Tho</i>	'be strong'	(Animate)
	<i>tša^h_te</i>	'be prone to biting (for a dog)'	(Animate)
Time/Age	<i>'rá'yo</i>	'be new'	
	<i>txu</i>	'be old (for a woman)'	(Human)
	<i>xíta</i>	'be old (for a man)'	(Human)
	<i>ts'únt'u</i>	'be young (for a man)'	(Human)
	<i>ntxúntsi</i>	'be young (for a woman)'	(Human)

(16b) **ka* = \emptyset = *níxt'í*

INT.PC=3.PRES=run

Intended meaning: 'She runs a lot'.

PC verbs can be further divided into two main subclasses using subtle morphological criteria. Verbs like *dóta* 'be big' belong to a first subclass, which I call "Class 1 PC verbs." This class is very similar to active verbs regarding inflection. On the other hand, verbs like *mě* 'be hard' belong to a second subclass, which abounds in inflectional idiosyncrasies. I call the second subclass "Class 2 PC verbs" or "Stative Verbs." These two subclasses are described in detail below.

3.1. Class 1 PC verbs. This is a closed class with only a few members. The verbs involved are presented in table 1, organized according to the categories advanced by Dixon (1982; 2004). The two verbs rendering the most prototypical instances of dimension (i.e., 'big' and 'small') are found in this class. Seven out of 11 members denote properties that are only applicable to a human or an animate entity. This is a relevant semantic feature of the class. The last three verbs in table 1 are lexemes that can also occur as nouns. In contrast to their use as verbs, the nouns identify the sex of the referent rather than his or her age. Such transcategorial lexemes are common across languages when they reflect properties related to Human Characteristics, e.g., Spanish *viejo* (adj/n) 'old man'; *vieja* (adj/n) 'old woman'; *joven* (adj/n) 'young'; etc. (for a full discussion, see Beck 1999; 2002, who treats them as words with nominal and verbal properties).

All the verbs in this class receive the special intensifier markers presented in (15) and (16) above. However, the verb *tx'úlo* 'be small' uses the diminutive prefix as a (first-order) intensifier, as shown in (17):

- (17) *ma ngǔ ∅ = (řá-)txí-tx'úlo*
 1POSS house 3.PRES=INT.PC-DIM-be.small
 'My house is (very) very small'.

As for inflection, the verbs of this class behave much like other active verbs. As an illustration, compare how the FWs used in the inflection of the verb *dótá* 'be big' are the same as the ones used with an active verb '*ófo* 'write', an intransitive verb. Both verbs have been inflected for all grammatical persons in the present (18a) and in the imperfect (18b):¹⁰

- | | | | |
|-------|-------------------------------|-------------------------------------|----------------------------------|
| (18) | 1st | 2nd | 3rd |
| (18a) | <i>dí = dótá'ófo</i> | <i>gí = dótá'ófo</i> | <i>∅ = dótá'ófo</i> |
| | 1.PRES=be.big/write | 2.PRES=be.big/write | 3.PRES=be.big/write |
| | 'I'm big/I write'. | 'You're big/You write'. | 'He is big/
He writes'. |
| (18b) | <i>n-dí = dótá'ófo</i> | <i>n-gí = dótá'ófo</i> | <i>mí = dótá'ófo</i> |
| | IMP-1.PRES
=be.big/write | IMP-2.PRES
=be.big/write | 3.IMP=be.big/write |
| | 'I was big/I was
writing'. | 'You were big/You
were writing'. | 'He was big/
He was writing'. |

Despite the similarities in (18), not all PC verbs of this class inflect in the tenses available for active verbs. For example, Class 1 PC verbs may receive past tense inflection, but when they do so they express a present state of affairs. This is shown in (19) with the verb *txu* 'be old (for a woman)':

- (19) *nǒ = r b'éhñó bí = n-txu*
 DEF.SG=SG woman 3.PST=NI-be.old(for.a.woman)
 'The woman is old' (lit., 'The woman got old').

In contrast, active verbs that depict states render a resultative state reading when they are inflected in the past but do not have a present reading. This is illustrated in (20):¹¹

- (20) *nú ma nǒno xa = bí = n-jóhyá*
 DEF 1POSS mother INT=3.PST=NI-be.glad
 'My mother got very happy' (resultative state).
 *'My mother is very happy'.

¹⁰ Third person has been translated for convenience as 'he' throughout this paper, although third person in Otomi does not entail gender or number distinctions.

¹¹ The verb *jóhyá* 'be glad' is one of a very few active verbs in Otomi that depict PCs.

However, other PC verbs of this class cannot receive past morphology, as shown by the impossibility of (21a) with the verb *dó'tá* 'be big', as the verb appears to express a stative reading only. The language uses an alternative construction to express state change, which involves the verb *te* 'grow', as shown in (21b):

(21a) **nú ma bötsí bi = dó'tá*
 DEF 1POSS child 3.PST=be.big

Intended meaning: 'My children got big'.

(21b) *nú ma bötsí bi = te Ø = dó'tá*
 DEF 1POSS child 3.PST=grow 3.PRES=be.big

'My children got big' (lit., 'My children grew, they are big').

3.2. Class 2 PC verbs or "stative verbs." This is an open and productive class of PC verbs (i.e., adjectives from Spanish enter the lexicon as members of this class). A nonexhausted list of members of this class in SI Otomi appears in Appendix A. Class 2 PC verbs have a number of morphological idiosyncrasies which make them look different at times from other Class 1 PC verbs or any other verb in general. Since they form a special subclass, I have labeled them "stative verbs."

The idiosyncrasies involve the following features: (a) their argument is encoded with object morphology; (b) they receive a morphologically conditioned nasal prefix; (c) they use a special set of FWs; and (d) they lack a morphosyntactic found form. These morphological idiosyncrasies identify stative verbs as a verbal subclass.¹² I describe each of the characteristics separately below.

3.2.1. The active/stative split. The argument of stative verbs is encoded with the same morphology that encodes objects of transitive verbs. Consider, for example, (22), where it is shown that object anaphora is encoded by means of suffixes:¹³

(22) *yá gá = tǝn-kagí, pé ga = pǝng-i*
 P 2.PST=win.S-1OBJ but 1.FUT=return-F
ga = tsǝx-'-i
 1.FUT=take.animate.S-2OBJ-F

'You beat me, but I'll come back to get you'.

¹² Despite their morphological idiosyncrasies, the reader should bear in mind that stative verbs are verbs that follow the markedness criteria established in 2. In other words, they are syntactic predicates WFM and are not unmarked modifiers of nouns.

¹³ Object markers display allomorphic patterns that are too complex to treat accurately here. See, for example, Voigtlander and Echegoyen (1985) or Lastra (1992; 1997) for more details.

The same morphology is also used to encode the argument of a stative verb, as illustrated in (23):

- (23a) *xi = m-pá-gagí*
 (3.)PERF=ST-be.hot-1OBJ
 ‘I’m hot’.
- (23b) $\emptyset = \check{r}á-ni^{\check{h}}-kagí$
 (3.)PERF=INT-be.fast.walking.s-1OBJ
 ‘I walk steady’.
- (23c) *xi = ñ-hěts’i-’í*
 (3.)PERF=ST-be.high-2OBJ
 ‘You’re tall’.
- (23d) *xi = no-’a’í = húj*
 (3.)PERF=be.fat.s-2OBJ=PL
 ‘You (PL) are tall’.

The pattern shown in (23) reveals that Otomi has an active/stative split involving its intransitive verbs, just like the one proposed for Guaraní by Mithun (1991). As for the active verbs in this split, a number of them display another type of split of the agent/patient type similar to those found in other Otomanguean languages like Amuzgo (Smith-Stark and Tapia García 2002) or Chocho (Veerman-Leichsenring 2000). Active/stative splits are based on an aspect dimension, while agent/patient splits are based on volition, control, etc. (for more details, see Mithun 1991). Consider, for example, (24), an instance of the agent/patient split:

- (24) *yá bi = <z>á^h-kagí*
 P (3.)PST=<NPS>get.tired.s-1OBJ
 ‘I got tired’.

Since FWs in Otomi cannot express Tense/Aspect/Mood separately from grammatical person, I have placed the grammatical third person between angled brackets in (23) and (24). In this way, the FWs used with stative verbs should be interpreted as encoding Tense/Aspect/Mood only.

The split is clearly observed when the verb is inflected in first or second person, as in (23). However, the language has a zero-marking of third-person object, as in (25); this necessarily implies that the split is covert when the verb has a third-person argument, as may be seen in (26):¹⁴

¹⁴I take the structure in (26) to be the underlying encoding pattern for all stative verbs inflected in third person. For this reason, all examples above that involve the stative verb *mě* ‘be hard’ can be understood as having the same underlying structure as (26).

- (25) *ntónsé hím-bi = <z>i-∅*
 then NEG-3.PST=<NPS>ingest-3OBJ
 ‘Then (the man) did not eat it’.

- (26a) *ká = ∅ = n-tsé-∅* *ri ngó*
 INT.PC=(3.)PRES=ST-be.cold=3OBJ 2POSS flesh/skin
 ‘Your skin is rather cold’.

- (26b) *pé hínjí ∅ = kǔhi-∅*,
 but NEG (3.)PRES be.tasteful-3OBJ
ká = ∅ = n-jú-∅
 INT=(3.)PRES-ST-be.bitter-3OBJ
 ‘But it’s not tasteful, it’s quite bitter’.

3.2.2. The stative prefix. A number of examples in the previous section show that stative verbs receive a nasal prefix. I have treated this nasal prefix as a “stative marker.”¹⁵ The prefix is not used when the stem has a nasal consonant in initial position (i.e., *mě* ‘be hard’, *noho* ‘be fat’, etc.).¹⁶ Examples of verbs with the stative prefix are given in (27):

- (27a) *bi = <z>i-∅* *ya t’éi nú*
 3.PST=<NPS>ingest-3OBJ PL wheat REL
xi = n-k’ángi-∅
 (3.)PERF-ST-be.green-3OBJ
 ‘They ate green wheat’.

- (27b) *ga = kún-∅ = hé* *njabú xi = ñ-’óni-∅*
 1.FUT=grind.S-3OBJ=1PL.EX this.way (3.)PERF-ST-be.dry-3OBJ
 ‘We’d grind it (the chili) that way, while being dry’.

- (27c) *’yá ∅ = řá-m-pá-∅*
 P (3.)PRES-INT.PC-ST-be.hot-3OBJ
 ‘It’s very hot’.

The stative prefix has three allomorphs with regular phonological conditioning: [n-] in (27a) is the default realization of the morpheme; [ɲ] (*ñ*) in (27b)

¹⁵ The language has at least seven other nasal prefixes. Palancar (forthcoming) studies three of these prefixes.

¹⁶ For cases like this, Scott Berthiaume has suggested that the prefix is still used with nasal stems but assimilates to the stem.

appears before a glottal ([ʔ, h]); and [m-] occurs before a labial consonant ([p, ph, b]), as in (27c).¹⁷

There are places in the inflectional paradigm of a stative verb where the nasal prefix is not used. For example, the marker is not found when the verb appears in the negative in the present, as in (28a), and in the imperfect with the FW *mí*, as in (28b):

(28a) *nǒ ma ngǔ híngí* $\emptyset = (*n-)$ *t'ǎxi- \emptyset*
 DEF.SG 1POSS house NEG (3.)PRES=(ST-)be.white-3OBJ
 'My house is not white'.

(28b) *hí-mí = (*n-)tsé- \emptyset*
 NEG-(3.)IMP=(ST-)be.cold-3OBJ
 'It wasn't cold'.

However, the prefix is used when the verb receives special PC verb morphology, presented in 3. This is seen, for example, in (29) with the use of the FW *már*, despite the co-occurrence of the negative:

(29) *hí-már = n-hú- \emptyset*
 NEG-(3.)IMP.PC=ST-be.heavy-3OBJ
 'It wasn't heavy'.

3.2.3. Special uses of FWs. Stative verbs use FWs available to other verbs, but they have a number of morphological deviations. To express a state of affairs in the present, stative verbs use perfect morphology, as in (30), rather than a zero FW, which is common with other verbs to express present, as in (31):

(30) *xí = ñ-hěts'i- \emptyset*
 (3.)PERF=ST-be.high-3OBJ
 'S/he is tall'.

(31a) *xa = $\emptyset = n-tsú = 'na$*
 INT=3.PRES=M=be.scared=QUOT
 'They say he's afraid'.

(31b) *ri páhni $\emptyset = 'rá'yo$*
 2POSS upper.garment 3.PRES=be.new
 'Your shirt is new'.

¹⁷ Additionally, when the nasal is attached to a stem with an initial glottal-labial consonant cluster (e.g., [#ʔbX] or [#hmX]), as in *běxii* 'dark' or *hmödi* 'expensive'), the labial is dropped, and the nasal assimilates the labial feature, as [m-ʔX] or [m-hX] (e.g., *m-'exii* or *m-hödi*).

Nevertheless, the use of the perfect to express present is not an anomaly; as the examples in (32) show, motion verbs and inchoative verbs commonly use perfect for this purpose:

(32a) *nǒ = r* 'bɛhñó *xí = <m>á = nú*
 DEF.SG=SG woman 3.PERF=<NPS>go=DISTAL.I

'The woman is now passing by'.

(32b) *nó = r* *thó* *yá* *xí = 'y>ót'-i*
 DEF.SG.3POSS=SG corn P 3.PERF=<NPS>dry-F

'His corn is dried'.

Moreover, a zero FW is still used with stative verbs in other circumstances to express present time. This is seen, for example, in the negative, as in (33), or when the verb receives PC intensifiers, as in (34):

(33) *yá* *híngí* $\emptyset = kǔhi-\emptyset$
 P NEG (3.)PRES=be.tasteful-3OBJ

'It doesn't taste nice'.

(34a) *kó* *xa = $\emptyset = <z>\check{\sigma}' = nǒ = r$* *ñǔní* \emptyset
 because INT=3.PRES=<NPS>arrive.here.B-B=DEF.SG=SG smell REL

$\emptyset = \check{r}\acute{a}-n-tx\ddot{o}-\emptyset . . .$

(3.)PRES=INT.PC-ST-be.stinky-3OBJ

'Because a very stinky smell arrives . . .'.

(34b) *ká = $\emptyset = n-tšé-\emptyset$*
 INT.PC=(3.)PRES=ST-be.cold-3OBJ

'It's rather cold'.

For the imperfect, stative verbs employ the FW *már*. As seen in 3, this is a characteristic of all PC verbs. However, in contrast to Class 1 PC verbs, this is the only FW available for a stative verb to express imperfect (in the positive polarity). Besides the FW *már*, these verbs also use the FW *dar* for the future, which is a marker unique to stative verbs. An example is given in (35):

(35) *nú* *ma* *bötsí* *dar = ñ-hěts'i- \emptyset*
 DEF(PL) 1POSS child (3.)FUT.ST=ST-be.tall-3OBJ

'My children will be tall'.

3.2.3.1. A word about PC intensifiers. When a PC verb occurs with the intensifier *řá-* ‘very’, the FWs *már* and *dar* drop the sonorant /t/ (*r*). An example is given in (36):

- (36) *gǎ'thó ha (bi) = <z>u^h-k-a = ma wá*
 all where 3.PST=<NPS>reach.S-1OBJ/DAT-B=1POSS foot/leg
má = řá-ň-’ú-∅
 (3.)IMP.PC=INT.PC-ST-be.painful-3OBJ

‘Everywhere where it (i.e., the hail) hit my legs was very painful’.

In (34) above, I showed that stative verbs use a zero FW with PC intensifiers to express a present state of affairs. Alternatively, the intensifier *ká* ‘rather’ may also be used with the perfect, but when it does so, it follows the FW *xi*. This is shown in (37):

- (37) *xi = ká = n-tsé-∅*
 (3.)PERF=INT.PC=ST-be.cold-3OBJ
 ‘It’s rather cold’.

The word order of FWs in (37) is unusual because in other circumstances *ká* ‘rather’ precedes FWs, as shown, for example, in (38):

- (38) *nú'mu n-dá = pěng-i ká = már = n-tsé-∅*
 when IMP-1.PST=return-F INT.PC=(3.)IMP.PC=ST-be.cold-3OBJ
 ‘When I came back, it was rather cold’.

The prefix *řá-* ‘very’ is not used when the verb is negated, as shown in (39a), although *ká* ‘rather’ may be used, as in (39b):

- (39a) **pórke híngí ∅ = řá-n-kǔhi-∅*
 because NEG (3.)PRES=INT.PC-ST-be.tasty-3OBJ
 Intended reading: ‘because it isn’t very tasty’.
- (39b) *yǒ gí = ta-∅-i pórke ká = híngi*
 IMPER.NEG 2.PRES=buy-3OBJ-F because INT.PC=NEG
∅ = kǔhi-∅
 (3.)PRES=be.tasty-3OBJ

‘Don’t buy it because it’s not that tasty’.

The distribution in (39) suggests that there is some semantic contrast involving the degree of intensity of the property conveyed. I have assumed that this degree is subtler for *ká* and have thus regarded it as closer to possible English counterparts like *rather* or *enough*. The prefix *řá-* appears to have a

more intensive reading, and I have thus translated it using English *very* or *much*. It is rather common to find both markers used together with the same PC verb, as, for example, in (40):

- (40) *nǒ = r* *déhé = 'á* *ká = ∅ = řá-ñ-'ǔxi-∅*;
 DEF.SG=SG water=3SG INT.PC=(3.)PRES=INT.PC-ST-be.salty-3OBJ
dá = tsǒñ = hé, *dá = n-txá = hé* *kó*
 1.PST=arrive.there.S=1PL.EX 1.PST=M-bathe.S=1PL.EX because
de_gē^hnú *ká = ∅ = řá-m-pá-∅*
 over.there INT.PC=(3.)PRES=INT.PC-ST-be.hot-3OBJ

'That water is very very salty; we got there and we went for a bath because it was so very hot over there'.

3.2.4. Lack of a morphosyntactic bound form. The vast majority of verbs in Otomi have a stem formative (e.g., *tsé + t'²-i* 'get cold', *k'ó' + ts'-i* 'go to see someone', *zó + n-i* 'weep', etc.). For the most part, such stem formatives are no longer semantically transparent. Palancar (2004) has shown that such verbs have two inflected forms, i.e., a "free form" and a "bound form." Free forms are phonologically free and are used before a prosodic or a clausal boundary. Bound forms are phonologically bound and occur intra-clausally. The examples in (41) illustrate the use of the forms (|| represents a prosodic boundary):

- (41a) Free Form
dá = hǒ^hk-∅-i
 1.PST=fix-3OBJ-F
 'I fixed it'.
- (41b) Bound Form
dá = hǒ^hk-∅-a = ma *ngǔ ||*
 1.PST=fix-3OBJ-B=1POSS house
 'I fixed my house'.

(41a) shows the free form of the verb *hǒ^hki* 'fix'; the verb occurs here at both a prosodic pause and a clausal boundary. In contrast, (41b) illustrates the use of the bound form of the same verb. The verb in (41b) occurs intraclausally (i.e., it is followed by an NP). The suffix *-i* is the default realization of a free form, while the suffix *-a* is the common exponent of a bound form.¹⁸ The

¹⁸ In general, when the morphological root of a verb bears a high tone, the root commonly receives an ascending tone in the bound form (e.g., *hǒ^hki* vs. *hǒ^hk-a* 'fix').

forms are not interchangeable. For example, a free form may not be used in the context of a bound form, as shown in (42):

- (42) **dá* = *hó^hk-∅-i* = *ma* *ngǔ* ||
 1.PST=fix-3OBJ-F=1POSS house
 ‘I fixed the house’.

A considerable number of stative verbs are verbs with stem formatives (i.e., *ts’ǔ* + *t’-i* ‘be thin (for wood, thread)’, *hě* + *ts’-i* ‘be high’, *thé* + *n-i* ‘be red’, etc.). In principle, such verbs should be eligible to have the two forms just like other verbs. However, this is not what we find (at least not in the context where the forms are expected). The pattern is illustrated in (43):¹⁹

- (43a) *ma* *ngǔ* *híngí* \emptyset = *t’áxi-∅* ||
 1.POSS house NEG (3.)PRES=be.white-3OBJ
 ‘My house isn’t white’.
- (43b) *híngí* \emptyset = *t’áxi-∅* *ma* *ngǔ* ||
 NEG (3.)PRES=be.white-3OBJ 1POSS house
 ‘My house isn’t white’.

If (41b) and (43b) are compared, we observe that the verb *t’áxi* ‘be white’ does not undergo any change in the context where the use of a bound form is expected. In fact, a form such as *t’ǎxa*, the predictable bound form of *t’áxi* ‘be white’, is impossible in the expected context:

- (44) **híngí* \emptyset = *t’ǎx-∅-a* = *ma* *ngǔ* ||
 NEG (3.)PRES=be.white-3OBJ-B=1POSS house
 Intended meaning: ‘My house isn’t white’.

Despite the fact that stative verbs deviate in this respect from other verbs, the impossibility of (44) does not mean that the verbs lack a morphological bound form. In 6, I show how stative verbs occur in bound form when they are members of nominal compounds (i.e., *t’ǎxa-ngǔ* ‘white house’).

3.3. PC verbs in other dialects. All dialects of Otomi have a similar large class of stative verbs that express PCs. In this respect, stative verbs are a dia-systemic feature of Otomi as a language. Much less is known, however, of the existence of a lexical subclass similar to Class 1 PC verbs in 3.1. In this respect, it is not known whether the different variants have a general class of PC verbs like SI Otomi or whether PC verbs are only circumscribed to stative verbs.

¹⁹ The zero for the third-person object anaphora is placed in PC verbs after the root because the final /i/ is not a morpheme in this construction, in contrast to (41) where it is.

and function to the Spanish prefix *re-* /*re*/, which is an adjective intensifier (i.e., *re-grande* ‘very big’). Prefixes of this kind occur only in the dialects of Mexico State and Querétaro State. It is unclear the relationship they bear, but I believe the prefix in SI Otomi emerged from the reanalysis of a genuine Otomi stative marker which at some point became an intensifier on the basis of the Spanish prefix by virtue of their functional and phonological resemblance.

Similarly, all dialects have a stative/active split involving these verbs. The pattern as illustrated in (46) is realized with the dialects’ own morphological idiosyncrasies (for convenience, I have adapted the glosses in 46*a* to the format used in this paper and have glossed 46*b* and 46*c*, which are not explicitly glossed in the sources):

- (46*a*) *m-∅-rá_n-nzæ[?]-ki*
 IMP-(3.)PRES-ST-be.strong-2OBJ
 ‘You were strong’. (San Felipe Santiago Otomi; Andrews 1993:36)
- (46*b*) *xú = n-hěts’i-ga = wi*
 (3.)ST.PRES=ST-be.high-1OBJ=DU(.IN)
 ‘We’re tall (you and me)’. (Sierra Otomi; Voigtlander and Echevoyen 1985:70)
- (46*c*) *rA-no-?kí-hí*
 (3.)CONT.PRES-be.fat-2OBJ-PL
 ‘Y’all are tall’. (Toluca Otomi; Lastra 1992:21)

As for the behavior or pertinence of a morphosyntactic free-form, little is known about this phenomenon in dialects other than SI, but one finds verbs in bound form sporadically when the verbs are members of compounds, e.g., Sierra Otomi *t’ax-a-hwä* (be.white-B-fish) ‘trout’ (Voigtlander and Echevoyen 1985:66), which suggests that forms such as *t’áxi* and the like are free forms.

Finally, it is extremely common to find cognates as lexical members of stative verbs across the dialects. The phenomenon is illustrated in (47) with five verbs picked at random from Hekking and Andrés de Jesús (1989), Hernández Cruz et al. (2004), Wallis (1956), Voigtlander and Echevoyen (1985), Andrews (1993), and Lastra (1992; 1997):

(47)	‘be tasteful’	‘be high’	‘be fat’	‘be cold’	‘be white’
San Ildelfonso Tultepec	<i>kũhi</i>	<i>hěts’i</i>	<i>noho</i>	<i>tsé</i>	<i>t’áxi</i>
Santiago Mexquititlán	<i>kuhi</i>	<i>hñets’i</i>	<i>noho</i>	<i>tse</i>	<i>t’áxi</i>
Mezquitil	<i>kũhi</i>	<i>hñets’i</i>	<i>noho</i>	<i>tsé</i>	<i>t’áxi</i>
Sierra	<i>k̄hi</i>	<i>hěts’i</i>	<i>noho</i>	<i>se</i>	<i>t’áxi</i>
San Felipe Santiago	—	(<i>n̄te</i> ‘grown’)	<i>noho</i>	<i>tsæ</i>	<i>t’áši</i>

TABLE 2
PC NOUNS IN SI OTOMI

		As a Noun		As a PC
Physical	<i>měfo</i> ¹	'a short person'	>	'short (for a person)'
	<i>ndó'yó</i>	'a skinny person'	>	'thin (for an animate)'
Psychological	<i>hōga-jō't</i>	'a good person'	>	'good (for a person)'
	<i>giint'ei</i>	'a jealous person'	>	'jealous'
	<i>bō^hka</i>	'a cruel/aggressive person'	>	'cruel'
	<i>zōne</i>	'an evil person/vampire'	>	'evil'
	<i>ndúxte</i>	'a naughty person'	>	'naughty'
	<i>xóngo/a</i>	'a silly person (man/woman)'	>	'silly'
	<i>bíbo</i>	'a clever person'	>	'clever'
	<i>rā_mpefi</i>	'a hard-working person'	>	'hard-working'
	<i>rā_mpeño</i>	'an industrious person'	>	'industrious'
	<i>t'ixfáni</i>	'a drunkard'	>	'drunk'
	<i>yozna</i>	'a homosexual'	>	'homosexual'
Sex	<i>ndō</i>	'a man'	>	'male'
	<i>'béhñó</i>	'a woman'	>	'female'
	<i>dōsya</i>	'an old man'	>	'male (for an old person)'
	<i>xíta</i>	'an old man'	>	'male (for an old person)'
	<i>txú</i>	'an old woman'	>	'female (for an old person)'
Marital status	<i>ts'ünt'ú_tho</i>	'just a young man'	>	'single (for a male person)'
	<i>ntxúntsi_tho</i>	'just a young woman'	>	'single (for a female person)'

¹The noun *měfo* 'a short person' appears most frequently in nominal compounds such as *měfo-ndō* 'short man' or *měfo-txú* 'short woman'. Compounds like these are treated in detail in 6 in the text.

Toluca	<i>kíhí</i>	<i>hǐtsi</i>	<i>nohó</i>	<i>cé /tsél/ t'ǎši</i>
Ixtenco	<i>kihi</i>	<i>hēc'í /hěts'í/</i>	<i>nohó</i>	<i>cé /tsél/ (n)t'ǎši</i>

4. PC nouns. In the previous sections, I have studied in detail the lexemes expressing PCs that belong to Group 1, as it was characterized in 2. I have presented evidence for the treatment of such lexemes as verbs and have described their characteristics as a subclass with respect to the bulk of intransitive verbs. In this section, I study the PC lexemes pertaining to Group 2, and I show that such lexemes are nouns. The lexemes in question express properties of Human Propensity of the type in Dixon (1982). A list of the most representative members of this group is given in table 2. All the lexical elements in table 2 express properties ascribable to human beings. This is one of the defining features of this class. The last four elements in table 2 may also occur as Class 1 PC verbs (table 1),²² but when they do so they express the age of a given person, whose sex is known.

²²The lexeme *txu* has high tone when it is a verb and low tone when it is a noun. Such a tonal pattern is unique to this word.

Nouns are heads of NPs and, in Otomi, the head of an extended NP occurs in final position. This may be seen in (48):

(48a) (*nǒ* = *r* *mǎ^hjǒ*)_{NP} *bi* = <*d*>*ú*
 DEF.SG=SG priest 3.PST=<NPS>die
 ‘The priest died’.

(48b) (*nú* *ma* = ‘*rǎ* *ya* *mbóho*)_{NP}
 DEF other=IND.PL PL mestize.man
 bi = <*z*>*ú^hk-Ø-i*
 3.PST=<NPS>get.somebody.out-3OBJ-F
 ‘The other men brought him out’.

The nouns in (48) work as syntactic actants WFM. This is a crucial characteristic of nouns. PC lexemes of Group 2 behave in this respect like other nouns in Otomi. This is illustrated in (49):

(49a) *yá* *xi* = <*m*>*á*= *nú* (*nǒ* = *r*
 P 3.PERF=<NPS>go=DISTAL.I DEF=SG
 günt’ei/xóngo/bó^hka/etc.)_{NP}
 jealous/silly/cruel/etc.

‘The jealous/silly/cruel (man) is passing by over there’.

(49b) *dí* = *hong-Ø-a* = *n’ǎ* ‘*bego*, *pe* *dí* = *ne-Ø*
 1.PRES=look.for.B-3OBJ-B=one worker but 1.PRES=want-3OBJ
 (‘*na* *řá mpefi*)_{NP}
 IND.SG hard.working

‘I’m looking for a worker, but I want a hard-working (one)’.

In contrast to verbs, nouns are not expected to function as predicates WFM. This also happens with Otomi nouns, as nouns require a special nominal predication construction to function as verbs, a syntactic measure. The construction in question is encoded by a set of special FWs together with an NP consisting of the noun plus a number determiner. (50a) illustrates a singular NP, whereas (50b) illustrates a plural one:

(50a) *nǒ* = *r* *ndǒ* || Ø = *ar* *mǎ^hjǒ*
 DEF=SG man 3.PRES=SG priest
 ‘The man is a priest’.

(50b) *nú* = ‘*u* *ya* *ndǒ* || Ø = *ya* *mǎ^hjǒ*
 DEF=DEM.PL.II PL 3.PRES=PL priest
 ‘These men are priests’.

Like nouns in nominal predication, the PC lexemes of Group 2 predicate their respective property by means of the same additional apparatus used for nouns, as in (50). This is illustrated in (51):

(51a) *nǒ = r jǒ'í* || $\emptyset = ar$ ***günt'ei***
 DEF=SG person 3.PRES=SG jealous
 'The man is jealous'.

(51b) *gǎ'thó nú ma míxi* || $\emptyset = ya$ ***ndúxte***
 all DEF(PL) 1POSS cat 3.PRES=PL naughty
 'All my cats are naughty'.

In (52), I present the full paradigm of this nominal predication construction with a singular NP for the present and the imperfect. The paradigm is illustrated both with the noun *mǎ^hjǒ* 'priest' and the PC noun *bǒ^hka* 'cruel':

(52) Present	Imperfect
1st	1st
<i>dá = r mǎ^hjǒ/bǒ^hka</i>	<i>n-dá = r mǎ^hjǒ/bǒ^hka</i>
1.PRES.NP=SG priest/cruel	IMP-1.PRES.NP=SG priest/cruel
'I'm a priest/cruel'.	'I was a priest/cruel'.
2nd	2nd
<i>gá = r mǎ^hjǒ/bǒ^hka</i>	<i>n-gá = r mǎ^hjǒ/bǒ^hka</i>
2.PRES.NP=SG priest/cruel	IMP-2.PRES.NP=SG priest/cruel
'You're a priest/cruel'.	'You were a priest/cruel'.
3rd	3rd
$\emptyset = ar$ <i>mǎ^hjǒ/bǒ^hka</i>	<i>má = r mǎ^hjǒ/bǒ^hka</i>
3.PRES=SG priest/cruel	3.IMP.NP=SG priest/cruel
'He's a priest/cruel'.	'He was a priest/cruel'.

Like PC verbs, PC nouns may receive the intensifier *řá-* 'very'. This is shown in (53). (53b) additionally shows that the singular determiner is dropped when the prefix is used:²³

(53a) *nǒ ma bǒtsí* $\emptyset = řá-bíbo$
 DEF.SG 1POSS boy 3.PRES=INT.PC-clever
 'My boy is very clever'.

²³ As shown in (36) above, the same phenomenon also occurs with PC verbs.

- (53b) *nǒ* = *r* *mbóhǒ* = 'á *ma* = *řá-t'ixfǎni*
 DEF.SG=SG mestize.man=3SG 3.IMP.NP=INT.PC-drunkard
 'That man was very drunk' (lit., '... a very drunkard').

The use of this intensifier *řa-* 'very' with such nouns is not a complete anomaly given that PC nouns convey a property that can be measured.²⁴ The prefix has lexicalized in two deverbal nouns, *řá_mpefi* 'a hard-working person' and *řá_mpeño* 'an industrious person'.²⁵

5. Overlaps across word classes expressing PCs. PC verbs and PC nouns show two main grammatical overlaps in SI Otomi. Predication of a third person in the present with the intensifier *řa-* 'very' is formally identical across the two classes. This is illustrated in (54):

- (54a) Class 1 PC verb
 Ø = *řá-dótá*
 3.PRES=INT.PC-be.big
 'He's very big'.
- (54b) Stative verb (Class 2 PC verb)
 Ø = *řá-ñ-hěts'i-Ø*
 (3.)PRES=INT.PC-ST-be.high-3OBJ
 'He's very tall' ~ 'It's very high'.
- (54c) PC noun
 Ø = *řá-bǒ^hka*
 3.PRES=INT.PC-cruel
 'He's very cruel'.

Similarly, the predication of a third person in the imperfect is again formally identical when this prefix is used, as shown in (55):

²⁴ The use of intensifiers with PC lexemes should NOT be taken to be a piece of evidence that points to the adjectival nature of such lexemes. The use of intensifiers is motivated only because of the semantics of PC lexemes, which express measurable properties.

²⁵ Apart from PC nouns, the prefix is found sporadically in a few nouns that express a mass (i.e., *řá-bojǒ* 'a lot of money') or nouns that express quantity (i.e., *řá-múndo* 'a (very big) heap'). A noun such as *zǒne* 'a male witch' or 'a vampire' is not typically treated as expressing a PC elsewhere, but since this noun may also be used to depict a person as evil, I have included it in table 2.

(55a) Class 1 PC verb

má = *řá-dótá*

3.IMP.PC=INT.PC-be.big

‘He was very big’.

(55b) Stative verb (Class 2 PC verb)

má = *řá-ñ-hěts’i-∅*

(3.)IMP.PC=INT.PC-ST-be.high-3OBJ

‘He was very tall’ ~ ‘It was very high’.

(55c) PC noun

má = *řá-bó^hka*

3.IMP.NP=INT.PC-cruel

‘He was very cruel’.

The grammatical overlap between PC verbs and PC nouns perhaps is a consequence of borrowing across paradigms. This topic deserves further study.²⁶

In the previous sections, I have described the characteristics of both verbs and nouns that express PCs in SI Otomi. These lexemes may also appear in nominal compounds. The study of such compounds is crucial to the present discussion because the construction in question has been taken in the literature to represent a nominal attributive construction. This is the topic of the next section.

6. PCs in nominal compounds: adjectives in Otomi? In 1.1, I presented two examples of nominal compounds that involve PC lexemes. More examples of the same structure are given in (56):

(56a) (*n’a* *dótá-dó*) *ja* = ∅ = *’bú^h* = *’pú* *nú*
 one be.big-stone EXIST=3.PRES=EXIST.S=DISTAL.II LOC
há = *r* *wá* *nǒ* = *r* *kwéba* = *’á*
 LOC.3POSS=SG foot/leg DEF.SG=SG cave=3SG

‘There’s one big stone at the foot of that cave’.

²⁶ The similarities in (54) and (55) appear to point to an extension of the use of the inflectional morphology proper of a nominal predication construction to the realm of PC verbal predication. From this extension, the language obtained a generalized marker to encode a PC lexeme in predication (in third-person imperfect). There is yet another piece of evidence that suggests this extension took place. The imperfect FW *már* used with PC verbs resembles the FW *má* plus the singular enclitic =*r* used in the nominal predication construction.

(56b) *dá* = *hand-∅-a*(= *nǒ* *ri* *tx'úlo-ngǔ*)
 1.PST=see-3OBJ-B=DEF.SG 2POSS be.small-house
 'I saw your small house'.

(56c) *ya dá* = *tǎng-∅-a*(= 'na *ma* *k'ǎxt'-a-pǎhni*)
 P 1.PST=buy.B-3OBJ-B=IND.SG 1POSS be.yellow-B-sweater
 'I bought my yellow sweater'.

The lexical combinations in (56) are instances of an extremely common grammatical strategy in Otomi to express PCs. I take lexical combinations like these to be nominal compounds, i.e., *dǒtá-dó* 'big-stone' in (56a), *tx'úlo-ngǔ* 'small-house' in (56b), and *k'ǎxt'-a-pǎhni* 'yellow-sweater' in (56c).

However, the (pseudo-) attributive context in which such PC verbs occur has led a number of authors to claim that such lexical combinations are adjectives in a natural attributive context. Accordingly, the theoretical controversy rests in treating the NPs in (56) as reflecting a syntactic parsing of the type (*n'a dǒtá dó*) 'one big house' for (56a), (*nǒ ri tx'úlo ngǔ*) 'your small house' for (56b), and (*'na ma k'ǎxt'a pǎhni*) 'a yellow sweater of mine' for (56c), where the lexemes *dǒtá* 'big', *tx'úlo* 'small', and *k'ǎxt'a* 'yellow' are seen as attributive adjectives. This is, for example, the position taken by Lastra (1992) and Andrews (1993), who invariably write them as phonological words.

Contrary to such a view, I claim here that these elements reflect nominal compounds rather than a syntactic (Adj N) structure, at least for SI Otomi and other dialects. In my analysis, the first member of such compounds is a PC verb (i.e., *dǒtá* 'be big' in 56a and *tx'úlo* 'be small' in 56b are Class 1 PC verbs, while *k'ǎxt'a* 'be yellow' in 56c is the bound form of the stative verb *k'ǎxt'i* 'be yellow'). Similar compounds involving PC nouns also exist. In order to provide evidence that supports my claims, I compare the combinations in (55) with other nominal compounds in the language.

SI Otomi has two types of (synthetic) nominal compounds depending on the constituent order between the head and the dependent. The two types are illustrated in (57), where the head appears in small caps:

(57a) Type 1
'BQTS'E-hmé
 basket-tortilla
 'tortilla basket'

(57b) Type 2
měfo-TXÚ
 short.person-old.woman
 'short old woman'

All compounds in (57) are nouns. Type 1 compounds commonly express function (i.e., ‘basket designed for the tortillas’, etc.) or source material (i.e., ‘table made of wood’, etc.). In contrast, Type 2 compounds instantiate a special compounding structure where the dependent member expresses a PC about the head. The lexical combinations in (56) are typical instances of Type 2 compounds. All lexemes expressing PCs, regardless of whether they are verbs or nouns, are in principle potential fillers for the dependent slot in this structure. Additionally, as we shall see in more detail in 6.4.4, there are also a number of acategorical lexemes which can only be used in Type 2 compounds.

Several characteristics identify nominal compounds in (SI) Otomi: (a) their syntactic behavior as head nouns in nominal predication; (b) the morphological adjusted forms they present; (c) their function as novel lexemes; (d) their occurrence in lexical pairs; and (e) their restrictions concerning internal modification. Such criteria are common in linguistic theory to define compounds (see Fabb 1998 or Spencer 1991). I examine each criterion separately in order to show that the lexical combinations in (56) are nominal compounds.

6.1. PC compounds as head nouns in nominal predication. Compounds behave as head nouns in the nominal predication construction introduced in 4. This construction is encoded by a set of special FWs together with an NP consisting of the noun plus a number determiner. The construction is illustrated again in (58):

- (58) *nú = ya* *řébéldé* *má(= ya* *bé)*_{NP}
 DEF=PL rebel 3.IMP.NP=PL thief

‘The rebels were thieves’.

In the nominal predication construction instantiated in (58), the head noun in the predicate has a generic reading depicting a kind. This is a property of the construction because nouns in predication cannot be modified by other determiners. The restriction is illustrated by the ungrammaticality of (59):

- (59a) **dá = nú = nŭ = r* *ndŏ* *há = r* *fóto*
 1.PRES.NP=DEF=DEM.SG.II=SG man LOC=SG picture

Intended meaning: ‘I’m that man in the picture’.

- (59b) **nŏ = r* *Tóño* *má = ri* *’bego*
 DEF.SG=SG Antonio 3.IMP.NP=2.POSS worker

Intended meaning: ‘Antonio was your worker’.

The impossibility of examples like (59) illustrates that the head noun in a predicative NP cannot be modified. The examples also show that the nominal predication construction in Otomi does not express identification, as is common in other languages like English or Spanish. The language has another construction devoted to expressing identification, which involves a cleft relative clause, as illustrated in (60):

- (60a) $n\check{o} = r$ *Tóño* ($g\acute{o} = \emptyset$ *ma* *'bego*
 DEF.SG=SG Antonio REL.FOC=3.PRES 1POSS worker
hm\check{o} = 'á)
 previously=3SG

‘Antonio was my laborer’ (literally, ‘. . . is the one who is previously my laborer’).

- (60b) $n\check{o} = r$ *mb\acute{o}h\acute{o}* ($g\acute{o} = m\acute{a} = r$ *b\acute{é}(= 'á)* \emptyset
 DEF.SG=SG mestize REL.FOC=3.IMP.NP=SG thief(=3SG) REL
d\acute{a} = hant = h\acute{ú})
 1.PST=see.s=PL(.IN)

‘The mestize man happened to be the thief we saw’.

- (60c) $n\acute{u} = n\acute{a} = r$ *j\ddot{o}'í* ($g\acute{é} = \emptyset = r$ *Xúwa = 'á*)
 DEF=DEM.SG.I=SG person REL.FOC=3.PRES=SG John=3SG
 ‘This man (here) is John’.²⁷

Compound nouns can also appear as head nouns in the nominal predication construction. This can be seen in (61) with Type 1 compounds, but more crucially in (62) with Type 2 compounds:

- (61a) $n\check{o} = r$ *ng\acute{o}* \emptyset *g\acute{a} = 'r\acute{a}^h-kagí* *m\acute{a} = (r*
 DEF.SG=SG meat REL 2.PST=give.to.1/2-1OBJ/DAT 3.IMP.NP=SG
ng\acute{o}-pts'údi)
 meat-pork.B

‘The meat you gave me was pork-meat’.

- (61b) $n\acute{u} = n\acute{a} = r$ *'b\acute{o}ts'e* $\emptyset = (ar$ *'b\acute{o}ts'e-hmé)*
 DEF=DEM.SG.I=SG basket-tortilla 3.PRES=SG basket-tortilla
 ‘This basket (here) is a tortilla basket’.

²⁷ I want to thank an anonymous referee for suggesting an example with a proper name in order to strengthen the argument that this construction expresses identification. The meaning in (60c) cannot be expressed with the ordinary nominal predication construction.

(62a) *nú = 'u ya ndǒ ∅ = (ya mǎfo-ndǒ)*
 DEF=DEM.PL.II PL man 3.PRES=PL short-man

‘Those men are short’ (lit., ‘. . . are short-men’).

(62b) *nǒ = r ts'oi ∅ bi = <d>ǎhm-i*
 DEF.SG=SG pot REL 3.PST=<NPS>break.volume-F

má = (r hǔ-ts'oi)
 3.IMP.NP=SG be.heavy-pot

‘The pot that broke was heavy’ (lit., ‘. . . was a heavy-pot’).

(62c) *nǒ ma míxi híng-∅ = (ar 'bó-mtxí)*
 DEF.SG 1POSS cat NEG-3.PRES=SG black-cat.B

‘My cat is not a black’ (lit., ‘. . . is not a black-cat’).

The examples in (62) show Type 2 compounds. Given the restriction of the nominal predication construction with regard to internal modification, the PC lexemes in (62)—*mǎfo* in (62a), *hǔ* in (62b), and *'bó* in (62c)—are not syntactic modifiers but genuine nominal compound members in a dependent slot.

6.2. The morphology of nominal compounds. A number of Otomi compounds bear a nasal interfix between the members involved in the combination. This interfix is shown in (63):

(63) <i>n'á-n-jéyá</i>	(one-INTER-year)	‘one year (i.e., as a span of time)’
<i>hǔ-n-zönö</i>	(three-INTER-month)	‘three months (i.e., as a span of time)’
<i>txí-dáda-n-hyádi</i>	(DIM-father-INTER-sun)	‘sun’ (lit., ‘honorable father sun’)
<i>txí-nǒnö-n-zönö</i>	(DIM-mother-INTER-moon)	‘moon’ (lit., ‘honorable mother moon’)

Such an interfix is occasionally found in PC compounds too, as shown in (64):

(64) <i>hǔ-n-dó</i>	(be.heavy-INTER-stone)	‘a type of stone which is heavy’
<i>(n-)xó-n-thé</i>	(ST-be.stinky-INTER-water.B)	‘sweat’

On the other hand, compounds normally display morphophonologically adjusted members. This is a typical feature of many compounds in Otomi, where both heads and dependents may display the adjustments. Examples of adjusted heads are shown in (65). This only happens in Type 2 compounds, where the head occurs in second (and final) position:

(65)	Adjustment of Head		Bound	Free
	Type 1		Morpheme	Morpheme
	<i>dö-THÉ</i> (big-water.B)	‘river’	<i>-thé</i>	<i>déhe</i>
	<i>'bó-MTXÍ</i> (black-cat.B)	‘black cat’	<i>-mtxí</i>	<i>míxi</i>
	<i>ndǒ-FNÍ</i> (man-horse.B)	‘stallion’	<i>-fní</i>	<i>fǎni</i>

Adjusted dependents may be found in both types of compounds:

(66)	Adjustment of Dependent		Bound	Free
	Type 1		Morpheme	Morpheme
(66a)	<i>NGŌ-pts'údi</i> (meat-pig.B)	'pork meat'	<i>-pts'údi</i>	<i>ts'údi</i>
	Type 2			
(66b)	<i>t'ǎx-a-NGŪ</i> (be.white-B-house)	'white house'	<i>t'ǎx-a-</i>	<i>t'ǎx-i</i>
(66c)	<i>nǒxk-a-NGŌ</i> (get.stale-B-meat)	'stale meat'	<i>-nǒxk-a-</i>	<i>nǒxk-i</i>

As shown in (66), dependent members in compounds can display adjusted forms. The dependents in (66b) and (66c) are verbs: *t'ǎxi* 'be white' is a stative verb, while *nǒxki* 'get stale' is an inchoative verb. Verbs as dependents in Type 2 compounds always occur in bound form, regardless of being inchoative verbs (e.g., *nǒsk-i* (get.stale-F) vs. *nǒsk-a* (get.stale-B); *'ǎxk-i* (get.sour-F) vs. *'ǎxk-a* (get.sour-B); etc., or stative verbs (e.g., *ts'ǔt'-i* (be.thin (for wood)-F) vs. *ts'ǔt'-a* (be.thin (for wood)-B); *t'ǎx-i* (be.white-F) vs. *t'ǎx(-a)* (be.white-B); etc. As I pointed out in 3.2.4, the bound form of stative verbs is used only in compounds and not in all contexts expected for other verbs.

There is a tonal contrast between the bound form *t'ǎx(-a)* (be.white-B) in (66b) if compared to the free form *t'ǎx-i* (be.white-F). The contrast involves an ascending tone vs. a high tone. The pattern is seen elsewhere, i.e., stative verb *thén-i* (be.red-F) ~ *thěng-a* (be.red-B), as in *thěng-a-xǒju* 'red ant'; inchoative verb *'ót'-i* (get.dry-F) ~ *'ót'-a* (get.dry-B), as in *'ót'-a-ndó^h-jó* (get.dry-B-seed.of.fruit-nopal.fruit.B) 'dry nopal-fruit seed'; etc. This contrast does NOT reflect a (suprasegmental) derivational operation by virtue of which a verb turns into an adjective. As mentioned in n. 17, verbs commonly receive an ascending tone in their bound form if they have a high tone in their free form.²⁸

6.3. Compounds as lexical units. One of the fundamental characteristics of compounds is their derivational ability to produce new lexemes. PC compounds are widely used in the language to produce new words that express subordinate members of the given category expressed in the head noun. When the compounds are conventionalized, it suggests that the subordinate categories they depict are well-established concepts expressing conventional cultural entities. This is the case, for example, for the different types of chilies found in the lexicon of SI Otomi; the most relevant ones are given in (67) (the translation of the different types is rendered in the

²⁸ Besides, elements in compounds often show a rather idiosyncratic behavior regarding tone; it is not rare to find the same elements bearing different tones in similar contexts, e.g., Mezquital Otomi *dáta-nā* (be.big-head) '(a) big head' vs. *dáta-xót'o* (be.big-sunflower) '(a large) sunflower'.

conventional name such chilies receive in the Spanish spoken in Central Mexico):

(67a)	Noun-Noun	<i>rǒxa-ñ'í</i>	(berry-chili)	'chile miracielo'
(67b)	Stative Verb-Noun	<i>dǒtá-ñ'í</i>	(be.big-chili)	'chile jalapeño'
(67c)		<i>thěng-a-ñ'í</i>	(be.red-B-chili)	'chile cascabel/guajillo'
(67d)		<i>gwéro-ñ'í</i>	(be.blond-chili)	'chile amarillo'
(67e)		<i>k'ang-a-ñ'í</i>	(be.green-B-chili)	'chilaca'
(67f)	Acategorial-Noun	<i>'bó-ñ'í</i>	(black-chili)	'chile pasilla'
(67g)		<i>t'úlo-ñ'í</i>	(small-chili)	'chile serrano'

The compounds in (67) reflect different structures. (67a) represents a noun-noun combination, while the examples in (67b)–(67e) show a combination of a stative verb plus a noun. The two examples in (67f) and (67g) are more difficult to categorize. As we shall see in more detail in 6.4.4, morphemes such as *'bó-* 'black' and *t'úlo-* 'small' are "acategorial" lexical elements.

On the other hand, once a lexical combination is well established and lexicalized, a semantic shift from the original compositional meaning commonly occurs. Because of this extension, the new meaning is no longer recoverable from the meanings of the parts, and the combination is stored as a set expression (i.e., *'ñü-jí* (path-blood) with the meaning 'vein'). Various examples of Type 2 compounds showing this type of lexical shift are given in (68):

(68)	<i>dǒtá-m'asmé</i>	(be.big-napkin)	'tablecloth'
	<i>dǒtá-déhe</i>	(be.big-water)	'sea'
	<i>dǒtá-'ñü</i>	(be.big-path)	'road'
	<i>xǒ-thé</i>	(be.stinky-water.B)	'stagnant water'
	<i>(n-)xǒ-n-thé</i>	(be.stinky-INTER-water-B)	'sweat'
	<i>t'ǎx-a-hwó</i>	(be.white-B-fish)	'trout'

However, not all lexical combinations are possible or extant. This serves as another important piece of evidence that such lexical combinations should be treated as compounds, and not as adjectives, in syntactic attribution. For example, potential combinations such as **hěts'-a-ngǔ* (be.tall-B-house) for 'high house', **pǐ'-libro* (be.thick.B-B-book) for 'thick book', **mǎ-t'ó* (be.long-stick) for 'long stick', **no-'-bǒtsí* (be.fat-B-child) for 'fat child', or **'ó-'-dú^{htu}* (be.rough.B-B-clothes) for 'rough cloth(es)', and many others are not possible. The same is true for **ndǒ-zǎ* (be.thick (of wood)-tree/wood) for 'thick wood', even though the mass noun *ts'út'-a-zǎ* (be.thin (of wood)-B-tree/wood) for 'thin wood' exists and designates a type of wood suitable for carving ceiling beams.

Similarly, compounds such as *hǔ-n-dó* (be.heavy-INTER-stone) and *hǔ-ts'oi* (be.heavy-pot) are actual compounds designating types of stones and

pots, which are heavy in contrast to others that are not. In contrast, a noun such as **hǔ-káhá* (be.heavy-box) is impossible because boxes are not heavy by essence but only by virtue of the contents they bear inside. Other relevant examples include nonexistent **tsé-pá* (be.cold-day) for ‘cold day’ or **tsé-yé* (be.cold-hand) for ‘cold hand(s)’, which do not express conventional concepts either. However, both the compounds *tsé-déhe* (be.cold-water) and *tsé-thé* (be.cold-water.B) do exist. The former depicts a type of refreshment, which is drunk cold, and the latter, a type of warm drink, which has become cold.

Finally, despite the fact that the language has *’ú-t’èi* (be.sweet-atole) ‘sweet atole’²⁹ with the stative verb *’ú* ‘be sweet’, Type 2 compounds often require the inchoative lexical pairs of stative verbs. For example, compounds exist with inchoative verbs such as *’ǔxk-a-déhe* (get.salty.B-water) ‘salty water/soup’, *’ǔxk-a-límo* (get.bitter-B-lemon) ‘bitter lemon’, and *nǒxk-a-ngó* (get.stale.B-meat) ‘stale meat’. On the other hand, combinations with their stative counterparts are not possible, such as *’úxi* ‘be salty’ (e.g., **’ǔx-a-déhe*), *’ixi* ‘be bitter’ (e.g., **’ix-a-límo*), or *nǒxi* ‘be stale’ (e.g., **nǒx-a-ngó*).

6.4. Attributive-like features of PC compounds. In the previous section, I have presented much evidence to support treating lexical combinatory structures of the type *t’ǎx-a-ngǔ* (be.white.B-house) ‘white house’ as nominal compounds rather than as adjectives and nouns. In this section, I look at other aspects of PC compounds which are more attributive-like.

6.4.1. Frequency and productivity of PC compounding. Type 2 nominal compounding is productive with all PC nouns, all inchoative verbs (e.g., *kúxt’i* ‘get hard (i.e., bread)’, *nǒxki* ‘get stale’, *’ixki* ‘get sour’, *’ót’i* ‘get dry’, *t’áxki* ‘get white’, etc.), and all Class 1 PC verbs (especially with verbs such as *dǒtá* ‘be big’, *tx’úlo* ‘be small’, and *’rá’yo* ‘be new’, which are rather common). In contrast, the use of stative verbs is usually restricted to conventional existing combinations, and in some cases, it is not allowed at all. For example, verbs such as *pǐdi* ‘be thick’, *noho* ‘be fat’, *’ódi* ‘be rough (texture)’, *’bexi* ‘be dark’ are never found in compounds. Nevertheless, the combinatory power of stative verbs expressing colors is far richer than other members, as well as the verbs *nzá-tho* ‘be beautiful’ or *zéhwa* ‘be silly (for an animal)’. (See 6.4.4 for the status of acategorial members.)

Overall, in the light of all the possible word classes that can occur as dependent members in Type 2 compounds, compounding is not a structure that establishes a relevant contrast between a potential class of adjectives and other word classes in Otomi.³⁰

²⁹ *Atole* is a type of hot sweet drink made of corn flour.

³⁰ I owe this important observation to my first anonymous referee.

6.4.2. Adding up dependent members. Type 2 compounds may have more than one dependent member expressing a PC. This is shown in (69):

- (69a) (*dǒtá*-(*'bó-dē-thǒ*))
 be.big-black-grain_corn
 'big black corn grains'
- (69b) (*tx'úlo*-(*t'ǎx*-*'bǒ'ts'e-hmé*))
 be.small-be.white.B-basket-tortilla
 'small white tortilla basket'
- (69c) (*'rá'yo*-(*'bó-kářo*))
 be.new-black-car
 'new black car'

The examples in (69) display instances of potential combinations, which always include free combinatory members. However, such compounds are NEVER actual words in the language and are NEVER encountered in natural texts.

6.4.3. Modification of a PC lexeme in a compound. Like in other positions, PC lexemes in compounds may be modified by the intensifier *řá*- 'very'. This phenomenon is illustrated in (70) with an instance of the nominal predication construction, in which we have the Class 1 PC verb *dǒtá* 'be big' as a member of the compound noun:

- (70) *nú = ná = r* *ngǔ* $\emptyset = \text{řá-dǒtá-ngǔ}$
 DEF-DEM.SG.I=SG house 3.PRES=INT.PC-be.big-house
 'This house is a very big house'.

This property is somehow expected given that dependent members in such compounds commonly express a gradable property. Nonetheless, the intensifier cannot be used to modify an internal member in compounds with more than one dependent. This is shown by the ungrammaticality of (71a) in contrast to (71b), which represents a possible alternative:

- (71a) **nú ma dǒtá-řá-'bó-kářo*
 DEF 1POSS be.big-INT.PC-black-car
 Intended meaning: 'My big, very black car'.
- (71b) *nú ma řá-'bó-kářo* \emptyset $\emptyset = dǒtá$
 DEF 1POSS INT.PC-black-car REL 3.PRES=be.big
 'My big, very black car'.

TABLE 3

			Examples	Gloss	Translation
Productive	<i>'bó-</i>	'black'	(a) <i>'bó-mtxí</i>	(black-cat.B)	'black cat'
	<i>t'úlo-</i>	'small'	(b) <i>t'úlo-ngũ</i>	(small-house)	'small house'
Nonproductive	<i>dö-</i>	'big' ¹	(c) <i>dö-thé</i>	(big-water.B)	'river'
			(d) <i>dö-'ñú</i>	(big-path)	'dirt-road'
			(e) <i>dö-zã</i>	(big-tree)	'plane tree'
	<i>m'ó-</i>	'blue'	(f) <i>m'ó-hmé</i>	(blue-tortilla)	'blue tortilla'

¹The morpheme *dö-* does not appear to have an inherent or consistent tone. In general, the tone in these elements is rather inconsistent and appears to be idiosyncratic to each compound.

Similarly, while PC verbs may be modified by means of the delimitative morpheme =*tho* 'just', when the verbs appear in the compounds, the morpheme cannot be used. This phenomenon is illustrated in the contrast between (72a) and (72b):

(72a) *xi = ñ- 'ú-∅ = tho*
 (3.)PERF=ST-be.sweet-3OBJ=DEL
 'It's rather/just sweet'.

(72b) **ma 'ú = tho-t'ëi*
 1POSS sweet=DEL-atole
 Intended meaning: 'My rather/just sweet atole'.

6.4.4. Acategorical lexemes.³¹ Like other dialects, SI Otomi has a number of bound morphemes expressing PCs that can only occur in Type 2 nominal compounds. They form a closed class with very few members, presented in table 3. The first two elements in table 3 have a free combinatory power, while the other two are no longer morphologically productive, that is, they have lexicalized in the formations in which they occur. Nevertheless, an element such as *dö-* 'big', as in forms (c)–(e) in table 3, appears in a large number of compounds (actual words) in all dialects of Otomi, whereas the element *m'ó-* 'blue' occurs only in the example provided (f), at least for SI Otomi. This contrast suggests that *dö-* was a largely productive item in historical times.³²

³¹ I want to thank Paulette Levy for her help with the ideas in this section.

³² Two other members should perhaps be included in the list, at least for SI Otomi. The form *hõga* 'good' occurs in only one compound: *hõga-jö'í* (good-person) 'good person'. This form appears to be a fossilized bound form of the verb *hõgi* 'be ripe, be at its best point, be good', whose current bound form is *ho'*. The second form is *mbõnga*, a productive acategorical lexeme used to produce compounds expressing types of wild animals. This element is cognate with *mbãnga* in Mezquital Otomi (Hernández Cruz et al. 2004:159–60), which is the bound form of the PC noun *mbãngi* 'wild (animal)'.

A number of these elements still bear a synchronic (segmental) resemblance to counterpart PC verbs. This resemblance can be seen in (73), which suggests some degree of historical relationship:

(73a) *dö-* ‘big’ ~ *dötá* ‘be big’

(73b) *’bó-* ‘black’ ~ *pothi* ‘be black’; *po’ts’i* ‘be dark colored/navy blue’

(73c) *t’úlo-* ‘small’ ~ *tx’úlo* ‘be small’³³

The elements in (73a) and (73b) resemble the morphological root of the verbs in question (the rest of the morphological material in the verbs represents petrified derivational morphology). This pattern may reflect an old stage of the language in which verbal roots could act as monomorphemic verbs in nominal compounds, a pattern which became opaque later. At times, minimal compound pairs are found involving cognates. Relevant ones are presented in (74) and (75). In such formations, the compounds with the elements in table 3 are older formations:

(74a) *dö-thé* big-water.B ‘river’

(74b) *dötá-déhe* be.big-water ‘sea’

(75c) *dö-’nú* big-path ‘dirt-road’

(75d) *dötá-’nú* be.big-path ‘road’

Most crucially, the elements in question cannot occur in predicative position. This is shown by the impossibility of (76):

(76a) *ma* *míxi* *xi* = **ñ-’bó-Ø/m-pothi-Ø*
 1POSS cat (3.)PERF=ST-black-3OBJ/ST-be.black-3OBJ
 ‘My cat is black’.

(76b) *nǒ* = *r* *ngǔ* = *’á* *már* = **dö/dötá*
 DEF.SG=SG house=3SG 3.IMP.PC=big/be.big
 ‘That house was big’.

The ungrammaticality of examples like (76) illustrates that the elements in question are lexemes that can only work as dependents of head noun compounds. Hajek (2004), elaborating on the proposal in Dixon (2004), has suggested that attributive-only elements are perhaps the best place in the grammar of a language to find true adjectives. According to such a view, the elements in table 3 look as if they were the alleged true “adjectives” of Otomi. However, I have shown that the structural slot filled by these elements is not a modifier slot in attributive position but as a dependent member

³³ Both lexical elements serve as the probable source for the various allomorphs of the diminutive prefix {*tí-*} (i.e., *t’ú-*, *t’í-*, and *txí-*).

in a nominal compound. The fact that the bound elements in (75) are not found anywhere else in the system proves that the lexemes they represent do not belong to a well-defined word class in the language. For this reason, I have opted to treat such elements as “acategorical” lexemes.

6.5. Compounds in other dialects. Type 2 compounds, as described in the previous sections for SI Otomi, are also found in all Otomi dialects. Nevertheless, neither their scope nor their details are well known for some of the dialects. If lexical idiosyncrasies are set aside, one could conclude that Otomi variants have very similar characteristics to the ones described. The dialects display elements cognate to the ones presented with regard to stative verbs and acategorical lexemes. Voigtlander and Echegoyen (1985) for Sierra Otomi and Bartholomew (2004) for Mezquital Otomi are the best treatments of Otomi compounds currently available. Bartholomew (2004) gives a number of combinations with stative verbs (called “adjectives” in her approach) such as *t’áxi* ‘be white’, *théni* ‘be red’, and *mă* ‘be long’. From the choice of these verbs, I suspect that stative verbs expressing colors are the most productive members of the class for forming compounds, just like in SI Otomi. Compounding in Mezquital Otomi appears to be especially productive, as shown by the data in Wallis (1956) and Hernández Cruz et al. (2004); but it may also be the case that a similar degree of productiveness has yet to be discovered for other dialects.

7. Conclusion. In this article, I have studied how PCs are encoded in Otomi. I have shown in detail that such concepts are not encoded as adjectives but rather as verbs and nouns. There are also special nominal compounds in which lexemes denoting PCs occur as dependent members. My analysis is developed on the basis of data from the dialect of San Ildefonso Tultepec Otomi, but I have also used data from other dialects to show that the phenomena described pertains to Otomi as a dia-system. Unfortunately, very little is known about the lexical treatment of PCs in other Otopamean languages. As for other Otomanguean languages, the analyses in Smith-Stark (2004), Rojas (2004), and López Cruz (2004) point to the existence of a well-defined class of adjectives in various Zapotec languages, but many adjectives in such languages display a very close relationship to stative verbs. Outside Otomanguean, other languages where PCs are related to verbs are Cora (Uto-Aztecán) and Tarascan (Isolate). In Cora, like in Otomi, there are no adjectives (see Vázquez 1994; 2004). As for Tarascan, Capistrán (2004) shows that the vast majority of lexemes expressing PCs are deverbal forms from active verbs. The lexical treatment of PCs in languages of the Americas remains for the most part an unexplored topic. This study is a cordial invitation to other linguists working on indigenous languages to contribute their findings to this fascinating area of the grammar.

APPENDIX A
STATIVE VERBS IN SI OTOMI

Dimension	<i>hěts'i</i>	'be high/tall'	(Human)
	<i>xidi</i>	'be wide'	
	<i>ndó'yó</i>	'be thin (for an animal)'	
	<i>xíni</i>	'be thin (for a thing)'	
	<i>ts'út'i</i>	'be thin (for wood, thread)'	
	<i>pǐdi</i>	'be thick'	
	<i>ndǒ</i>	'be thick (for wood, thread)'	
	<i>mǎ</i>	'be long'	
	<i>noho</i>	'be fat'	(Human)
Physical property	<i>'ódi</i>	'be rough'	
	<i>hǔ</i>	'be heavy'	(Human)
	<i>pá</i>	'be hot'	(Human)
	<i>tsé</i>	'be cold'	(Human)
	<i>k'á</i>	'be wet'	
	<i>xǎ</i>	'be humid'	
	<i>'óni</i>	'be dry'	
	<i>tsé_tho</i>	'be insipid'	
	<i>'ú</i>	'be sweet'	
	<i>'úxi</i>	'be salty'	
	<i>'íxi</i>	'be sour'	
	<i>jú</i>	'be bitter'	
	<i>'í</i>	'be spicy/hot'	
	<i>kǔhi</i>	'be tasty/flavorful'	
	<i>túdi</i>	'be soft/weak'	
	<i>tsédi</i>	'be strong/firm'	(Human)
	<i>mě</i>	'be thick (i.e., for a liquid)'	
	<i>mě</i>	'be hard'	
	<i>nǒxi</i>	'be stale'	
	<i>'bexú</i>	'be dark'	
<i>ndü-pá</i>	'be lukewarm'		
<i>ndü-tse</i>	'be weak'	(Human)	
<i>xó</i>	'be stinky'		
<i>'ú</i>	'be painful, be in pain'		
<i>hěti</i>	'be noisy/talkative'	(Human)	
<i>řá_níhi</i>	'be fast, quick, light-footed'	(Human)	
Color	<i>pothi</i>	'be black'	
	<i>po'ts'i</i>	'be dark colored/navy blue'	
	<i>t'áxi</i>	'be white/silvery'	
	<i>théni</i>	'be red'	
	<i>k'áxt'i</i>	'be yellow'	
	<i>k'ángi</i>	'be blue/green'	
	<i>hwě'ts'i</i>	'be golden'	

	<i>káfé</i>	‘be brown’	
	<i>řósá</i>	‘be pink’	
	<i>gwéero</i>	‘be blond’	(Human)
Human Propensity	<i>ts’ó_tho</i>	‘be ugly/bad’	(Human)
	<i>nzá_tho</i>	‘be beautiful’	(Human)
	<i>zēhwa</i>	‘be silly (for an animal)’	
Evaluation	<i>hó</i>	‘be good’	(Human)
Value	<i>hmödi</i>	‘be expensive’	

REFERENCES

- ANDREWS, HENRIETTA. 1993. The Function of Verb Prefixes in Southwestern Otomí. Summer Institute of Linguistics and University of Texas at Arlington Publications in Linguistics, no. 115. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- BARTHOLOMEW, DORIS. 2004. Notas sobre la gramática. Diccionario del hñähñu (otomí) del Valle de Mezquital, Estado de Hidalgo, Vocabularios Indígenas, no. 45, ed. Luis Hernández Cruz, Moisés Victoria Torquemada, and Donaldo Sinclair Crawford, pp. 473–95. Mexico, D.F.: Instituto Lingüístico de Verano.
- BECK, DAVID. 1999. The typology of parts of speech systems: The markedness of adjectives. Ph.D. dissertation, University of Toronto.
- _____. 2000. The syntax, semantics, and typology of Upper Nexaca Totonac. *Linguistic Typology* 4:213–50.
- _____. 2002. *The Typology of Parts of Speech Systems: The Markedness of Adjectives*. New York: Routledge.
- CAPISTRÁN, ALEJANDRA. 2004. Words functioning as adjectives in Tarascan. Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- CHAFE, WALLACE. 2004. Do all languages have adjectives? Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- DIXON, R. M. W. 1982. Where have all the adjectives gone? *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax*, ed. R. M. W. Dixon, pp. 1–62. Berlin: Mouton.
- _____. 2004. Adjective classes in typological perspective. *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon and Alexandra Aikhenvald, pp. 1–49. Oxford: Oxford University Press.
- DIXON, R. M. W., AND ALEXANDRA AIKHENVALD, eds. 2004. *Adjective Classes: A Cross-Linguistic Typology*. Oxford: Oxford University Press.
- ECKER, LAWRENCE. 1952. *Compendio de gramática otomí: Introducción a un diccionario otomí-español*. Mexico, D.F.: Instituto Nacional de Antropología e Historia, Secretaría de Educación Pública.
- ENFIELD, N. J. 2004. Adjectives in Lao. *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon and Alexandra Aikhenvald, pp. 323–47. Oxford: Oxford University Press.
- FABB, NIGEL. 1998. Compounding. *The Handbook of Morphology*, ed. Andrew Spencer and Arnold M. Zwicky, pp. 66–83. Oxford: Blackwell.

- HAJEK, JOHN. 2004. Adjective classes: What can we conclude? *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon and Alexandra Aikhenvald, pp. 348–61. Oxford: Oxford University Press.
- HEKKING, EWALD. 1995. El otomí de Santiago Mexquititlán: Desplazamiento lingüístico, préstamos y cambios gramaticales. *Studies in Language and Language Use*, no. 17. Amsterdam: Institute for Functional Research into Language and Language Use.
- _____. 2002. Desplazamiento, pérdida y perspectivas para la revitalización del hñāñho. *Estudios de Cultura Otopame* 3:221–48.
- HEKKING, EWALD, AND SEVERIANO ANDRÉS DE JESÚS. 1984. *Gramática otomí*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- _____. 1989. *Diccionario español–otomí de Santiago Mexquititlán*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- HENGEVELD, KEES. 1992. *Non-Verbal Predication: Theory, Typology, and Diachrony*. Berlin: Mouton de Gruyter.
- HERNÁNDEZ CRUZ, LUIS; MOISÉS VICTORIA TORQUEMADA; AND DONALDO SINCLAIR CRAWFORD, eds. 2004. *Diccionario del hñāñhu (otomí) del Valle de Mezquital, Estado de Hidalgo*. *Vocabularios Indígenas*, no. 45. Mexico, D.F.: Instituto Lingüístico de Verano.
- HESS, H. HARWOOD. 1968. *The Syntactic Structure of Mezquital Otomi*. *Janua Linguarum, Series Practica*, no. 43. The Hague: Mouton.
- KRUSPE, NICOLE. 2004. Adjectives in Semelay. *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon and Alexandra Aikhenvald, pp. 283–305. Oxford: Oxford University Press.
- LASTRA, YOLANDA. 1992. *El otomí de Toluca*. Mexico, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- _____. 1997. *El otomí de Ixtenco*. Mexico, D.F.: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- LÓPEZ CRUZ, AUSENCIA. 2004. Adjectives in San Pablo Güilá Zapotec. Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- MARTIN, S. E. 1992. *A Reference Grammar of Korean*. Rutland and Tokyo: Tuttle.
- MITHUN, MARIANNE. 1991. Active/agentive case marking and its motivations. *Language* 67:510–46.
- PALANCAR, ENRIQUE L. 2004. Verbal morphology and prosody in Otomi. *IJAL* 70:251–79.
- _____. Forthcoming. Intransitivity and the origins of middle voice in Otomi. *Linguistics* 44, no. 3 (2006).
- ROJAS, ROSA MARÍA. 2004. *Hacia la categorización léxica de los conceptos de propiedad en el zapoteco de Santa Ana del Valle, Oaxaca*. Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- SMITH-STARK, THOMAS. 2004. Property concepts in San Baltasar Chichicapan Zapotec. Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- SMITH-STARK, THOMAS C., AND FERMÍN TAPIA GARCÍA. 2002. El amuzgo como lengua activa. *Del cora al maya yucateco: Estudios lingüísticos sobre algunas lenguas mexicanas*, ed. Paulette Levy. Mexico, D.F.: Universidad Nacional Autónoma de México.
- SOHN, HO-MIN. 2004. The adjective class in Korean. *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon and Alexandra Aikhenvald, pp. 223–41. Oxford: Oxford University Press.
- SPENCER, ANDREW. 1991. *Morphological Theory*. Oxford: Blackwell.

- Soustelle, Jacques. 1993. *La Familia Otomí Pame del México Central*. México, D.F.: Fondo para la Cultura y las Artes. [Spanish translation by Nilda Mercado of Jacques Soustelle, *La Famille Otomi-Pame du Mexique Central*. Paris: Institut D'Ethnologie, 1937.]
- SUÁREZ, JORGE A. 1983. *The Mesoamerican Indian Languages*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- VÁZQUEZ, VERÓNICA. 1994. Los conceptos de propiedad en cora: modificación, predicación y marcación de número. *Investigaciones Lingüísticas en Mesoamérica*, ed. Carolyn J. Mackay and Verónica Vázquez. Mexico, D.F.: Universidad Nacional Autónoma de México.
- _____. 2004. Property concepts in Cora (Uto-Aztecan). Paper presented at the annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), January 8–11, 2004, Boston.
- VEERMAN-LEICHSENRING, ANNETTE. 2000. *Gramática del chocho de Santa Catarina Ocotlán, Oaxaca*. Leiden/Mexico, D.F.: Universiteit Leiden/Universidad Nacional Autónoma de México.
- VOIGTLANDER, KATHERINE, AND ARTEMISA ECHEGOYEN. 1985. *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. Mexico, D.F.: Instituto Lingüístico de Verano. [First impression in 1979 without author names.]
- WALLIS, ETHEL EMELIA. 1956. *Diccionario Castellano–Otomí Otomí–Castellano*. Cuadernos del Valle de Mezquital, vol. 1, no. 1. (With the collaboration of Victor Manuel Arroyo.) Itzmiquilpan/Mexico: Patrimonio Indígena del Valle de Mezquital and Instituto Lingüístico de Verano.

Intransitivity and the origins of middle voice in Otomi*

ENRIQUE L. PALANCAR

Abstract

San Ildefonso Tultepec Otomi has three distinctive nasal morphemes that share a substantial number of features. The morphemes are the following: a middle marker, a detransitivizer, and a marker characterizing a closed subclass of intransitive verbs. In this article, I propose an analysis that relates these three morphemes diachronically, and links them to a common source. My analysis is able to account for both the commonalities and differences among the three morphemes. In the article, I claim that the middle marker in Otomi emerged from the reanalysis of an old inflectional morpheme characterizing a subclass of intransitive verbs depicting activities performed by humans. This analysis has important theoretical implications for the understanding of middle voice morphemes, given that, up until now, reflexives have been the only well known source for the emergence of middle markers.

1. Introduction

Otomi is an Otomanguean language spoken in Central Mexico that belongs to the Otopamean family. The article studies an interesting grammatical phenomenon found in the dialect spoken at San Ildefonso Tultepec, a large village in the state of Querétaro. I refer to this dialect as SI Otomi. Despite the great dialectal diversification of the language, the claims made for SI Otomi could be extended to other varieties.

In this article, I study the meaning and distribution of three verbal nasal morphemes in SI Otomi. The first nasal morpheme has been treated in Palancar (2004a) as a morphological marker of middle voice, for example, *m-pa't'i* 'warm oneself (i.e. by the fire)' < *pa't'i* 'heat something'. The second nasal morpheme derives intransitive verbs from transitive ones, and could be characterized as a "detransitivizer," for example, *m-peni* 'do the laundry' < *peni* 'wash clothes'. In broad terms, these two

morphemes could be said to be derivational.¹ In contrast, the third morpheme is inflectional, and characterizes a subclass of intransitive verbs, for example, *m=foge* ‘bark’ versus an open class of unmarked intransitives, for example, *tsudi* ‘hang’.

In the article, I first analyze these morphemes as independent morphological units in the language, and I give a substantial body of evidence to support such an analysis. Despite their difference, in my analysis I claim that the morphemes developed from a common source. The goal of the analysis is to show how they did it. The general picture looks as follows: I propose that there was first an ancestor morpheme for all three.² This old morpheme was somehow similar in function to the inflectional morpheme. From this common source, both the detransitivizer and the inflectional morpheme developed. The middle morpheme emerged in turn from the inflectional one.

The article is structured as follows. In the following section, I give a brief introduction to middle voice as a grammatical phenomenon. This section serves as theoretical background for Section 3, which introduces the middle morpheme described in Palancar (2004a). In Section 4, I introduce the other two nasal morphemes presented above and compare them to the middle morpheme. A summary of the facts observed in the language is given in Section 5. In Section 6, I present some comparative data from the dialect of Sierra Otomi, which are intended to clarify the origin of the middle morpheme in SI Otomi. The diachronic process that gave rise to the nasal morphemes is treated extensively in Section 7. The article concludes in Section 8, where I review the analysis presented in the study and advance the theoretical implications of the findings.

2. Middle voice

The study of middle voice stemmed from the need to characterize the special verbal morphology present in some old Indo-European languages. Lyons (1968: 373) proposes that middle voice applies when the “action or state affects the subject of the verb or his interests.” Examples in (1) from Greek illustrate the phenomenon (the first two examples are adapted from Kazenin [2001: 917]):³

- (1) a. *lú-omai*
wash-1SG.M (INDIC.PRES)
‘I wash myself.’
- b. *lú-ō*
wash-1SG.ACT (INDIC.PRES)
‘I wash.’

- c. *lú-omai* *khitōn-a*
 wash-1SG.M robe-ACC.SG
 'I wash the robe (for myself).'

The middle form of the verb in (1a) expresses a washing self-centered action that directly affects the subject. This form contrasts with the active form in (1b), which does not have this semantic implication and merely states that the speaker is involved in a washing activity that is directed by default to another entity (e.g. clothes). The middle form in (1a) may additionally take an overt direct object, as shown in (1c). In this case, the subject not only carries out the action of washing the shirt, but also indirectly benefits from it.

The literature on middle voice agrees that middle morphology applies to verbs rendering actions involving the subject (Klaiman 1992; Maldonado i.p.). Actions in middle voice commonly reflect a conceptual focus on the subject's own domain (Maldonado 1999; Nava and Maldonado 2005). This contrasts with active transitive voice that depicts situations where two individual and well-differentiated participants interact, most commonly as agent and patient. Kemmer (1993, 1994) has proposed a group of semantic situation types where middle marking is expected to happen crosslinguistically. Most of such types reflect actions that happen within the subject's own domain. Kemmer's situation types can be organized in the five main categories presented in Table 1, adapted from Shibatani

Table 1. *Middle situation types*

Body Action middles:	
Grooming actions	<i>afeitarse</i> 'shave'; <i>lavarse</i> 'wash'; <i>vestirse</i> 'dress'
Body function actions	<i>mearse</i> 'be pissing'; (mex.) <i>venirse</i> / (pen.) <i>correrse</i> 'come'
Change in body posture actions	<i>sentarse</i> 'sit down'; <i>tumbarse</i> 'lie down'; <i>arrodillarse</i> 'kneel'
Nontranslational motion actions	<i>estirarse</i> 'stretch out'; <i>reclinarse</i> 'lean back'
Translational motion actions	<i>irse</i> 'go away'; <i>marcharse</i> 'leave'; (mex.) <i>regresarse</i> 'come back'
Psychological middles:	
Emotion middles:	(mex.) <i>enojarse</i> / (pen.) <i>cabrearse</i> 'get angry'; <i>reirse</i> 'laugh'
Cognition middles:	<i>imaginarse</i> 'imagine'; <i>confundirse</i> 'be mistaken'
Indirect middles:	<i>conseguirse</i> / <i>agenciarse</i> / (pen.) <i>apañarse</i> 'get for oneself'
Reciprocal middles:	<i>entenderse</i> 'understand each other'; <i>pelearse</i> 'fight/quarrel'
Spontaneous middles:	<i>romperse</i> 'break'; <i>abrirse</i> 'open'; <i>desvencijarse</i> 'fall apart'

(2002). All types are illustrated with a few verbs in Spanish. The verbs are given in their infinitival form plus the marker *se*; a reflexive of third person that Maldonado (1992, 1999) analyzes as a middle morpheme.⁴

Table 1 shows the five major semantic realms where middle morphology happens across languages. These semantic realms are called “middle situation types” in Kemmer (1993). With the sole exception of spontaneous middles, all middle situation types comprise actions carried out by humans, at least in the default value. In general, middle marking prototypically occurs in verbs that render a physical or a psychological experience of the subject. In the following section, I show that Otomi has a nasal prefix that can be characterized as a morphological exponent of middle voice.⁵

3. Otomi middle morpheme

Palancar (2004a) has shown that Otomi has a marker of middle voice. This marker is a nasal morpheme that is prefixed to verbs. I refer to this morpheme as {*N-*}. Even though the data in this article come from SI Otomi, the pattern could be treated to some extent as diasystemic to the language. It may be observed in other dialects, such as Santiago Mexquititlán Otomi (Hekking and Andrés de Jesús 1984; Hekking 1995), Mezquital Otomi (Bernal Pérez 1986), Sierra Otomi (Voigtlander and Echegoyen 1985), Ixtenco Otomi (Lastra 1997, 1998), and the varieties spoken in Mexico State (Andrews 1993; Lastra 1989, 1992).

Before giving an illustration of the grammatical function of {*N-*}, it is convenient first to study its form. Examples in (2) show the three possible morphs realizing {*N-*}:⁶

- (2) a. *n-tsi* (in) ‘be afraid’
 b. *ñ-’ani* (in) ‘get informed’
 c. *m-pa’t’i* (in) ‘warm oneself (i.e. by the fire)’

The allomorphs of {*N-*} in (2) are motivated by a regular phonological conditioning: [n] in (2a) has the broadest distribution and is the default realization of the morpheme; [ɲ] (written *ñ*), as in (2a), appears before a glottal ([ʔ, h]); and [m] is used before a labial consonant ([p, ph, b]).

The analysis in Palancar (2004a) is based on a study of a sample of 72 middle verbs that show the nasal morpheme in (2). The semantics of these verbs reflect actions proper of middle voice. Attending to the middle situation types in Table 1, these 72 verbs show the following distribution: 31 verbs express body actions (43 percent); eight verbs psychological actions (11 percent); eight verbs reciprocal actions (11 percent); and two verbs

spontaneous actions (3 percent).⁷ Otomi does *not* use middle voice to express the area covered by indirect middles in other languages. Apart from these cases, 23 verbs in the sample do not directly conform to the types expressed, but express actions involving a human or animate subject, e.g. *m-’edi* (in) ‘get lost’ or *n-jut’i* ‘struggle to get free’. Among these 23 verbs, we find eight that express self-centered actions — for example, *m-pa’t’i* (in) ‘warm oneself’ in (2c) — which are typically encoded by reflexives in languages that have such morphemes.

As for morpholexical processes, 46 out of these 72 verbs are clearly deverbal, at least on morphological grounds. However, the semantics leading to the creation of the novel middle lexeme has become obscure for many of these 46 verbs. The remaining verbs (28 out of 72) represent cases where the derivational process is not transparent. The verbs in (2) are derived cases, as shown in (3):

	Source		Middle
(3) a.	<i>tsü</i> (in) ‘get frightened’	>	<i>n-tsü</i> (in) ‘be afraid’
b.	<i>’ani</i> (tr) ‘ask a question’	>	<i>ñ-’ani</i> (in) ‘get informed’
c.	<i>pa’t’i</i> (tr) ‘heat something’	>	<i>m-pa’t’i</i> (in) ‘warm oneself’

Examples in (3) show the lexical and morphological sources for the middle verbs in (2). The source verb in the derivation may be intransitive (as in [3a]) or transitive (as in [3b] and [3c]). Derivation from an intransitive applies to 30 percent of derived verbs in Palancar (2004a), while transitive sources represent the remaining 70 percent. In the derivational process, we observe three different meanings of {*N*-} that contribute to the creation of a middle verb. The first case is the “middle of intensity,” treated in Section 3.1; the second one is the “middle of internal energy,” shown in Section 3.2; and the third one is the “self-centered middle,” presented in Section 3.3. Despite cases of semantic transparency, middle derivation is no longer a productive morphological pattern in the language, as will be seen in Section 3.4. Because of that, many middle verbs have an unpredictable meaning. Other realms where {*N*-} is also found in the grammar of Otomi are treated briefly in Section 3.5.

3.1. *Middle of intensity*

The middle of intensity applies in the following circumstances. In most instances, the source verb expresses a punctual telic action. Here, the middle profiles the action of the source verb as imperfective, removing the telic feature. With this imperfective construal, the maintenance of a given state of affairs conveys a sense of repetition of the action, and it thus

brings about a nuance of intensity. The subject of the middle verb is now interpreted as being fully involved in the process. The verb *tsü* (in) ‘be/get frightened’ in (3a) may serve as an example. When *tsü* receives middle morphology, it surfaces as *n-tsü* (in) and means ‘be afraid’. Here the sensation of being frightened from the source verb is intensified. The transitive verb *beni* (tr) ‘remember something/think about something’ derives the middle *m-beni* (in) ‘ponder/reflect’. The act of remembering of the source verb is construed as a punctual mental experience where something comes to mind. The middle intensifies that experience and makes it unbound, leading to an interpretation as an act of thinking. Other examples may be seen in Palancar (2004a).

3.2. *Middle of internal energy*

A middle verb may focus on the internal energy of the subject while performing the action. This reading contrasts with the external perspective present in the active source verb. The verb ‘*ani* ‘ask a question’ in (3b) serves to illustrate this meaning. The verb derives the middle verb *ñ-’ani* ‘get informed’. The source verb here renders an action that is directed to another person. Haiman (1983: 803) calls such verbs “extroverted.” With the middle, the perspective is shifted to the inner world of the subject. The action is no longer portrayed as extroverted. The subject of the middle still requires information by asking, but the verb now emphasizes the subject’s own internal motivation to know about a certain topic. Another clear example of this meaning is the verb *fěhmi* (tr) ‘hit something with ample movements of the arms’. The verb refers to the hitting action typically effected on a carpet to remove its dust, or on sown beans to detach seeds from their dried sheaths. This verb serves as source for the middle verb *m-fěhmi* (in) ‘flutter wings’. The semantics of this middle verb expresses the nontranslational motion effected by some stationary wounded insects or nonflying birds. The iterative hitting motion present in the semantics of the source verb serves to construe the energetic internal motion displayed by these winged animals when they flutter their wings wanting to fly, but fail to do so.

3.3. *Self-centered middle*

The application of {*N-*} to a given extroverted verb may render an action performed on the self, that is, a self-centered action. The verb *pa’t’i* ‘heat something’ in (3c) may serve to illustrate this meaning. This verb is a

causative verb derived from the noun *pa* ‘heat’. Its derived middle *m-pa’t’i* ‘warm oneself’ has a clear self-centered meaning where the subject places himself close to a source of heat in order to get warm. The following verbs may additionally serve as further examples of this middle meaning. The source verbs *he’ts’e* (tr) ‘cover someone (e.g. with a blanket)’ and *xöki* (tr) ‘uncover someone (e.g. removing a blanket)’ derive the two middle verbs *ñ-he’ts’e* (in) and *n-txöki* (in), which now mean to put a blanket on or to remove it from oneself. Similarly, a source verb like *sü* (tr) ‘take care of somebody’ derives a verb *n-tsü* (in) ‘take care of oneself’, which now denotes a caring action oriented to the self.⁸

Prototypically, self-centered actions are expressed by reflexive markers in languages that have such morphemes. However, not all markers that express such notions should be taken to be reflexive markers.⁹ In Palancar (2004a), I have extensively argued that the nasal morpheme in these verbs is NOT a reflexive. There are several crucial facts in the language that support this analysis. For example, typical “direct reflexive types” like ‘see oneself’ or ‘hit oneself’ are conveyed by a transitive construction. This is shown in example (4) for the verb *handi* (tr) ‘look at/see’:

- (4) a. *dá=hand-Ø-a=ma hmi ka=r ñhe*
 1.PST=see-3OBJ-B=1POSS face LOC=SG mirror
 ‘I looked at myself in the mirror.’
 (Lit. ‘I looked at my face in the mirror.’)
- b. *dá=ñ-hand-a=ka=r ñhe*
 1.PST=M-see-B=LOC=SG mirror
 Intended reading: *‘I looked at myself in the mirror.’
 Actual reading: ‘I was facing the mirror.’

Example (4a) shows that a self-centered action, such as looking at oneself in a mirror, is constructed with the transitive construction. Alternatively, the active verb has a lexical correlate with the middle marker, as shown in (4b). However, this middle verb does not express the action of seeing something as self-centered. The verb means ‘face in a certain direction’ and not ‘see oneself’. The same applies to other “reflexive” hitting actions. Besides, only eight verbs in Palancar’s sample of 72 verbs reflect typical self-centered actions.

The expression of “indirect reflexive types” serves as another piece of evidence that {*N-*} is not a reflexive marker. Languages with reflexive pronouns commonly use these morphemes to express events denoting the genuine interest of the subject in performing the action, as in English *John built a house for ‘himself’*. To express this meaning, Otomi uses a special set of pronouns within a PP, but not {*N-*}. This is illustrated in (5):

sections. Another example would be the current meaning of the middle verb *n-jut'i* 'struggle to get free', which derives from the verb *jut'i* (tr) 'pull horizontally', where the source verb does not include any reference to the restraints to which the subject of the middle is subjected. A last example, and perhaps the most striking one, concerns the verb *pe'ts'i* (tr) 'possess'. This verb is used to render the possession of alienable things, and serves as the morphological base for the middle verb *m-pe'ts'i* (tr) 'mourn a deceased relative'. However, the language offers no clue as to what may have been the cultural bridge to link the semantics of these two verbs.

4. Middle verbs and intransitivity

I have pointed out in Section 3 that a source verb can be transitive or intransitive, despite the fact that transitive sources are more common (70 percent in a sample of 46 derived verbs). In contrast to the pattern in source verbs, middle verbs are overwhelmingly intransitive. Intransitive derived verbs represent 91.4 percent of all cases. The same tendency applies to middle verbs in general. Palancar (2004a) reports that 90 percent of middle verbs in his sample are intransitive. This is a highly significant percentage, which I take as relevant for the understanding of the origin of {*N*-}. As I pointed out in the beginning of the article, Otomi has another two nasal morphemes that are also associated with intransitivity.

4.1. *The middle morpheme and the detransitivizer*

A nasal morpheme similar in form to the middle marker appears in other grammatical contexts in the language. This morpheme derives intransitive verbs denoting activities from transitive ones that denote accomplishments. This is illustrated in (6):

- (6) a. *dí* = *peng-Ø-a* = *ma* *pahni*
 1.PRES=wash.clothes.B-3OBJ-B=1POSS shirt
 'I'm washing my shirt.'
- b. *dí* = *m-pen-i*
 1.PRES=DETR-wash.clothes.F-F
 'I'm doing my washing.'
- c. **dí* = *m-peng-Ø-a* = *ma* *pahni*
 1.PRES=DETR-wash.clothes.B-3OBJ-B=1POSS shirt
 Intended reading: 'I'm washing my shirt (for myself).'

In example (6a), we have the transitive verb *peni* ‘wash (clothes)’.¹² From this verb, an intransitive verb is derived, as seen in (6b), through the application of a yet another nasal morpheme, which is similar in form to the middle morpheme. The intransitive verb now denotes an activity. The ungrammaticality of example (6c) shows that the verb is intransitive. Additionally, (6c) also shows that this potential verb *cannot* become a middle verb similar to the Greek example in (1c).¹³

As illustrated by examples in (6), the main function of this nasal morpheme is derivative. In this light, I treat it as a “detransitivizer” morpheme. This morpheme occurs naturally with other transitive verbs such as *tai* ‘buy’, *pefi* ‘work’, *koki* ‘sweep’, *pe* ‘weave’, *pë* ‘steal’, *t’abi* ‘clean the cornfield’, *kox’t’i* ‘plough the cornfield’, etc., and it always produces an intransitive verb that renders an activity. In order to distinguish it from the middle marker {*N-*}, in this article I treat the “detransitivizer” as {*N*_{DETR-}}.

We can now look into the similarities of both {*N-*} and {*N*_{DETR-}} in more detail. On the one hand, {*N*_{DETR-}} behaves like middle {*N-*} on the morphological level, that is, it is a prefix and shows the same alternants in (2). On the other hand, given the fact that most middle verbs are intransitive, both morphemes could be seen to have the same function as ways of deriving intransitive verbs. In the light of this evidence, it would make sense to treat both {*N-*} and {*N*_{DETR-}} as the same morpheme, even though they may have slightly different functions. Agreeing with such a view, I argue in Section 7 that this must have been the case in historical times, and that both morphemes emerged from a common source.

However, I claim that the morphemes should be kept well apart in the grammar of synchronic Otomi. There are various reasons that support such an analysis. Even though a middle verb is overwhelmingly intransitive, we find verbs that are only transitive, such as *m-pe’ts’i* ‘mourn a deceased relative’, *n-kaxt’i* ‘rub with cream’, and *n-tsuti* ‘put earrings on’. Beside these cases, some other verbs may be transitive or intransitive, such as *m-fi’ts’i* ‘wear a piece of clothing over another one’, *ñ’eni* ‘play’, and *ñ-hüü* ‘lay eggs’. The transitivity of these verbs proves that {*N-*} is different from {*N*_{DETR-}}. On the other hand, while {*N*_{DETR-}} exclusively applies to transitive sources, as in (6), the middle may apply to an intransitive base to derive a middle verb, for example, the case of *tsü* (in) ‘be/get frightened’ and *n-tsü* (in) ‘be afraid’ in (2a) and of thirteen other cases in Palancar’s sample.

Furthermore, the different semantics displayed by the middle morpheme, as treated in the previous section, prove that {*N-*} is functionally independent from {*N*_{DETR-}}. A last argument in favor of treating these

two morphemes separately comes from the fact that many middle verbs no longer have a clear morphological source in the language, and thus function as fully deponent verbs being independent lexical items, for example, *ñ-’eni* (lab) ‘play’, *n-tso’mi* (in) ‘worry’, *n-sente* (in) ‘be sad’, etc.

4.2. Nasal intransitives

To complicate matters, Otomi has also a number of intransitive verbs that receive a nasal morph similar in form to $\{N-\}$ and $\{N_{\text{DETR}}-\}$. The application of the nasal shows a restriction concerning tense. The nasal is only used when the verb inflects for other tenses except the present tense (and the imperfect). This is illustrated in (7) and (8):

- (7) a. *hin-dí = gon-i*
 NEG-1.PRES=snore-F
 ‘I don’t snore.’
 b. **hin-dí = n = gon-i*
 NEG-1.PRES=N=snore-F
 Intended reading: ‘I don’t snore.’
- (8) a. *hin-dí = ne-Ø* *ga = n = gon-i*
 NEG-1.PRES=want-3OBJ 1.FUT=N=snore-F
 ‘I don’t want to snore.’
 b. **hin-dí = ne-Ø* *ga = gon-i*
 NEG-1.PRES=want-3OBJ 1.FUT=snore-F
 Intended reading: ‘I don’t want to snore.’

In example (7a), the verb *goni* (in) ‘snore’ is inflected in the present tense and appears without a nasal. When the verb inflects for tenses other than the present, the nasal is required, as shown in (8a) in contrast to the impossibility of (8b). These examples additionally show that the use of this nasal morph is not arbitrary. Even though it looks similar to both $\{N-\}$ and $\{N_{\text{DETR}}-\}$, it is necessary to keep them as distinct morphemes. There are various reasons that support this analysis.

On the one hand, the nasal in example (8a) is a clitic rather than a prefix. The morph may fluctuate between two phonological hosts. This is shown in (9), where || indicates a possible prosodic boundary:

- (9) a. *ga* || *n = gon-i*
 1.FUT N=snore-F
 ‘I’ll snore.’

- b. $ga = n$ || $gon-i$
 1.FUT= N snore-F
 'I'll snore.'

The morpheme *ga* in (9) expresses future tense and first person subject. Morphemes like *ga* are function words, which serve as exponents of tense/aspect/mode and of subject agreement. Such function words often occur as proclitics to verbs. However, when speakers hesitate, the function words receive stress and appear as nucleus of a phonological phrase, as shown in (9). The nasal morpheme we are studying in this section may be pronounced together with the verb, as in (9a), or with the function word, as in (9b). The former sequence is produced in careful speech, the latter in natural discourse. In contrast, both middle $\{N-\}$ and $\{N_{DETR}-\}$ are always prefixes. This is illustrated by the impossibility of both examples in (10):

- (10) a. $*no = r$ *mixi* $mí = n$ || *te'ts'e*
 DEF.SG=SG cat 3.IMP=M lick-F
 Intended meaning: 'The cat was licking herself.'
- b. $*dí = m$ || *pen-i*
 1.PRES=DETR wash.clothes-F
 Intended meaning: 'I'm doing my washing.'

The impossibility of prosodic pauses in the examples in (10) shows that neither $\{N-\}$ nor $\{N_{DETR}-\}$ behaves like the nasal in (9).

Besides this difference, the nasal morpheme in (9) also shows a relevant morphophonological difference relative to the prefixes. In example (2b) in Section 3, I showed that $\{N-\}$ is sensitive to the phonological nature of the first consonant in the verbal stem, and that it surfaces as a palatal nasal in contact with a glottal ([h,ʔ]). In contrast to the prefix, the nasal clitic in (9) does not commonly undergo this type of assimilation, and most of the times it occurs as [n], which is the default value of this morpheme.

We have also seen that the application of the nasal in (9) is restricted by tense. Consequently, it does not apply to the full paradigm of the intransitive verb. This contrasts with the distribution of $\{N-\}$ and $\{N_{DETR}-\}$, which do not show such a restriction. This behavior additionally suggests that the nasal clitic in (9) is inflectional rather than derivational. For convenience, I treat this new nasal clitic morpheme as $\{N_{INFL}=\}$ (or simply $N=$).

The morpheme $\{N_{INFL}=\}$ applies to a number of intransitive verbs in the language, but not to all of them. In this sense, the verbs that receive $\{N_{INFL}=\}$ form a closed lexical class of intransitive verbs. I call such class "nasal intransitives." Table 2 shows a non-exhaustive list of the members of this class:

Table 2. *Nasal intransitives in Otomi*

Body action:		
Grooming	(N=)' <i>öxi</i>	'have a hair-cut'
	(N=)' <i>e-ño</i>	'comb long hair' (<i>ño</i> 'head')
Body function	(N=)' <i>tsakni</i>	'spit (due to having much saliva)'
	(N=)' <i>gö-dä/ne/gü</i>	'be/become blind/dumb/deaf' ¹⁴
	(N=)' <i>goni</i>	'snore'
	(N=)' <i>hehe</i>	'sneeze'
	(N=)' <i>poho</i>	'defecate'
Translational motion	(N=)' <i>sagi</i>	'jump'
	(N=)' <i>tih</i>	'walk fast'
Psychological:		
Cognition	(N=)' <i>johya</i>	'be/get happy'
	(N=)' <i>pühni</i>	'forget/do wrong'
Emotion	(N=)' <i>seya</i>	'be/get angry'
	(N=)' <i>thede</i>	'laugh'

Miscellaneous:	(N=)' <i>ohni</i>	'cross oneself'
	(N=)' <i>tüxfani</i>	'be/get drunk'
	(N=)' <i>fox-ka-böhai</i>	'get stuck in mud' ¹⁵
	(N=)' <i>xädi</i>	'pray'
	(N=)' <i>peti</i>	'beg'
.....		
	(N=)' <i>pem'da</i>	'play music'
	(N=)' <i>ofo</i>	'write'
	(N=)' <i>wedi</i>	'sew'
	(N=)' <i>töte</i>	'win'
	(N=)' <i>föge</i>	'bark'
	(N=)' <i>zuni</i>	'sound'
	(N=)' <i>zö</i>	'burn'

The verbs in Table 2 have been classified according to the middle situation types presented in Table 1 in Section 1. Attending at least to the semantics of the middle types, the table shows many verbs that display prototypical middle semantics. They include verbs for grooming actions (*N=)*'*öxi* 'have hair cut' or (*N=)*'*e-ño* 'comb long hair' (lit. 'comb-head'); body functions such as (*N=)*'*poho* 'defecate' or (*N=)*'*tsakni* 'spit'; and other psychological functions such as (*N=)*'*pühni* 'forget/do wrong', (*N=)*'*johya* 'be/get happy', (*N=)*'*seya* 'be/get angry'.¹⁶ Apart from such verbs, we have other verbs that render a miscellany of actions, most of them performed by humans. There are some of them that could be construed as middle, such as (*N=)*'*ohni* 'cross oneself' or (*N=)*'*tüxfani* 'be/get drunk', but there are also a number of verbs here that do *not* present middle semantics in a clear way, for example, (*N=)*'*ofo* 'write', (*N=)*'*wedi* 'sew', or (*N=)*'*föge* 'bark'. The thin line in Table 2 indicates this phenomenon. The fact that middle verbs in the table surpass

Table 3. *A few intransitive verbs (of the open class)*

<i>hwet'i</i>	'be/get off (i.e. a light)'
<i>tsudi</i>	'hang'
<i>'ot'i</i>	'be/get dry'
<i>k'ots'i</i>	'come off the skin (thorn or splinter)'
<i>ju't'i</i>	'drip'
<i>kohi</i>	'stay, remain, be left over'
<i>hqi</i>	'fall (a mass)'
<i>hwangi</i>	'fall on the same plane (person)'
<i>tagi</i>	'fall on a different plane'
<i>yaxt'i</i>	'slid'
<i>thogi</i>	'pass by'
<i>tsqho</i>	'arrive here'
etc.	

nonmiddle verbs may be only accidental. In this way, the class of nasal intransitives does not only include verbs with middle semantics, which may aid to differentiate further $\{N-\}$ from $\{N_{INFL}=\}$.¹⁷

Nasal intransitives form a closed class that contrasts with the open class of unmarked intransitive verbs in the language. A few verbs of this open class are illustrated in Table 3.

Table 3 includes prototypical examples of unmarked intransitive verbs in SI Otomi. They show the following pattern, which is characteristic of the class: the verbs may either have animate or inanimate subjects. Besides, most of the verbs usually render punctual actions (e.g. the different 'falling' verbs, etc.). With the exception of a few middle and nasal intransitives that indicate manner of moving, most motion verbs in the language appear in this open class, especially those indicating the trajectory of the motion, for example, *thogi* 'pass by' or *tsqho* 'arrive here'. Similarly, most change of state verbs also occur in this class, and most of them have inanimate subjects, for example, *hwet'i* 'for a light to be/get off', *hqi* 'for a mass to fall', *'ot'i* 'be/get dry', etc.

5. Contrasting Otomi nasal morphemes

In the previous sections, I have shown that Otomi has three nasal morphemes: the middle prefix $\{N-\}$, the detransitivizer prefix $\{N_{DETR}-\}$, and the inflectional clitic $\{N_{INFL}=\}$ proper of the lexical subclass of nasal intransitives. I have presented convincing arguments both from their form and their meaning that such nasal morphemes should be kept apart in Otomi as representing three distinct morphemes. Despite their differences, their similarities are, nonetheless, remarkable.

In general, all of these morphemes are strongly associated with intransitive verbs. While {*N-*} occurs in some transitive formations, the vast majority of middle verbs in the language are intransitive. Similarly, {*N_{DETR-}*} and {*N_{INFL=}*} only occur with intransitives. While derivation is characteristic of {*N_{DETR-}*}, {*N-*} may also display a strong derivational function. And despite the fact that {*N_{INFL=}*} is not derivational, there are some nasal intransitives that can be transitive. In Table 2 above, the verbs (*N=*)*pūhni* ‘forget’, (*N=*)*xadi* ‘pray’, and (*N=*)*peti* ‘beg’ show this behavior. However, the verbs NEVER receive {*N_{INFL=}*} when they occur as transitive verbs.

Regarding semantics, the verbs having these morphemes show some relevant overlap. {*N-*} has been characterized as a middle morpheme occurring in verbs that display typical middle semantics. Many nasal intransitives also show middle semantics, even though this feature is not common to all of them. Two of the verbs in Table 1 even have middle correlates: (*N=*)*’ōxi* ‘have hair cut’ is related to *ñ-’ōxi* ‘cut one’s hair’; and (*N=*)*’eño* ‘comb long hair’ is related to *ñ-’eño* ‘comb short hair’. At times, subtle middle semantic nuances are observed in nasal intransitives, which are lacking in the large class of unmarked intransitives. For example, the verb (*N=*)*tsakni* ‘spit’ is a cognate of the verb *tsats’i* (in) ‘spit’. (*N=*)*tsakni* renders a spitting action where the subject has too much saliva in his mouth as due to a physical reaction. In contrast, in *tsats’i* the subject spits because he feels like it. On the other hand, some middle verbs with {*N-*} recall verbs with {*N_{DETR-}*}. Verbs derived with {*N_{DETR-}*} do not display middle semantics, but they behave like activities, that is, the action they refer to is construed as unbounded. An aspectual feature of imperfectivity may be observed in middle verbs having an intensive meaning, as treated in Section 3.1.

Finally, regarding formal aspects, all three morphemes share having a nasal as a distinctive feature. The morphs realizing {*N-*} and {*N_{DETR-}*} are homophonous, and despite the differences in distribution and behavior, {*N_{INFL=}*} occurs in similar contexts to the prefixes.

In the following sections, I explore the details of these phenomena in more depth. In Section 7, I propose that nasal intransitives are the existing remains of a larger class of verbs in the language. I argue that both nasal prefixes in the language, middle {*N-*} and detransitivizer {*N_{DETR-}*} emerged from a reanalysis of the ancestor of {*N_{INFL=}*}. Additionally, in this section I also give an account of the possible dynamics involved in the emergence of these functions.

However, before the diachrony of this process is presented in Section 7, the reader may find it useful or convenient to review first some relevant aspects of the verbal morphology of Sierra Otomi. This is done in the

next section with the intention of showing how a conservative Otomi system looks like. Sierra Otomi is the most conservative variant of the language (Yolanda Lastra pers. comm.). It still presents complex lexical classes present in the older dialect described by Cárceres (1907 [1580?]) in the sixteenth century.¹⁸

6. The nasal morpheme in Sierra Otomi verbs

According to Voigtlander and Echegoyen (1985), Sierra Otomi has four verbal classes that they call “conjugations.” The classes are established according to two main features: (a) verbs may receive a different set of function words;¹⁹ (b) function words may receive a nasal suffix according to the class and the syntactic valency of the verb. This nasal is the cognate of nasal morphemes in SI Otomi. The different classes are illustrated in (11), (here function words are not written as proclitics):

- (11) a. Conjugation I:
dí *pé'ts'i*
 1.PRES(I) put.away
 ‘I put it away.’
- b. Conjugation II:
dí-n *zófo*
 1.PRES(II) talk.to.him
 ‘I talk to him.’
- c. Conjugation III:
dí-dí *hóki*
 1.PRES(III) fix
 ‘I fix it.’
- d. Conjugation IV:
dí-dí-m *péfi*
 1.PRES(IV) work
 ‘I work.’

All verbs in (11) are inflected for first person present (indicative). Example (11a) shows that for a transitive verb of conjugation I, the function word for first person present is *dí*. In verbs of conjugation II, the function word receives a nasal suffix. As shown in (11c) and (11d), verbs belonging to conjugations III and IV are formed based on conjugations I and II, they simply add the suffix *-dí* to all persons in the present and the imperfect.²⁰ All conjugations, except conjugation IV, have both transitive and intransitive verbs. Besides, conjugation I and III have each two subclasses of intransitive verbs. The presence or absence of the nasal suffix is an important feature to distinguish conjugation I from II. However, its

Table 4. *Verbal classes in Sierra Otomi*

Conj.	Type of Verb	FW + Ø-V	FW + {-N}-V
I/III	tr.	all tenses	—
	in. subclass 1	all tenses	—
	in. subclass 2	present/imperfect	all other tenses
II	tr.	past	present/imperfect/future
	in.	—	all tenses
IV	in.	—	all tenses

application shows a number of contextual restrictions based on tense. The distribution of the nasal is presented in more detail in Table 4 (FW means “function word”).

Table 4 shows that all intransitive verbs of conjugation II and IV receive the nasal suffix in all tenses. This distribution recalls intransitive middle verbs in SI Otomi. Subclass 2 of intransitive verbs from conjugations I and III receive the nasal suffix when the verb is inflected in any tense other than the present or the imperfect. In this sense, verbs of this subclass remind one of nasal intransitives in SI Otomi. Unmarked intransitives of conjugations I and III would correspond to unmarked verbs in SI Otomi shown in Table 3.²¹

The different conjugation classes found in Sierra Otomi are arranged according to morphological features that are no longer semantically productive. Voigtlander and Echegoyen (1985) treat the conjugations as pure grammatical classes organizing the lexicon. In other words, the classes do not have semantic transparency. An exception to the pattern is found in conjugation IV, which still contains verbs expressing self-centered actions, reciprocal actions, as well as verbs the authors characterize as “actions as a whole.” The latter category reflects aspectual activities similar to the ones derived by {*N*_{DETR-}} in SI Otomi, as in example (6b).

The authors offer a list of 160 verbs to illustrate the conjugations. Out of these 160 verbs, 79 are marked with the nasal morpheme. However obscure the classes may have become in Sierra Otomi, a semantic pattern can be traced for most of the 79 verbs that receive the nasal morpheme. A clear pattern reveals itself when nasal verbs are contrasted with unmarked ones. I propose that the nasal morpheme characterizing the subclass of nasal intransitives in Sierra Otomi applied first to verbs depicting activities carried out by humans (or animates).²² Such a characterization still prevails when one compares the semantics of nasal intransitives with unmarked cases.

The overwhelming majority of unmarked verbs in Voigtlander and Echegoyen (1985) depict either spontaneous actions or telic motions. The same pattern is observed for the open class of intransitives in SI Otomi in

Table 3. Some examples of unmarked verbs are: 'yǎni 'wiggle'; 'wége 'get close to'; kóhi 'stay at'; póni 'exit'; 'yudi 'for a thing to peek'; gǔi 'stretch out'; jǔ't'i 'drip'; xógi 'open'; xót'i 'collapse'. Unmarked verbs contrast with nasal intransitives in that the latter depict activities, for example, kǔni 'grind corn'; něi 'dance'; toge 'ride a horse'; túhu 'sing'; 'wédi 'sew'; péni 'do the washing'; du 'carry, transport'; 'yánt'u 'be kidding'; zéte 'work for somebody else'; fúki 'do spiritual cleansings'; 'oxpu 'sleep at a given place at night'; hánni 'look from one side to the other'; háte 'cheat'. To these one may add the few transitive verbs of conjugation II: yǎfi 'chew'; nú 'look'; and wāpi 'clean the cornfield'.

In addition to the difference concerning aspectual semantics, many Sierra Otomi verbs with the nasal morpheme have middle semantics. This is true of all conjugations. All verbs showing middle semantics in Sierra Otomi appear in Table A in the Appendix. These middle verbs together with the ones that render activities constitute the vast majority of nasal verbs in Voigtlander and Echegoyen (1985).²³ Given that all these verbs are marked with a nasal morpheme, one could claim that the nasal attached to function words is a middle marker in Sierra Otomi. In the light of such a proposal, the only difference between Sierra Otomi and SI Otomi would lie in the fact that the marker in SI Otomi has functioned as a fully derivative morpheme producing middle verbs, whereas it seems that this was not the case for Sierra Otomi, at least attending to Voigtlander and Echegoyen (1985).

The data from Sierra Otomi that I have presented in this section reveal a previous frozen stage to the reanalysis I describe in the following section. The original subclass of nasal verbs in Sierra Otomi appears to have involved at first only intransitive verbs, and later extended to a few transitive verbs that eventually became conjugation II. The extension appears to have gone this direction given the very few verbs in conjugation II and the many restrictions this conjugation has on the application of the nasal marker. Verbs in Sierra Otomi preserved old lexical splits that are old to Otomi. They are found in the paradigm of certain verbs in Cárceres (1907 [1580?]). I propose that the lexical split found in Sierra Otomi and SI Otomi intransitives emerged from an old class of intransitive verbs that rendered imperfective actions performed by human subjects. The relevant diachronic process is described in the next section.

7. The origin of the nasal markers in SI Otomi

In this section, I propose that both {N-} and {N_{DETR}-} in Otomi emerged from the reanalysis of the direct predecessor of {N_{INFL}=}. Since all

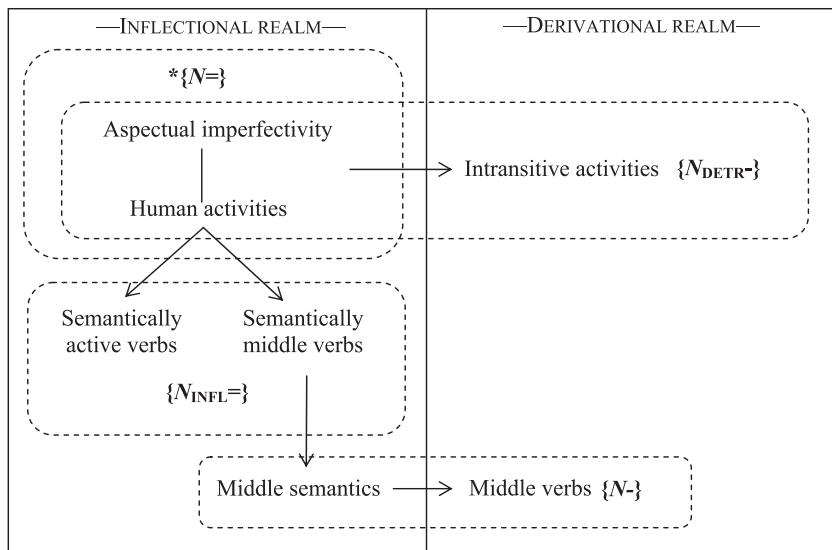


Figure 1. *The different outcomes of $*\{N=\}$*

morphemes emerged from this old source, I refer to it as $*\{N=\}$. The dynamics of this development is complex, and is summarized in Figure 1.

Figure 1 shows that Otomi had an old inflectional marker $*\{N=\}$ which was first associated with intransitive verbs that depicted imperfective nonpunctual actions typically performed by human beings.²⁴ The marker was then extended to other verbs expressing actions that pertained to the semantic realm of middle voice. At this early stage, however, the marker was not yet a middle morpheme. This is the situation inherited by $\{N_{INFL}=\}$ in SI Otomi and by all intransitive verbs marked with a nasal in Sierra Otomi.

Parallel to these semantic associations, the marker developed a derivative function, so that new verbs were created in the lexicon. On the one hand, the marker kept its original aspectual profile and became a detransitivizer producing intransitive activities from transitive accomplishments. This is the story of $\{N_{DETR}=\}$ as indicated in Figure 1, which depicts a morphological function still present in many dialects.

On the other hand, as the verbs depicting middle semantics grew in the lexicon, the marker was metonymically associated with this special semantics. Because of this association, it became a marker of middle voice, and came to express all the semantic idiosyncrasies that middle voice has in Otomi (see Sections 4.1 and 4.2). Parallel to the process described for the detransitivizer, the new exponent of middle voice also became a

derivative morpheme. This is the moment where the middle marker {*N-*} develops in SI Otomi. Even though {*N-*} extended to a few transitive verbs (e.g. the few transitive middles in SI Otomi that correlate with conjugation II in Sierra Otomi), the association with intransitivity remained a tight bond throughout the process, and new derived middle verbs remained within the intransitive realm.

7.1. *The stages of the process*

The whole diachronic process can be described in different stages:

- Stage 1. Here we find the predecessor of all the three nasal morphemes treated in the study. The marker was probably a clitic that had an inflectional function. At this stage, the marker served as a morphological exponent to distinguish a special subclass of imperfective intransitive verbs. The common association of the morphemes to intransitivity originates here. This nasal morpheme is still treated as an intransitive marker. The clitic was applied to verbs whose state of affairs denoted imperfective dynamics. In this way, the semantics of this proto-clitic was already associated with aspectual imperfectivity from the very early stages in the dynamics.²⁵
- Stage 2. The language acquired a “detransitivizer” (which corresponds to synchronic {*N_{DETR-}*}) once the old marker developed a function to create lexical contrasts based on aspectual imperfectivity. In this way, {*N_{DETR-}*} is a direct inheritor of the aspectual semantics present in the ancestor. {*N_{DETR-}*} derives intransitive verbs denoting activities by removing the telicity implicit in the actions depicted by the source verbs. On the other hand, the emergence of {*N_{DETR-}*} is clearly independent of what was going to happen in other stages. In other words, its emergence is independent of both {*N_{INFL=}*} and middle {*N-*}. Nevertheless, I suggest that once the process of its creation had been started, the existence of this morpheme served as a powerful motivation for the emergence of the middle marker as a fully derivative morpheme.
- Stage 3. Independently of Stage 2, the old nasal morpheme applies to a greater number of semantically middle verbs, and it becomes strongly associated with the semantics of such verbs via metonymy. Consequently, the morpheme is gradually viewed by speakers as an exponent of middle semantics.
- Stage 4. The situation described in Stage 3 could have remained like this for a long time because it represents a stable state of affairs. Nevertheless, for some unknown reason, either the marker was

applied to more verbs having middle semantics, or its association with middle semantics became stronger. As a direct consequence of such associations, the morpheme was reanalyzed as a middle marker. This change may have been very sudden (see Section 7.2). The marker was still nonderivative at this stage. I suggest this is the stage where we find Sierra Otomi verbs, but a similar description applies to the group of nasal intransitives in SI Otomi (see Table 2 in Section 4.2). The dotted line in Table 2 indicates a crucial inflectional point in the process.

- Stage 5. The association of $\ast\{N=\}$ with semantic middle becomes a definite one. $\ast\{N=\}$ is reanalyzed at this stage as a morpheme used to create lexical contrasts based on middle semantics, and hence starts to develop a derivative function. Once this new function grammaticalizes, we obtain $\{N-\}$ as a derivative middle morpheme.

My analysis accounts for the formal and functional similarities observed between the nasal morphemes treated in this study. It explains why the morphs realizing $\{N-\}$ and $\{N_{\text{DETR}}-\}$ are formally similar. Through the reanalysis, the morphemes became prefixes. In other words, $\ast\{N=\}$ was first a clitic that enjoyed the relative freedom of alternating between its two hosts, namely, the verbal stem and the function word. As the derivation was in progress, the morpheme became closer to the verb as the exponent of lexical meaning. Because of this, the morphs realizing the morpheme became more bound units and underwent regular phonological adjustments with the verbal stem (e.g. the prefixes in SI Otomi undergo regular palatalization of the nasal before a glottal, while the clitic does not). Due to their distribution, the morphs became prefixes. In contrast, $\{N_{\text{INFL}}=\}$ remained a clitic in SI Otomi nasal intransitives, and apparently became a suffix to function words in Sierra Otomi.

On the functional and semantic level, the analysis also accounts for the similarities found among the three morphemes. Intransitivity is a semantic and distributional feature linking all the three morphemes due to their common source in $\ast\{N=\}$. While $\{N_{\text{DETR}}-\}$ does not have middle semantics, the map explains the overlapping middle semantics found between $\{N-\}$ and $\{N_{\text{INFL}}=\}$, treated in Sections 4.2 and 5. It also explains why certain middle verbs having an intensive meaning (see Section 3.1) portray the actions as imperfective or unbound. This semantic feature relates middles to verbs derived with $\{N_{\text{DETR}}-\}$. I propose that the nuance of intensity in such middle verbs emerges from the imperfective construal inherited from the ancestor. Like $\{N_{\text{DETR}}-\}$, $\{N-\}$ removes the telicity implicit in the source verbs, for example, *'öni* (in) 'hide' (punctual) > *ñ-'öni* 'hide away' (unbound ~ intensive).

On the other hand, the dynamics also illustrate why we find an overlap between $\{N_{\text{DETR-}}\}$ and $\{N_{\text{INFL=}}\}$. As treated in Section 5, some intransitive verbs with $\{N_{\text{INFL=}}\}$ function as labile verbs, for example, ($N=$)*pūhni* ‘forget’, ($N=$)*xadi* ‘pray’; etc. In this sense, the semantics of such labile verbs resemble the activity verbs that $\{N_{\text{DETR-}}\}$ prototypically derives (e.g. *tai* ‘buy’, *pefi* ‘work’), making $\{N_{\text{INFL=}}\}$ look like $\{N_{\text{DETR-}}\}$. From data like these, it seems that the derivational function emerged from a reanalysis of labile cases. As the nasal was extended to all tenses, speakers were confronted with two verbal stems with similar referential meaning but with a different syntactic profile. As the nasal was the marked case, they reanalyzed it as a derivative morpheme and extended it to other verbs.

7.2. *Towards a formal representation of the dynamics*

The emergence of the middle function did not happen smoothly. Instead, such a process must have taken place in an abrupt manner, which could be characterized as “catastrophic” in the sense used in “catastrophe theory” by Thom (1972).²⁶ Thom’s theory studied sudden changes in natural dynamic systems, and a “catastrophe” is the generic term in his theory for any sudden qualitative change. The theory provided a complex topological model to account for such changes in the form of various “catastrophic maps.” The complex dynamics involved in the process I described in the previous sections could have been defined by a set of conditions or parameters, as seen in (12):

- (12) a. Semantic parameter:
 aspectual imperfectivity ↔ middle voice semantics
 b. Functional parameter:
 intransitive marker ↔ middle marker
 c. Morpholexical parameter:
 inflectional marker ↔ derivative marker

The parameters in (12) interact in a complex way giving way to Stages 1, 3, 4, and 5 presented in the previous section. To visualize the dynamics created by these parameters, I have used a topological map known as the “cuspid catastrophe,” which portrays a three-dimensional space regulated by three parameters. For this purpose, consider Figure 2.

The topological map in Figure 2 is intended as a way to capture the dynamics of the origin of the middle morpheme in SI Otomi. The map represents a vast semantic space. The numbers in the figure represent the relevant stages in the previous section. These stages are represented as

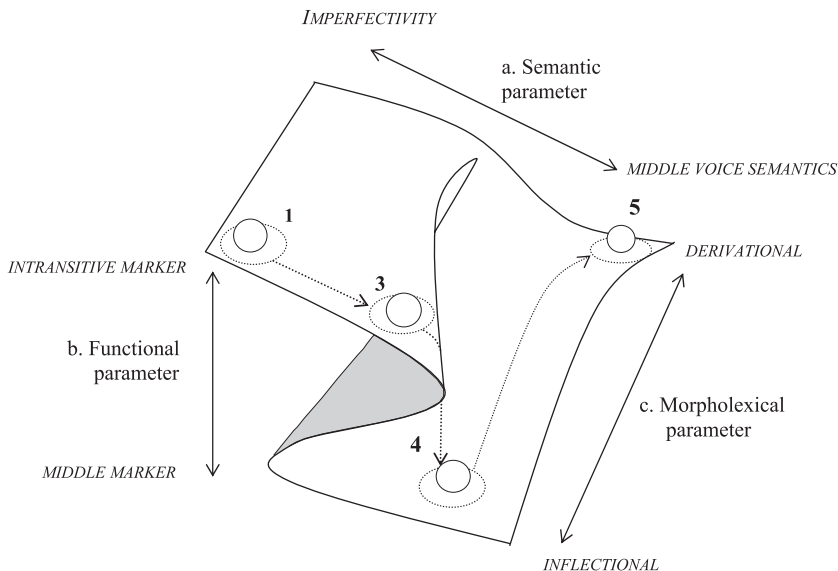


Figure 2. *The origin of the middle marker in SI Otomi*

balls in the map. One should want to visualize the balls rolling to different places in the rugged plain as if they were linguistic entities moving to different semantic spaces, as they acquire different functions. The map includes a leap from Stage 3 to Stage 4. This leap represents an abrupt change as the marker is reanalyzed in the minds of speakers as a middle marker. Middle voice derivation is no longer productive in SI Otomi (see Section 3.4), and the language has many deponent middle verbs for which a morphological source is now difficult to discern. This implies that the diachrony, once it reached Stage 5, became an unproductive lexical mechanism.

8. Concluding remarks: theoretical implications

In this article, I have presented a description of three nasal morphemes present in SI Otomi. The morphemes function as distinctive grammatical units in the language, but share a substantial number of features. The morphemes under study were the following: the middle marker ($\{N_{MID}-\}$), the detransitivizer ($\{N_{DETR}-\}$), and a marker characterizing a closed subclass of intransitive verbs ($\{N_{INFL}=\}$).

I have advanced an analysis that relates these three morphemes diachronically, and links them to a common source, treated as $*\{N\}$. This

old morpheme was similar in function to synchronic $\{N_{\text{INFL}}=\}$, but applied to intransitive verbs depicting human activities. In my analysis, I have also provided a model for the diachronic process at work. The type of analysis I give in this article is able to account for both the commonalities and differences among the three morphemes. The similarities were motivated by the common origin of the morphemes, while their divergences emerged as the three entities gained their own independent semantic and functional space in the system.

The Otomi case I presented in this article proves that middle markers can emerge from sources other than reflexives. This should have important implications for our understanding of middle markers. Up to now, reflexive pronouns have been the only well known sources for the emergence of middle markers. Most of our data in this respect come from well-documented Indo-European languages, even though the case may be seen in other nonrelated families. Examples of the phenomenon are seen for instance in Old Norse, where middle *-sk* developed from an older independent reflexive *sik*; in the Russian middle *-s'/-sja*, which emerged from the predecessor of the reflexive *sebja* (for more details, see Klaiman 1992 and Kemmer 1993). As I have shown, this was never the case for Otomi, given that the middle marker in the language did not developed formally from a reflexive marker $\{N\}$, nor from its marginal reflexive meaning.

Similarly, Kemmer (1993, 1994) proposes an interesting conceptual cline that motivates the emergence of such middle markers developing from reflexives. The cline is based on the principle of “distinguishability of participants.” Verbs denoting killing or hitting actions have clearly distinguished agent and patient entities. In contrast, this is not the case with verbs denoting actions that the subject directs to a part of his own body, that is, grooming actions. Middle voice is prototypically associated with the latter type of verbs.

In the morphological domain, reflexives undergo phonological attrition as they become associated with verbs depicting middle semantics. In this type of system, verbs with reflexives are often transitive while verbs with middle markers often are not. This may be seen in Norwegian (Bokmål) in (13):

- (13) a. *Peter se-r seg selv*
 Peter see-PRES 3REF.I 3.REF.II
 ‘Peter sees himself.’
- b. *Peter vaske-r seg*
 Peter wash-PRES 3REF.I
 ‘Peter washes (himself).’

- c. *Peter undre-r seg*
 Peter wonder-PRES 3REF.I
 ‘Peter is surprised.’
- d. *Peter undre-s på, om det er hend-t noe*
 Peter wonder-M on if it is happen-P.P something.N
gal-t
 crazy-N
 ‘Peter wonders if something wrong has happened.’
 (Haugen 1965)
- e. *Peter trive-s på fest-en*
 Peter enjoy-M in party-DEF.C
 ‘Peter enjoys himself at the party.’

The language has both a middle marker and reflexive pronouns. The reflexive pronoun illustrated in the examples in (13) is *seg* (an anaphoric for third person), which may receive additional marking as in (13a) by means of the morph *selv* ‘self’. The language uses the double reflexive in reflexive situations where there is high distinguishability of participants, as for example in (13a).²⁷ As the situation approaches the middle realm, a single reflexive is then employed, as shown in (13b). Notice that the verbs here are transitive, and the reflexive works as a syntactic object. A number of cognition middle verbs show an intermediate position, and may receive the single reflexive in (13c), or the middle *-s* in (13d). However, other verbs can only receive the middle, as seen in (13e). Verbs with the middle morpheme are typically intransitive.

Norwegian is one language that shows a clear example of a continuum of reflexive-middle markedness proposed for Swedish by Tohno (1999) and Sundman (1987) and employed by Shibatani (2002) from a more typological perspective. Such a continuum inherits the ideology from Kemmer’s analysis of middle voice based on the distinguishability of participants. In the scenario, transitivity (as realized and inherited by reflexive markers) is seen as the source from which the intransitivity of both middle morphemes and middle verbs develops. This scenario accounts for phenomena found in languages where the middle marker clearly develops from a reflexive.

However, the model does not account for cases like Otomi middle verbs. As I have shown in this article, both the marking of middle voice and the emergence of a middle morpheme in the language developed directly from the grammatical realm of intransitivity, and from here, it extended to a few transitive verbs.

The case of Otomi middle voice represents an interesting challenge for linguistic theory on middle voice, and poses the question of how

common it might turn out to be in other languages, especially in Mesoamerican languages. The research in this area in languages other than Indo-European is relatively recent in current linguistic theory, and for this reason, the finding and proper study of middle voice phenomena in indigenous languages of Mexico is still rather new. Consequently, much of this area still lies under cover. Middle voice marking has been reported for Uto-Aztecan in Pima by Estrada Fernández (2005), and for Tarascan by Nava and Maldonado (2005). The latter authors show that Tarascan middle voice is a basic nonderivative voice, and that the multiple middle morphemes in the language cannot be said to have developed either from reflexives or from reflexive scenarios. A somewhat similar phenomenon is found in Pima, where the middle marker evolved from a verbal morpheme expressing a generic third person singular. Nevertheless, we still do not know at this stage how extensive the pattern is in other families, even though the phenomenon may be more common than it was ever thought to be.

Received 3 July 2003
 Revised version received
 26 April 2004

Universidad Autónoma de Querétaro

Appendix

Table A. *Middle verbs in Sierra Otomi*

Situation type			Conj.
Body action:			
Grooming	<i>hi</i>	(in) 'bathe'	I.2
	<i>xúda</i>	(in) 'wash one's face'	I.2
	<i>hógi</i>	(in) 'get dressed'	IV
Body posture	<i>'bái</i>	(in) 'stand'	I.2
	<i>'bēni</i>	(in) 'lie'	I.2
	<i>jwāni</i>	(in) 'straighten up'	III.2
Motion:	<i>'yo</i>	(in) 'walk'	I.2
	<i>dūnni</i>	(in) 'roll'	III.2
	<i>sági</i>	(in) 'hop'	I.2 // IV
	<i>tíhi</i>	(in) 'run'	I.2
	<i>tíhni</i>	(in) 'go running'	III.2
	<i>xáha</i>	(in) 'swim'	IV
Psychological:			
Cognition	<i>jóed-hya</i>	(in) 'be happy'	III.2
	<i>bídi</i>	(in) 'be/get frightened'	II
	<i>su</i>	(in) 'be afraid'	II

Table A (Continued)

Situation type			Conj.
	<i>kwe</i>	(in) 'be angry'	II
	<i>'wĩ</i>	(in) 'dream'	II
	<i>'yomfěni</i>	(in) 'think'	IV
	<i>sadi</i>	(in) 'learn'	IV
	<i>běni</i>	(tr) 'remember'	II
	<i>numánho</i>	(tr) 'please'	II
	<i>nde</i>	(tr) 'want'	II
Emotion	<i>'yúya</i>	(in) 'be miserable'	I.2
	<i>'uni</i>	(in) 'suffer'	IV
Miscellany:	<i>'bú</i>	(in) 'live'	I.2
	<i>zǎi</i>	(in) 'accustom'	II
	<i>'áhā</i>	(in) 'sleep'	I.2
	<i>sāya</i>	(in) 'rest'	I.2
	<i>zoni</i>	(in) 'weep'	I.2
	<i>hě</i>	(in) 'be quiet'	II
	<i>hěni</i>	(in) 'be quiet'	IV
	<i>'bėdi</i>	(in) 'get lost'	IV
	<i>yani</i>	(tr) 'defend'	II
Reciprocal:	<i>hwěki</i>	(in) 'have pity of'	IV
	<i>fats'i</i>	(in) 'help each other'	IV
	<i>poni</i>	(in) 'exchange'	IV
	<i>súi</i>	(in) 'fight each other'	IV
	<i>tóhni</i>	(in) 'push each other'	IV
	<i>thāti</i>	(in) 'marry'	IV
	<i>zengwa</i>	(in) 'greet each other'	IV

Notes

- * The present study was financed by the Universidad Autónoma de Querétaro and by the CONACYT (project no. 47475). I benefited greatly from Scott T. Berthiaume's insightful comments on an earlier version of this paper. I am heartily grateful to him for his help, patience, and time. I also want to thank Ricardo Maldonado for his wise comments and for providing the inspiration to dig for middle voice marking in Otomi and other related phenomena, and for making me see that middle voice is a complex area of meaning that can be approached through unexpected paths. My thanks to my two anonymous referees and to Paulette Levy, who helped me understand some aspectual details in the dynamics I propose. All errors and deficiencies remain my own responsibility. Correspondence address: Facultad de Lenguas y Letras, Universidad Autónoma de Querétaro, Cerro de las Campanas s/n, Querétaro 76010, Mexico. E-mail: palancar@uaq.mx or epalancar@hotmail.com.

1. See the description in Section 3 for more details.
2. Otomi has up to eight nasal morphemes depending on the dialect. I study here only three of them, which I treat as diachronically related. The other five morphemes are: an exponent of imperfective tenses; an exponent of some nonpresent stems; an

exponent of some impersonal stems; a stative marker in participles and stative verbs, which denote property concepts; and a denominal prefix which derives intransitive verbs. See Palancar (2004b) for more details.

3. Abbreviations: ‘-’ morpheme boundary; ‘.’ cumulative morpheme; ‘=’ clitic; ACT: active; B: bound form; C: common gender; Conj: conjugation; DEF: definite; DETR: detransitivizer; F: free form; FOC: focus; FUT: future; IMP: imperfect; IND: indefinite; INDIC: indicative; INFL: inflectional; in: intransitive; LOC: locative preposition; M: middle; mex: Mexican Spanish; N: neuter gender; NEG: negative; pen: Peninsular Spanish; POSS: possessive; PRES: present; PST: past; REF: reflexive (pronoun); S: suffixal form; SG: singular; tr: transitive.
4. Not present in Shibatani, the table further includes the categories of “body function actions” in Maldonado (1999), and “reciprocal actions” in Kemmer (1993).
5. Those interested in knowing more about middle voice may consult Kazenin (2001) for an extensive review of the phenomenon, or other well known sources such as Givón (1981), Shibatani (1985), Klaiman (1988, 1991, 1992), Fox and Hopper (1994), and Kemmer (1993, 1994).
6. In the transcription of Otomi, I have used the orthography suggested by Hekking (2002). Deviations from the IPA convention: ‘/?’; f/p^h; j/k^h; y/jj; x/ʃ; tx/tʃ; ñ/ɲ; r/ɾ; a/a; a/ɔ; e/ɛ; o/ɔ; u/i. Umlaut indicates nasalization (i.e. ä/ã; ü/ü, etc.). For SI Otomi, lexical tone is only represented here in function words.
7. Middle morphology has become obligatory for verbs rendering reciprocal actions. Most such verbs additionally receive dual morphology as a way to mark the involvement of the other party in the action. Verbs like *n-thōti* (in) ‘marry’ or *n-k’ō’ti* (in) ‘see each other’ could serve as examples of reciprocal verbs. Both verbs derive from transparent sources, namely: *thōti* (tr) ‘answer someone’ and *k’ō’ti* (tr) ‘see someone’. While middle is required for reciprocals, Otomi does not make extensive use of {N-} in verbs rendering spontaneous actions. There are very scanty instances of such an extension in the language; one of them is *n-tūxni* (in) ‘be/get entangled’ from *tūxni* (tr) ‘entangle’ (see Palancar [2004] for more details).
8. The middle verbs *n-txōki* [ntʃō^hki] (from *xōki* [ʃō^hki]) and *n-tsū* [ntsū] (from *sū* [sū]) reflect the following morphophonological adjustment before {N-}: [ʃ, s > tʃ, ts / n_____].
9. The linguistic literature makes at times a confusing use of the term “reflexive.” Kazenin (2001) uses this term to refer to any type of morphology that may express a “direct reflexive.” He distinguishes two types of “verbal reflexives”: the “derived one,” which emerges historically from an independent reflexive pronoun (e.g. Old Norse *-sk* stemming from an older pronoun **sik*) and the “underived type” like the Greek example in (1). In view of these categories, I argue that Otomi is of the second type. However, I do not see the reason why one would want to call {N-} a “verbal reflexive” if the language employs this meaning very marginally.
10. This pronoun cross-references a participant already established in the clause.
11. The language has zero object anaphora, but agreement object markers for first and second person are suffixal and always have been — cf. Cárceres (1907 [1580?]), e.g. SI Otomi *Ø=tōn-kagi* [3.PRES=win.S-1OBJ] ‘he beats me (as in a game)’.
12. The form *penga* in (6a) is the prosodically bound form of the verb. See Palancar (2004c) for more details.
13. Roberto Zavala (pers. comm.) has suggested to me that I should treat this morpheme as an “antipassive” morpheme. I think this is an insightful suggestion to be taken into account for the future, but it still requires further study.
14. *də* means ‘eye’; *ne* ‘mouth’; *gü* ‘ear’.
15. *ka* is a general locative preposition and *bəhai* means ‘earth, mud’.

16. Furthermore, under the miscellaneous category we have the verb (*N=*)'ohni 'cross oneself' which renders a middle action commonly encoded in other languages with a reflexive pronoun.
17. Ricardo Maldonado (pers. comm.) has suggested to me that "nasal intransitives" are all middle verbs, and that this class may be an ancient class depicting middle voice. However, I have preferred not to treat all these verbs as having middle semantics because I could risk making the discussion circular.
18. In this article, I do not use such historical sources because they do not show clear data on the phenomenon I study here.
19. As in SI Otomi, such function words are often proclitics, but in Sierra Otomi they are written as independent phonological words. In other dialects, similar function words appear to have become prefixes, as for example in the varieties of Mexico State (Andrews 1993; Lastra 1989, 1992).
20. The suffix is accidentally similar here to the function form, but it does not represent reduplication.
21. In contrast to Sierra Otomi or Ixtenco Otomi (Lastra 1997), SI Otomi does not have the subclass of verbs with the suffix *-dí*. Consequently, there is no distinction of conjugations I and III, or II and IV for that matter.
22. According to Merlan (1985), this appears to be a fundamental pattern in any intransitive split.
23. The rest of the 79 verbs, not mentioned in the discussion, are given here: the nasal applies to verbs rendering (verbal) communicative acts by humans, for example the intransitives *yä* or *mā* 'talk'; *yähni* 'for everybody to talk at once'; or the transitive verbs *zōfo* 'talk to someone' and *zōhni* 'call somebody'. These verbs are related semantically to a group of verbs depicting voice sounds but not verbal communication, such as *mbáhnì* 'shout' or *mbáfi* 'scream', as well as other acoustic verbs such as *ts'ixni* 'creak' and *zūni* 'toll'. The nasal is also found in a few verbs denoting internal change such as *te* 'grow'; *ja* 'happen'; *pádi* 'transform'; in a few spontaneous verbs such as *zō* 'burn' and *tōni* 'break in two', and in *zūdi*, a verb denoting the state of hanging. Apart from such verbs, the nasal is also found in a number of verbs denoting a routine activity, such as *gúrbi* 'be leader'; *hmū* 'be chief'; *pā* 'be a seller'; *tāi* 'be a buyer'; and *xāmbāte* 'be a teacher'. The last verb in the list is the transitive *yá'ts'i* 'peel', which has no middle meaning at all.
24. It is possible that the imperfective semantic feature is the oldest of all features in the proto-morpheme. All dialects of Otomi have a nasal prefix attached to a given paradigm of morphemes expressing person and tense in order to convey imperfective nuances. In SI Otomi, for example, the function word *dí* is a cumulative exponent of first person and present tense, as in *dí xōni* 'I am in a hurry' (lit. 'I hasten'). The imperfect tense for first person is encoded by means of this function word plus a nasal prefix, as in *n-dí xōni* 'I was in a hurry' (lit. 'I was hastening').
25. This raises the question of whether the marker had any other grammatical meaning besides a classificatory one before receiving this semantics. Such a question is difficult to answer because it raises the open question of the origin of the marker before this ancient grammaticalization stage.
26. "Catastrophe theory" has been applied to semantic analysis by Wildgen (1982) and Bernárdez (1994). The cuspid catastrophe — presented in Figures 2 and 3 — has been used elsewhere in Palancar (1996, 1999, 2002).
27. An anonymous referee pointed out that situations like (13a) depict just one participant, but playing two distinct different roles. This is a distinction not clearly made in Kemmer (1993).

References

- Andrews, Henrietta (1993). *The Function of Verb Prefixes in Southwestern Otomí*. The Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 115. Dallas, TX: SIL and the University of Texas at Arlington.
- Bernal Pérez, Felipe (1986). *Hñähñu: Recopilación de 505 Verbos Otomíes Conjugados*. Cardonal, Hidalgo, Mexico: Centro Social de Cardonal.
- Bernárdez, Enrique (1994). De la “lingüística catastrofista” a la lingüística cognitiva. *Revista de Filología Alemana* 2, 181–199.
- Cárceles, Fray Pedro de (1907 [1580?]). *Arte de la lengua otomí*. In *Boletín del Instituto Bibliográfico Mexicano*, Nicolás León (ed.), 1–64. Mexico DF: Instituto Bibliográfico Mexicano.
- Estrada Fernández, Zarina (2005). The pronominal form *-a* as a middle marker in Pima Bajo. *International Journal of American Linguistics* 71(3), 277–302.
- Fox, Barbara; and Hopper, Paul J. (eds.) (1994). *Voice: Form and Function*. Typological Studies in Language 27. Amsterdam: John Benjamins.
- Givón, Talmy (1981). Typology and functional domain. *Studies in Language* 5, 163–193.
- Haiman, John (1983). Iconicity and economic motivation. *Language* 59(4), 781–819.
- Haugen, Einar (1965). *Norwegian English Dictionary*. Oslo: Universitetsforlaget.
- Hekking, Ewald (1995). *El Otomí de Santiago Mexquititlán: Desplazamiento Lingüístico, Préstamos y Cambios Gramaticales*. Studies in Language and Language Use 17. Amsterdam: Institute for Functional Research into Language and Language Use.
- (2002). Desplazamiento, pérdida y perspectivas para la revitalización del hñähñu. *Estudios de Cultura Otopame* 3, 221–248.
- ; and Andrés de Jesús, Severiano (1984). *Gramática Otomí*. Querétaro: Universidad Autónoma de Querétaro.
- Kazenin, Konstantin I. (2001). Verbal reflexives and the middle voice. In *Language Typology and Language Universals*, Vol. 2, Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), 916–927. Berlin and New York: Mouton de Gruyter.
- Kemmer, Suzanne (1993). *The Middle Voice*. Typological Studies in Language 23. Amsterdam: John Benjamins.
- (1994). Middle voice, transitivity and the elaboration of events. In *Voice Form and Function*, Barbara Fox and Paul J. Hopper (eds.), 179–230. Amsterdam: John Benjamins.
- Klaiman, M. H. (1988). Affectedness and control: a typology of voice systems. In *Passive and Voice*, Typological Studies in Language 16, Masayoshi Shibatani (ed.), 25–83. Amsterdam: John Benjamins.
- (1991). *Grammatical Voice*. Cambridge Studies in Linguistics 59. Cambridge: Cambridge University Press.
- (1992). Middle verbs, reflexive middle constructions, and middle voice. *Studies in Language* 16(1), 35–61.
- Lastra, Yolanda (1989). *Otomí de San Andrés Cuexcontitlán, Estado de México*. Archivo de Lenguas Indígenas 13. Mexico DF: El Colegio de México.
- (1992). *El Otomí de Toluca*. Mexico DF: Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México.
- (1997). *El Otomí de Ixtenco*. Mexico DF: Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México.
- (1998). *Ixtenco Otomí*. Languages of the World: Materials 19. Munich: Lincom Europa.
- Lyons, John (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

- Maldonado, Ricardo (1992). Middle voice: the case of Spanish 'se'. Unpublished doctoral dissertation, University of California at San Diego.
- (1999). *A Media Voz: Problemas conceptuales del clítico SE*. Mexico DF: Universidad Nacional Autónoma de México.
- Maldonado (i.p.). Grammatical voice. In *Handbook of Cognitive Linguistics*, Dirk Geeraerts and Hubert Cuyckens (eds.). Oxford: Oxford University Press.
- Merlan, Francesca (1985). Split intransitivity: functional oppositions in intransitive inflection. In *Grammar inside and outside the Clause*, Johanna Nichols and Anthony Woodbury (eds.), 324–362. Cambridge: Cambridge University Press.
- Nava, Fernando; and Maldonado, Ricardo (2005). Basic voice patterns in Tarascan. In *Language, Culture and Mind*, Michel Achard and Suzanne Kemmer (eds.), 461–478. Stanford: Stanford University Press.
- Palancar, Enrique L. (1996). Effectors: more than mere instruments. *Estudios Ingleses de la Universidad Complutense* 4, 67–90.
- (1999). Instrumental prefixes in amerindian languages: an overview to their meanings, origins and functions. *Sprachtypologie und Universalienforschung* 52(2), 151–166.
- (2002). *The Origin of Agent Markers*. Studia Typologica 5. Berlin: Akademie Verlag.
- (2004a). Middle voice in Otomi. *International Journal of American Linguistics* 70(1), 52–85.
- (2004b). Minimizing morphology, maximizing function: Nasal morphemes in Otomi. Plenary talk at the 1st Conference of Otomanguan and Oaxacan Languages, March 19–21, University of California at Berkeley.
- (2004c). Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70(3), 251–278.
- Shibatani, Masayoshi (1985). Passives and related constructions. *Language* 61, 821–848.
- (2002). Synchronic typology and diachronic explanations: the case of middle voice. Plenary talk at the VII Encuentro Internacional de Lingüística del Noroeste, Universidad de Sonora.
- Sundman, Marketta (1987). *Subjectival och Diates i Svenskan*. Åbo: Åbo Academis förlag.
- Thom, René (1972). *Stabilité Structurelle et Morphogénèse*. New York: Benjamin.
- Tohno, T. (1999). Middle voice in Swedish. Unpublished MA thesis, Kobe University.
- Voigtlander, Katherine; and Echegoyen, Artemisa (1985). *Luces Contemporáneas del Otomí: Gramática del Otomí de la Sierra*. Mexico, DF: Instituto Lingüístico de Verano. [First impression in 1979 without author names].
- Wildgen, Wolfgang (1982). *Catastrophe Theoretic Semantics*. Amsterdam: John Benjamins.

MIDDLE VOICE IN OTOMI¹

ENRIQUE L. PALANCAR

UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

In this article, I claim that Otomi has a nasal morpheme that serves as an exponent of middle voice. My data come from the dialect of San Ildefonso Tultepec, but a similar phenomenon can be observed in other dialects, so that the claims made here could be extended to Otomi as a diasystem. The study is based on a sample of 72 verbs that display the middle morpheme, a sample intended as a quantitative reference for the qualitative patterns found in the language. About two-thirds of the verbs in this sample still show traces of lexical derivation. However, the derivative process has become unproductive and most middle verbs in Otomi function as independent units in the lexicon, showing just a morphological connection to their source verbs. My sample also contains cases where the source of the derivation was difficult, if not impossible, to reconstruct.

[KEYWORDS: Otomi, Otomanguean, middle voice, intransitivity, reflexives]

1. Introduction. In this article, I claim that Otomi has a morphological marker of middle voice. Otomi is spoken in Central Mexico and belongs to the Otopamean family, which is the northernmost branch of Otomanguean. The data in this article come from the dialect spoken in the village of San Ildefonso Tultepec, in the state of Querétaro, referred to here as San Ildefonso Otomi (SI Otomi). Even though Otomi displays considerable dialectal diversification, the phenomenon of middle voice is diasystemic and may be observed in other varieties of the language, e.g., Santiago Mexquititlán Otomi (Hekking and de Jesús 1984 and Hekking 1995), Mezquital Otomi (Bernal Pérez 1986), Sierra Otomi (Voigtlander and Echegoyen 1985), Ixtenco Otomi (Lastra 1997; 1998), and varieties spoken in Mexico State (Soustelle 1937, Andrews 1993, and Lastra 1989; 1992). Nevertheless, the descriptive literature on such dialects has not treated this phenomenon in a consistent fashion. In this sense, the claims made in this study are novel because they treat various facts of Otomi verbal morphology as exponents of middle voice.

¹ I benefited immensely from Scott T. Berthiaume's insightful comments on an earlier version of this article and am most grateful to him for his help. I also want to thank Ricardo Maldonado, who provided the inspiration to dig into an uncovered area of the Otomi language. Without his valuable previous research on the topic, I would never have become involved in this enterprise, let alone understood it. My sincere thanks as well to Paulette Levy, who commented on various aspects of middle voice systems. All of her suggestions improved this article substantially. Errors and deficiencies remain my own responsibility.

[*IJAL*, vol. 70, no. 1, January 2004, pp. 52–85]
© 2004 by The University of Chicago. All rights reserved.
0020-7071/2004/7001-0003\$10.00

Research in the area of middle voice in languages other than Indo-European is relatively recent and so proper study of middle voice phenomena in indigenous languages of Mexico is quite new. Middle voice marking has been reported in Pima (Uto-Aztecan) by Estrada (2002) and in Tarascan (isolate) by Friedrich (1969) and Le Cron Foster (1969) and more recently by Nava and Maldonado (forthcoming), who claim that Tarascan middle voice is a basic nonderivative voice in the language. However, we presently do not know how common the expression of middle voice is in other families of Mesoamerica.²

The study of middle voice developed to characterize the special verbal morphology present in Indo-European languages like Greek, Latin, and Vedic. The most popular definition of this phenomenon is by Lyons (1968:373), who argues that middle voice morphology applies when the “action or state affects the subject of the verb or his interests.” The phenomenon may be illustrated with a well-known example from Classical Greek (also used in Kazenin 2001:917): the middle form *lú-omai* [wash-1SG.MIDDLE (PRESENT.INDICATIVE)] ‘I wash myself’ expresses an action of washing that directly affects the subject. This form contrasts with the active form *lú-ō* (wash-1SG.ACTIVE (PRESENT.INDICATIVE)] ‘I wash’, which does not have this semantic implication. The middle form may additionally take an overt object. In this case, the transitive event is carried out by the subject for his/her own interest, e.g., *lú-omai khitōn-a* [wash-1SG.MIDDLE robe-ACCUSATIVE.SG] ‘I wash the robe (for myself)’.

Maldonado (forthcoming) points out that middle voice contrasts with active transitive voice. In the active voice, situations are depicted where two individual and well-differentiated participants interact, most commonly as agent and patient. In contrast, middle morphology is found in verbs which express actions involving the subject (cf. Smyth 1920, Klaiman 1992, and Maldonado [forthcoming]). In this sense, actions in middle voice commonly reflect a conceptual focus on “the subject’s own domain” (Maldonado [forthcoming] and Nava and Maldonado [forthcoming]). From a typological perspective, Kemmer (1993; 1994) proposes a group of semantic situation types where middle marking occurs cross-linguistically. Most of these reflect actions that take place within the subject’s own domain. Kemmer’s situation types can be grouped into the five main categories presented in table 1, adapted from Shibatani (2002). Each type is illustrated by Spanish verbs

²Paulette Levy has pointed out that Mel’čuk (1993:40–41) not only gives for Huastec (Mayan) examples (from Constable 1990) that parallel the semantic functions of the ones in Otomi but comments that “[t]his immediately reminds one of the notorious middle voice in classical languages.”

TABLE 1
MIDDLE SITUATION TYPES

Body Action middles	
Grooming actions	<i>afeitarse</i> 'shave'; <i>lavarse</i> 'wash'; <i>vestirse</i> 'dress'
Body function actions	<i>mearse</i> 'be pissing'; (mex.) <i>venirse</i> /(pen.) <i>correrse</i> 'come'
Change in body posture actions	<i>sentarse</i> 'sit down'; <i>tumbarse</i> 'lie down'
Nontranslational motion actions	<i>estirarse</i> 'stretch out'; <i>reclinarse</i> 'lean back'
Translational motion actions	<i>irse</i> 'go away'; <i>marcharse</i> 'leave'; (mex.) <i>regresarse</i>
Psychological middles	
Emotion middles	<i>enojarse</i> /(pen.) <i>cabrear</i> 'get angry'; <i>reírse</i> 'laugh'
Cognition middles	<i>imaginarse</i> 'imagine'; <i>confundirse</i> 'be mistaken'
Indirect middles	<i>conseguirse/agenciarse</i> /(pen.) <i>apañarse</i> 'get for oneself'
Reciprocal middles	<i>entenderse</i> 'understand each other'; <i>pelearse</i> 'fight/quarrel'
Spontaneous middles	<i>romperse</i> 'break'; <i>abrirse</i> 'open'; <i>desvincijarse</i> 'fall apart'

with the marker *se*, a reflexive clitic of third person that Maldonado (1992; 1999) has analyzed convincingly as an exponent of middle voice.³

Table 1 shows the five main semantic realms where middle morphology occurs cross-linguistically⁴—called “middle situation types” in Kemmer (1993). With the exception of spontaneous middles, all middle situation types comprise actions carried out by humans (at least in the default value). Prototypically, middle marking occurs in verbs that render a physical or a mental experience. The first type—body action middle—is a general label for a number of subtypes that depict actions involving the human body in a variety of (cultural) contexts.⁵

Middle voice is often equated with reflexives, which can lead to serious misunderstandings. This equation occurs because, in some languages, reflexive markers have developed such a wide range of semantic functions that

³ Table 1 also includes the categories “body function actions” from Maldonado (1999) and “reciprocal actions” from Kemmer (1994).

⁴ Middle situation types provide a good starting point for the study of middle voice phenomena, especially for poorly described languages. If these languages do have middle morphology, these types serve as useful methodological tools to find it.

⁵ For more information on middle voice, see Kemmer (1993; 1994), Maldonado (1992; 1999; forthcoming), and Kazenin (2001); also, Givón (1981), Shibatani (1985), Klaiman (1988; 1991; 1992), and Fox and Hopper (1994).

they also function as exponents of middle voice. This is the case in Spanish, where the marker *se* in the examples in table 1, often identified as a reflexive, is taken by Maldonado (1992; 1999) to be a morpheme expressing middle voice. In other languages, including English, reflexives have a limited range of functions. Since English lacks any other marker to do the job, it does not have middle voice among its relevant grammatical categories. The morphology expressing middle voice in many languages does not function as typical reflexive markers. Such languages may have an additional reflexive morpheme. For example, middle voice morphology in Greek is part of the paradigm of verbs, yet the language also has a set of reflexive pronouns, such as *eaútú* ‘himself’, which cover a similar limited range of functions as English reflexives. In this sense, a marker of middle voice may well not be a reflexive marker, even though this is common in languages like Spanish and German.

Middle voice is fundamentally a semantic phenomenon which is not easy to grasp employing grammatical diagnoses. Middle voice should not be taken to mean “diathesis.”⁶ However, some middle markers develop functions that enter the realm of syntax and thus come to serve a syntactic function where certain arguments may be projected to subject or where the valency of the verb is increased or decreased. Kemmer’s typological approach to middle voice is very convenient because it stipulates the different realms where middle voice morphology is expected to occur across languages and provides a means to categorize as middle any mysterious morphology showing the consistent range of functions that middles usually cover cross-linguistically.

In this article, I show that SI Otomi has a morpheme expressing middle voice. This morpheme is a nasal prefix that attaches to verbs. In the following sections, I identify the various functions of this nasal prefix and also propose that the marker in question is NOT a reflexive marker. In this sense, the Otomi system resembles the Greek system. Similarly, the middle marker in SI Otomi does not involve syntactic diathesis.

In 2, I briefly describe Otomi verbal morphology. In 3, I indicate sources for the data used in this analysis, while in 4, I introduce the middle morpheme and give some examples of its function as a middle voice marker in verbs depicting dressing actions. Middle derivation in Otomi is treated extensively in 5. There are at least four variables to take into account in such

⁶ Diathesis is defined by Trask (1993:82) as “[t]he relation between the semantic roles (deep cases, theta roles) subcategorized by a lexical verb or predicate and the surface expression of those roles as grammatical relations. Thus, for example, the verb *assault* requires an agent and a patient which, in the unmarked (active) case, are realized as subject and direct object, respectively.” In this sense, diathesis also involves changes in the relation between semantic and grammatical roles.

a process: the valency of the source verb (5.1), the valency of the middle verb (5.2), the semantics of the source verb (5.3), and the semantics of the middle verb (5.5). In 6, I present cases where derivation is no longer a transparent process in the system before concluding with a summary of this study's main findings and proposals (7 and 8).

2. San Ildefonso Otomi verbal morphology. Here, I present some facts about Otomi verbs that are relevant to understanding how middle morphology works in this language. SI Otomi verbs inflect both for tense/aspect and for the person of the subject by means of a complex system of function words that precede the verbal stem,⁷ illustrated in (1).⁸ Otomi is a tone language; however, tone is represented only in function words:⁹

- (1) *ne de ge=?pi ?ja bi posta ga?tho, ga?tho ja*
 and of IDE=DIST DP 3.PST bet all all PL
kosa Ø mí pɛ?ts'i
 thing REL 3.IMP possess

'And from there on he bet everything, everything he had'.

(1) illustrates the use of the function words *bi* (labeled 3.PST) and *mi* (3.IMP). Both morphemes express a third-person subject but differ in tense/aspect. The first conveys the past tense and the second the imperfect tense (imperfective aspect in the past).

Object marking is shown in the examples in (2):

- (2a) *ntonse him-bì <z>its'i(-Ø)*
 then NEG-3.PST NPS.take.animate(-3OBJ)¹⁰

'Then (the man) did not take her'.

⁷ For some dialects, such function words have been characterized as "proclitics" (see Hess 1968, Voigtlander and Echegoyen 1985, Hekking and de Jesús 1984, etc.). However, such a label is inadequate for SI Otomi because these words may receive full stress and appear at the edge of a prosodic unit. For other dialects, authors have treated them as prefixes, showing the outcome of a grammaticalization cline from an independent word into an affix.

⁸ Abbreviations are: B bound form of a verb; CP completive particle; DAT dative; DEF definite; DIM diminutive; DIST distal; DU dual; F fused form of a verb; FEM feminine; FUT future; IDE identifier; INDEF indefinite; IMP imperfect; INCL inclusive; IND indicative; INT intensifier; LOC locative preposition; M middle; NPS nonpresent stem; NEG negation; OBJ object; PL plural; POSS possessive; PRES present; PROG progressive; PST past; REC reciprocal; REL relative; SG singular; THV thematic vowel.

⁹ Tone does not play any role in the lexical phenomenon I treat in this article.

¹⁰ Despite the fact that the existence of a morphological empty unit is theoretically debatable, zero object anaphora is represented here with a zero morph (\emptyset) following the base.

- (2b) *ja gá tɔn-kagi, pe gá pengi*
 CP 2.PST win.F-1OBJ but 1.FUT return

gà tsiš-ŋi
 1.FUT take.animate-F-2OBJ

‘You beat me, but I will come back to get you’.

- (2c) *hingi ∅ heʔts’i-ŋi*
 NEG 3.PRES high-2OBJ

‘You’re not tall’.

SI Otomi has zero object anaphora on third person, as seen in (2a).¹¹ Object anaphora on first and second person is marked by suffixes on the verb, as shown in (2b). (2c) illustrates an intransitive verb that denotes a property concept. When this property is predicated on a first or second person, the latter is encoded as object. Before discussing the middle marker, I give some information about the sample and the methodology that I used in this study.

3. The sample. This research is based on a sample of 72 verbs from SI Otomi. These verbs come from oral texts I collected in the village of San Ildefonso Tultepec (Querétaro, Mexico) during elicitation sessions with speakers who were monolingual in the language in their youth. Originally, the sample contained 100 verbs with a nasal prefix, but 28 of them were removed because they did not have the morpheme {*N-*} in all tenses and all persons. This was important because SI Otomi has other nasal prefixes whose occurrence is conditioned by tense. Such cases are studied in more detail in Palancar (2003) and are not discussed further here.¹² I use the sample to give a quantitative profile to the various qualitative patterns middle verbs display in the lexicon.

4. The middle marker in SI Otomi. To introduce the middle marker, I present a brief review of dressing verbs in SI Otomi. Dressing verbs denote grooming actions, which qualify as the most common targets for middle voice morphology cross-linguistically. Otomi has a number of dressing verbs that are syntactically transitive. At the semantic level, these verbs refer to actions that involve a person who is dressing (the subject) and a

¹¹ Occasionally, an enclitic for third person is attached to the verbal complex and anaphorically refers to the NP object. Its distribution is still unclear.

¹² Most of these are members of a closed class of intransitive verbs that depict actions performed (mainly) by humans.

given piece of clothing (the object). The subject performs such dressing actions on him/herself. This is illustrated in (3):

- (3a) *ma gà kəde(-∅)*
 PROG 1.FUT put.something.on.below.waist(-3OBJ)

‘I’m going to put it on (i.e., a skirt)’.

- (3b) *no=r bɔtsi dā tɪt’=já thɪza*
 DEF=SG boy 1.FUT put.shoes.on.B=PL.3POSS shoe

‘The child will put his shoes on’.

The speaker in (3a) and the child in (3b) perform dressing actions involving a skirt and a pair of shoes respectively. The subjects are directly affected by the actions because they are dressed once the action is completed. In contrast, dressing actions performed on another person do not convey the same degree of affectedness onto the subject. In Otomi, such actions are conveyed by a different verb (4a) or by means of extra dative marking on the transitive verb (4b):

- (4a) *gà kəta=ma*
 1.FUT put.something.on.below.waist.onto.another.person.B=1POSS

bɔtsi
 child

‘I’ll put it (i.e., skirt) onto my child’.

- (4b) *dá tɪt’-wa=já thɪza ma bɔtsi*
 1.PST put.shoes.on-3DAT.B=PL.3POSS shoe 1POSS child

‘I’ll put the shoes on the child’.

The subjects of the verbs in (4) perform dressing actions just like the subjects in (3). However, they are not affected by them; they function as mere causers of the dressing event. In this sense, the inherent semantics of the transitive verbs in (3) involve actions within the subject’s own domain. In other words, the transitive verbs in (3) can be said to display middle semantics. A full list of dressing verbs in SI Otomi is given in Appendix A.

In addition, we find four dressing verbs in SI Otomi that have a nasal prefix. I treat this nasal morpheme as {*N-*}. The morpheme is illustrated in (5) (where in. = intransitive, lab. = labile, and tr. = transitive):

- (5a) *n-kə* (in.) ‘dress below waist’
 (5b) *ñ-he* (in.) ‘dress above waist’
 (5c) *m-phi?ts’i* (lab.) ‘put one piece of clothing on top of another’
 (5d) *n-tsiti* (tr.) ‘put earrings on’

As shown in (5), {*N*-} has three alternants, /*n*/ (5*a* and 5*d*), /*ɲ*/ <*ñ*> (5*b*), and /*m*/ (5*c*), which are phonologically conditioned: /*n*/ has the broadest distribution and is the default realization of the morpheme; /*ɲ*/ <*ñ*> occurs before a glottal consonant (/ʔ/ and /h/); and /*m*/ occurs before a labial (/b/, /p/, and /ph/).

The first two verbs in (5) are intransitive verbs and render the dressing action a routine activity without foregrounding any particular piece of clothing, as illustrated in (6):

- (6*a*) *dá* *n-kə*
1.PST {*N*}-dress.below.waist

‘I got dressed (I put pants/skirt on)’.

- (6*b*) **dá* *n-kə* *ma* *zeškho*
1.PST {*N*}-dress.below.waist 1POSS trousers

Intended meaning: ‘I put my trousers on’.

The verb *n-kə* ‘dress below waist’ in (6*a*) expresses a routine dressing action carried out by the speaker as s/he applied an unspecified piece of clothing to his/her lower body. The ungrammaticality of (6*b*) shows that the verb is intransitive.

The intransitive verbs *n-kə* ‘dress below waist’ (5*a*) and *ñ-he* ‘dress above waist’ (5*b*) are morphologically related to the intransitive verbs *kəde* ‘put x on below waist’ (3*a*) and *he* ‘put x on above waist’. As shown by the morphology, only *ñ-he* (5*b*) could be treated as clearly derived from the transitive *he* ‘put x on above waist’ by means of {*N*-}. As for *n-kə*, its relation to transitive *kəde* is more complex (see below).

The nasal morpheme is characteristically associated with intransitive forms but is not restricted to them. For example, the verb *m-phiʔtsʔi* ‘put one piece of clothing on top of another’ (5*c*) is labile, that is, it may be used as a transitive or an intransitive verb. However, the verb *n-tsiti* ‘put on earrings’ (5*d*) is transitive ONLY, as shown in (7):

- (7*a*) *dá* *n-tsita=ma* *ntsʔingu*
1.PST {*N*}-wear.earrings.B=1POSS earring

‘I wore my earrings/I put on my earrings’.

- (7*b*) *dá* *n-tsiti(-∅)*
1.PST {*N*}-wear.earrings(-3OBJ)

‘I put them on (i.e., my earrings)’/*‘I wore earrings’.

(7*a*) shows the transitive use of the verb. In contrast, (7*b*) is felicitous only if the verb is read as transitive, that is, with zero object anaphora, where the object (i.e., the earrings) is retrievable from context.

In this article, I propose that {*N*-} is a marker of middle voice in SI Otomi. The morpheme is found in verbal lexemes that convey typical middle situation types. For the middle situation types listed in table 1, the 72 verbs in my sample show the following distribution: 31 verbs express body actions (up to 43 percent of all verbs); 8 verbs psychological actions (11 percent); 8 verbs reciprocal actions (11 percent) (see 7.1); and 2 verbs spontaneous actions (3 percent) (see 7.2). Otomi does NOT use middle voice to express the area covered by indirect middles in other languages. Twenty-three verbs in the sample do not directly conform to any of the types listed. Nevertheless, these verbs also express actions affecting a human or animate subject, e.g., *m-ʔedi* (in.) ‘get lost’ or *n-khiʔi* (in.) ‘struggle to get free’.

A substantial number of verbs in SI Otomi acquire middle semantics when they receive {*N*}. For example, the verbs *m-phiʔtsʔi* (5c) and *n-tsiti* (5d and 7) derive from the verbs *phitsʔi* (tr.) ‘put something on top of another thing’ and *tsiti* (tr.) ‘hang something’, but they function now as full-fledged middle lexemes. {*N*-} places the action of a given source verb within the realm of middle voice, that is, within the subject’s own domain—more specifically, for these verbs, within a dressing–grooming scenario.

On the other hand, the data suggest that middle function is very old in SI Otomi because there are also a substantial number of verbs with {*N*-} that do not correlate clearly to a given derivative source. This phenomenon is treated in more detail in 6. For a brief illustration of this, look at the intransitive verb *n-kə* (5b). As I pointed out above, this verb is related to *kəde* (tr.) (3a), but it is not derived from it. Both verbs derive from an earlier source, reconstructible as **kə*, which no longer exists in the lexicon. The next section explores derivative processes in more detail.

5. Middle derivation in SI Otomi. Middle voice may be a derivative category, so that new middle lexemes are created out of morphologically unmarked verbs. This is the case for Greek, as well as many verbs in Otomi. In the sample, 47 out of 72 verbs have a lexical unmarked correlate. For such cases, the derivational process can be said to be transparent, at least on morphological grounds (see 5.4.4 for the nonproductivity of middle verbs and other obscure semantic processes). These cases make up 65 percent (65.2) of the sample. There are at least four variables to take into account in deverbal middle derivation: the valency of the source verb; the valency of the middle verb; the semantics of the source verb; and the semantics of the middle verb. All these variables interact closely in the derivational process, but it is convenient to consider them separately. A full list of the 47 derived verbs in the sample is given in Appendix B.

5.1. Valency of the source verb. A source verb in the derivation is usually a transitive verb: 33 out of 47 source verbs are transitive. However,

derivation from intransitive verbs is also common, as attested in the remaining 14 cases. Transitive derivation is illustrated in (8):

(8a) *no=r miši mí n-teʔts'e*
 DEF=SG cat 3.IMP M-lick (in.)

‘The cat was licking herself’.

(8b) *nu ma či lɛle ∅ teʔšá=r t'aphi*
 DEF 1POSS DIM baby 3.PRES lick.3POSS.B (tr.)=SG sweet

‘My baby is licking his lollipop’.

The middle verb *n-teʔts'e* (8a), which describes the self-centered action that cats do to themselves, derives from the transitive verb *teʔts'e* (8b).¹³ (9) shows the same case but with an intransitive source verb:

(9a) *no=r ndə mí n-khwaša=ka=r za*
 DEF=SG man 3.IMP M-lean.B=LOC=SG tree

‘The man was leaning on the tree’.

(9b) *no=r za mí khwaša=ka=r ngu*
 DEF=SG tree 3.IMP lean.B=LOC=SG house

‘The tree was leaning on the house’.

The verb *n-khwats'i* (9a) expresses a middle event that involves a human being as he is leaning with his body onto a given surface.¹⁴ The verb derives from the intransitive verb *khwats'i* (9b), which denotes the same event but for an inanimate entity showing neither volition nor control with respect to its locative state. Middle derivation from intransitive verbs is an expected pattern within middle systems. According to Maldonado (forthcoming: sec. 7), “[i]n languages having an intransitive/middle contrast, the middle designates an extra feature, a semantic specification not present in the plain intransitive.”

5.2. Valency of the middle verb. A derived middle verb is overwhelmingly intransitive. We find 43 such verbs in the sample, out of a total of 47. Thus, the resulting verb is intransitive in 93.5 percent of all cases. As expected, all intransitive source verbs yield a derived intransitive verb. Table 2 shows the correlation between the valency of a transitive source verb and the resulting middle verb.¹⁵

Table 2 shows that 87.7 percent of transitive source verbs derive an intransitive middle verb, which again illustrates that intransitivity is the default value for a middle verb in SI Otomi. These results are expected considering “the distinguishability of participants” hypothesis advanced by Kemmer

¹³ The morph *teʔša=* is a morphophonemically conditioned bound form.

¹⁴ The morph *khwaša=* is a morphophonemically conditioned bound form.

¹⁵ I thank an anonymous referee for suggesting this table to me.

TABLE 2
DERIVED CASES FROM A TRANSITIVE SOURCE VERB

Middle Verb	tr.	%
Transitive	3	9.0
Labile	1	3.3
Intransitive	29	87.7
Total	33	100.0

(1993). Kemmer proposes that middle voice conceptually decreases the number of participants involved in a given transitive event due to its main semantic function, which is to increase focus on the subject, making it the only relevant participant in the event. On the other hand, if we consider all 72 verbs in the sample, we find 66 middle verbs which are intransitive and only 6 which are transitive. This means that more than 90 percent of the verbs in the sample are intransitive.¹⁶

5.3. Semantics of the source verb. In some cases, the semantics of a source verb inherently depicts an action extended to another participant. Such verbs are called “extroverted verbs” by Haiman (1983:803), who argues that these verbs “describe actions which the subject usually performs toward others.” Extroverted verbs can be characterized as semantically active. When the middle morpheme is applied to an active verb, the action is immediately profiled as involving only the subject, and as taking place within the subject’s own domain. Examples in (10) and (11) illustrate this:

(10a) *dá šint’a=ma řueda*
 1.PST turn.something.around.B (tr.)=1POSS wheel

‘I turned around my wheel’.

(10b) *dá n-čint’i*
 1.PST M-turn.around (in.)

‘I turned around’.

(11a) *dá ?bedi(∅)*
 1.PST lose(-3OBJ) (tr.)

‘I lost it (i.e., my wallet)’.

(11b) *dá m-?edi*
 1.PST M-lose (in.)

‘I got lost’.

¹⁶ Among the 43 intransitive cases, we have the labile verb *m-pfi’ts’i* (5c) together with two other verbs that express a reciprocal action (see 7.1).

In these examples, the source verbs *šint'i* 'turn something around' (10a) and *ʔbedi* 'lose something' (11a) can be characterized as extroverted verbs: they portray actions performed by the subject onto external objects that do not pertain to the sphere of the subject.¹⁷ In contrast, the action is introverted when the middle morpheme is used. In (10b), the verb *n-čint'i* 'turn around' renders the nontranslational motion that the subject performs with its own body.¹⁸ Similarly, the middle verb *m-ʔedi* (11b) portrays the subject in an event where the subject himself is disoriented and lost.¹⁹

In other cases, the source verb has middle semantics of its own because it already refers to an action happening within the subject's own domain. Haiman (1983:803) calls such verbs "introverted verbs" because they "refer to actions which one generally performs upon one's self." These verbs could also be called "semantically middle." An example of a semantically middle verb is given in (12):

- (12a) *dá kɔʃt'a=ma ʔje*
 1.PST rub/massage.B (tr.)=1POSS hand/arm
 'I massaged my hand'.
- (12b) *dá kɔʃt'-wá=r ʔje*
 1.PST rub/massage-3DAT.3POSS.B (tr.)=SG hand/arm
 'I massaged his hand'. (Lit., 'I massaged him his hand'.)
- (12c) *dá n-kɔʃt'a=r ñɛthi*
 1.PST M-rub/massage.B (tr.)=SG cream
 'I put some cream on'.
- (12d) *dá n-kɔʃt'i(∅)*
 1.PST M-rub/massage(-3:OBJ) (labile)
 Transitive: 'I put it on (i.e., the cream)'.
 Intransitive: 'I gave myself a massage'.

¹⁷ An event of losing something may additionally be construed with the middle voice to express the affectedness of the subject's interests, e.g., in Spanish, when the verb *perder=se* 'get lost' is used transitively, the object of the loss is presented as something important for the subject. However, this does not happen in Otomi.

¹⁸ The affricate in the stem results from the phonological adjustment a postpalatal undergoes in contact with a nasal. The middle verb *n-čint'i* renders a turning action performed by a given figure on its own axis. Additionally, the language has other 'turning' verbs that are not middle, e.g., *gɔts'i* (in.) 'go around a given place so as to avoid passing through it' and *ʔbot'i* (lab.) 'turn your body but not around/go around a place'.

¹⁹ The labial in the consonantal cluster /ʔb/ is lost when the nasal is used. This is a general phonological process found elsewhere in the language. For example, the nonpresent stem of the verb *ʔboi* 'stand' is *mʔoi*, through the application of another nasal prefix in the language.

(12a) illustrates the use of the transitive verb *kɔʃt'i* 'rub/massage', as it conveys a rubbing action carried out by a given subject and directed to one of his own body parts. For this reason, the action remains within the subject's own domain and the verb qualifies as an inherently introverted or semantically middle verb. An extroverted rubbing action (directed to another participant) requires extra dative marking on the verb, as shown in (12b). In (12c), the verb receives the middle morpheme. With {*N*}, the verb is still transitive but now it takes a noun denoting some cream as its object. In this new reading, a middle meaning is clearly at work because the rubbing action takes place on the subject's own body by default. In other words, there is no need to specify a body part lexically, as in (12a). Alternatively, this middle verb may be used intransitively, as in (12d). In (12d), the verb expresses a self-centered action where the subject gives his own body a massage.

A similar case is illustrated in (13) with the verb *phɛʔts'e* 'trip':

(13a) *dá* *phɛʔsa=ma* *ʔwa*
 1.PST trip.B (tr.)=1POSS leg/foot
 'I tripped'. (Lit., 'I tripped my leg'.)

(13b) *dá* *m-phɛʔts'e*
 1.PST M-trip (in.)
 'I tripped'.

(13a) is an example of the transitive verb *phɛʔts'e*, which evokes an involuntary event of 'tripping'. Here a noun referring to the subject's affected body part, namely his leg or his foot, functions as the object of the verb. The same event may be rendered in intransitive terms by means of the middle form (13b). Speakers report there is no real difference in meaning between these two expressions.

In summary, I present, in table 3, the involvement of both active and semantically middle verbs as sources of derived middle verbs in the sample. Thirty-three source verbs are extroverted, that is, they have active semantics as in examples (10) and (11), representing 70 percent of the derived cases. Table 3 also shows that most extroverted source verbs are transitive. The other 14 cases are introverted or semantically middle verbs, representing the remaining 30 percent of derived verbs, which is a significant percentage and indicates that they play an important role in the middle derivation. At this stage, it is not clear what the motivation is for producing a morphological middle verb if the language already has alternative lexical means to cover middle semantics by means of introverted verbs. I consider this further in 5.4, where I explore the semantics of a derived middle verb in detail.

5.4. Semantics of the middle verb. In the derivation, a new middle verb is produced. The data suggest that the middle marker in SI Otomi can con-

TABLE 3
SEMANTICS OF SOURCE VERBS

	Total	%	tr.	in.
Extroverted	33	70	27	6
Introverted	14	30	6	8
Total	47	100	33	14

tribute in at least three different ways to the creation of the new lexeme. The first meaning is of intensity and repetition of the action. Such a reading is presented in 5.4.1. The second meaning, treated in 5.4.2, is more schematic. It reflects a shift in conceptual perspective toward the inner self of the subject and its internal energy in performing the action. In 5.4.3, I study the self-centered meaning, by which typical extroverted actions are turned into introverted ones and are directed to the subject's self. Despite these meanings, most middle lexemes have gained autonomy in the lexicon through the process of derivation; they are now old words, as it were, or lexically listed rather than derived. Accordingly, some of them have become semantically rather distant from their respective morphological sources. For such cases, it is difficult to reconstruct the path that led to the semantic extension. Consequently, most such verbs have unpredictable meanings synchronically. Relevant cases of this type are discussed in 5.4.4.

5.4.1. Middle of intensity. As illustrated above, a source verb may already express middle semantics, and yet it can also serve as a base for the derivation of a new verb which is both semantically and morphologically middle. In many cases, when a morphological middle verb is derived from a semantically middle one, the new verb develops an intensive meaning. In most instances, the source verb expresses a punctual action. The middle profiles the action of the source verb as imperfective. With the imperfective construal, the maintenance of a given state of affairs conveys a sense of repetitive action and brings about a nuance of intensity. A few examples illustrate this point.²⁰

The transitive verb *pət'e* refers to a situation where a person accidentally hits some part of his body against something in his environment (i.e., a stone, a table, etc.). The verb takes a noun denoting the body part hit as its object. When the verb receives middle morphology, it surfaces as *m-pət'e* and is no longer transitive. In middle voice, the verb conveys a situation where the subject is profiled as harming himself by inflicting injury on his body with a knife or stick in an iterative manner.

²⁰ Ricardo Maldonado suggested this dimension to me. I am grateful to him for his insightful observations.

Another example is the semantically middle source verb *tsu* ‘be frightened (by something or someone)’. When this verb receives middle morphology, it surfaces as *n-tsu* and means ‘be afraid’. Here, the sensation of being frightened from the verb *tsu* is intensified and it is profiled as happening through an unbounded stretch of time. The same could be said of the transitive verb *beni* ‘remember something/think of something’. The derived middle is *m-beni* ‘think (over)’. The act of remembering is first construed as a “punctual” (i.e., distinct in time) mental experience when something comes to mind. The morphological middle makes that punctual experience unbounded and intensifies it in order to interpret an act of thinking in such terms (Ricardo Maldonado, personal communication). An intensive reading occurs in the examples in (14):

(14a) *dá šɔt'i*
 1.PST scratch (in.)
 ‘I scratched (punctually)’.

(14b) *dá n-čɔt'i*
 1.PST M-scratch (in.)
 ‘I scratched all over (because it was itching a lot)’.

The verb *šɔt'i* ‘scratch’ in (14a) shows the intransitive use of this verb. Additionally, the verb could also take an object expressing the subject’s body part to which the scratching action is directed. The verb in (14a) expresses a situation in which the subject scratches his body once or twice to relieve himself from a punctual itching sensation—a common human response to an equally common corporal experience. From this intransitive use, the verb may receive middle morphology, as shown in (14b).²¹ Here the focus has shifted toward the internal world of the subject. The subject is seen as scratching a lot because, apparently, his body is itching acutely from a rash or some other internal cause.

Similar intensive readings are found in the verbs in (15):

Source	Middle
(15a) <i>tsithe</i> (in.) ‘drink’	<i>n-tsithe</i> (in.) ‘drink to get drunk’
(15b) <i>ts'oni</i> (tr.) ‘get some-thing dirty’	
(in.) ‘get dirty’	<i>n-ts'oni</i> (in.) ‘urinate/defecate on oneself’
(15c) <i>tsɔ</i> (in.) ‘feel’	<i>n-tsɔ-wi</i> (in.) ‘fight to defend oneself’
(15e) <i>ʔɔni</i> (in.) ‘hide’	<i>ñ-ʔɔni</i> (in.) ‘hide away not to be seen’

²¹ /j/ <š> becomes /tʃ/ <č> in contact with /n/.

The source verb in (15a) is not an introverted, semantically middle verb. However, in its derived middle form, the verb expresses a drinking action that involves the subject as a whole and bears an intensive reading when the subject drinks so much (or so frequently) for the sole purpose of getting drunk. The source verb in (15b) is labile. When used as an intransitive, it denotes a middle scene where the agent gets a stain on his clothes accidentally. From the intransitive, we obtain the verb *n-ts'oni* 'urinate or defecate on oneself (as the result of having an illness, being old, etc.)'. This reading renders both an intensive and an extensive reading of the pertinence of the unpleasant state of affairs inherited from the source verb. In (15c), the verb *tsɔ* 'feel' derives a middle along with the suffix *-wi*, a (dual) inclusive. The morphological middle verb expands and intensifies the sensation to refer to an act of self-defense. The inclusive marker additionally introduces an implicit reference to an unspecified attacking entity. Finally, when the intransitive verb *?oni* 'hide' receives middle marking, as in (15e), it denotes an action of hiding away so that the subject will not be seen by anyone (in the case of, e.g., a hermit, a refugee in danger, a terrorist, etc.).

5.4.2. Middle of internal energy. Alternatively, the derived middle verb may focus on the internal energy of the subject while performing the action. This reading contrasts with the external perspective present in the active source verb. As with the middle of intensity, there is also a nuance of repetition, illustrated by the four verbs in (16):

Source	Middle
(16a) <i>phehmi</i> (tr.) 'hit'	<i>m-phhmi</i> (in.) 'flutter wings'
(16b) <i>phet'i</i> (tr.) 'hit'	<i>m-phet'i</i> (in.) 'swim'
(16c) <i>khit'i</i> (tr.) 'pull horizontally'	<i>n-khit'i</i> (in.) 'struggle to get free'
(16d) <i>?oni</i> (tr.) 'ask a question'	<i>ñ-?oni</i> (in.) 'get informed'

The verb *phehmi* (16a) conveys a hitting action similar to the one typically used on a carpet to remove dust or on sown beans to make the seeds to come off their dried sheaths. The middle verb refers to the nontranslational motion effected by nonflying birds or stationary (wounded) insects. The iterative hitting motion present in the semantics of the source verb serves to convey the energetic inner motion displayed by these winged animals when they flutter their wings in a (failed) attempt to fly again. In (16b), the verb *phet'i* denotes a hitting action performed with large movements of the arm, as, for example, when hitting a drum.²² The translational middle verb *m-phet'i* 'swim' was coined from this verb. The motion is now well established in the lexicon and is no longer linked by speakers to its source verb

²² In fact, the verb has extended its meaning and is used for 'play a drum'.

in the derivation. The verb was probably first used to express a swimming action by foregrounding the manner of moving the arms in the water and by defocusing the extroverted hitting action. In other words, the perspective shifted from the external to the internal world of the subject. (16c) serves as yet another example showing this shift in conceptual perspective. The source verb expresses an action of pulling something on a horizontal axis. The middle verb *n-khít'i* 'struggle to get free' describes a tied person or animal as he/she/it struggles to get free. The effort of moving an object (i.e., a piece of furniture, a rock, a door, etc.) performed by the agent of the source is now shifted to the internal world of the subject of the middle, who is profiled as physically struggling to get free from his/her ties. Finally, the verb *ʔɔni* 'ask a question' in (16d) derives the middle verb *ñ-ʔɔni* 'get informed'. The inherently extroverted action of the active verb is put in the background as the perspective shifts to the inner psychic world of the subject. The subject still requires information by asking a question, but the middle verb profiles his desire to know about a topic of his own motivation.

5.4.3. Self-centered middle. The application of the middle morpheme to an extroverted verb may turn the action into an action performed on the self, termed a self-centered action, as illustrated by the verbs in (17):

Source	Middle
(17a) <i>heʔts'e</i> (tr.) 'cover someone'	<i>ñ-heʔts'e</i> (in.) 'cover oneself'
(17b) <i>šɔki</i> (tr.) 'uncover someone'	<i>n-čɔki</i> (in.) 'uncover oneself'
(17c) <i>sɯ</i> (tr.) 'take care of someone'	<i>n-sɯ</i> (in.) 'take care of oneself'
(17d) <i>teʔts'e</i> (tr.) 'lick something'	<i>n-teʔts'e</i> (in.) 'lick oneself'
(17e) <i>tsiti</i> (tr.) 'weigh something'	<i>n-tsiti</i> (in.) 'weigh oneself'

The transitive verbs *heʔts'e* (17a) and *šɔki* (17b) mean either to put blankets on or to remove blankets from someone. Their middle counterparts are intransitive verbs that mean to put blankets on or remove them from oneself. Similarly, the middle form of the verb *sɯ* (17c) 'take care of somebody' denotes a self-care action, also seen in the verbs in (17d) and (17e).

Self-centered actions like those in (17) are commonly expressed by reflexive markers in other languages (note the English glosses). This raises the question of whether the nasal prefix {*N-*} is indeed a reflexive marker showing the typical extensions reflexives undergo in other languages (Kemmer 1993; 1994). In other words, the presence of typical reflexive semantics in these examples could indicate that {*N-*} should be treated as a reflexive. If so, the derivation of verbs like *n-teʔts'e* 'lick oneself (i.e., a cat)' (17e) or *n-tsiti* 'weigh oneself' (17f) would be the transparent outputs of the reflexive function of the morpheme. Against such a view, I argue in the next sec-

tion that {*N-*} is NOT a reflexive marker, and I support my arguments with evidence from various domains.²³

5.4.3.1. Why middle {*N-*} is not a reflexive. One of the most common tests for reflexives is their occurrence in so-called direct reflexive types, such as ‘see oneself’ or ‘hit oneself’. Otomi does not make use of the nasal morpheme in the typical situation types associated with reflexive markers. For such situations, Otomi uses the transitive construction. This is illustrated in (18):

- (18a) *no=r* *?behño* *bì* *k’qt’á=r* *hmi*
 DEF=SG woman 3.PST look.at.3POSS.B=SG face

ka=r *ñhe*
 LOC=SG mirror

‘The woman looked at herself in the mirror’. (Lit., ‘The woman looked at her face in the mirror’.)

- (18b) **no=r* *?behño* *bì* *n-k’qt’a=ka-r* *ñhe*
 DEF=SG woman 3.PST M-look.at.B=LOC=SG mirror

Intended meaning: ‘The woman looked at herself in the mirror’.

(18a) shows that the transitive construction expresses a reflexive action, such as looking at oneself in the mirror. A middle verb for such a situation is not possible, thus the ungrammaticality of (18b). Consider these other examples:

- (19a) *dá* *handa=ma* *hmi* *ka=r* *ñhe*
 1.PST see.B=1POSS face LOC=SG mirror

‘I looked at myself in the mirror’. (Lit., ‘I looked at my face in the mirror’.)

- (19b) *dá* *handa=ka=r* *ñhe*
 1.PST see.B=LOC=SG mirror

‘I looked at myself in the mirror’. (Lit., ‘I looked (it) in the mirror’.)

²³ The literature sometimes uses the term “reflexive” in a confusing way. Kazenin (2001) uses this term to refer to any type of morphology that renders a “direct reflexive reading.” He distinguishes two types of “verbal reflexives”: the “derived one,” which emerges historically from an independent reflexive pronoun (e.g., Old Norse *-sk* from an older pronoun **sik*) and the “underived type” like the Greek example in (1). In view of these categories, I would argue that Otomi falls within the first type, but there is little reason to believe that {*N-*} was once a reflexive.

- (19c) *ma bɔtsi/nkhwat'i mí ñ-handa=ka=r khodo*
 1.POSS child/chair 3.IMP M-see.B=LOC=SG wall
 'My child/chair was facing to the wall'.

(19a) is a natural lexical alternative to the expression in (18a). The verb *handi* 'look at/see' can also be used without an object, as in (19b). When this happens, the clause conveys a situation in which the subject is looking at his whole body. In contrast to *k'ɔt'i* 'look at' (18), *handi* may indeed receive middle marking, as shown in (19c). However, the result is not a reflexive reading, which one might expect if {*N*-} were a reflexive. The middle verb in (19c) describes a situation where a given figure (animate or inanimate) is facing toward a given ground. The middle verb is no longer a looking verb (for the unpredictability of the semantics of middles in Otomi, see 5.4.4).

The verb *phei* 'hit' in (20) shows a similar phenomenon:

- (20a) *dá pɬeʔ=ma ʔje*
 1.PST hit.B=1POSS hand
 'I hit myself on the hand'. (Lit., 'I hit my hand/arm'.)
- (20b) *dá pɬeʔ-se*
 1.PST hit.F-ALONE
 'I hit myself'.
- (20c) **dá m-pɬei*²⁴
 1.PST M-hit

Intended meaning: 'I hit myself'.

As with looking verbs, (20a) shows that the self-centered act of hitting oneself is conveyed by the transitive construction. Speakers may use the alternative morpheme *-se* to emphasize that the subject performed the hitting action alone, without the help of anybody else, as in (20b), while (20c) shows that the reflexive reading cannot be expressed by the verb plus {*N*-}.²⁵

Another strong argument against a reflexive interpretation of {*N*-} comes from the fact that the semantic extension of the verbs in (17) occurs in only

²⁴ The form *m-pɬei* does exist, but it realizes an impersonal inflection which additionally shows complex morphophonological processes. A few verbs take a nasal morpheme, similar in form to the middle marker. These two morphemes may be related historically. However, the study of this relation falls out of the scope of the present article.

²⁵ Additionally, the language has a middle verb to describe a death by suicide that derives from the verb *ho* 'hit/kill', the derived verb *ñ-ho-se* 'commit suicide'. This form is marked morphologically. The verb requires the suffix *-se*, the same one as in (20b). The affix informs the listener that the event of killing happened because of the subject's own actions. A form such as **ñ-ho*, without the morpheme *se*, does not exist.

7 verbs out of 72, which represents slightly under 10 percent of all verbs. Meanwhile, the rest of the verbs show typical extensions of middle voice systems.

Similarly, the expression of “indirect reflexive types” provides another piece of evidence that {*N-*} is not a reflexive marker. Languages with reflexive pronouns commonly use them to express events denoting the genuine interest of the subject in the action. Such reflexives are called “indirect reflexives” and occur in English in predications such as *John built a house for himself*. For this function, Otomi uses a special set of pronouns, but not {*N-*}, as illustrated in (21):

(21a) *dá tɔŋga=?na=r ŋɟu pa ge-ke*
 1.PST buy.B=INDEF=SG house for IDE-1SG
 ‘I bought a house for myself’.

(21b) **dá n-tɔŋga=?na=r ŋɟu*
 1.PST M-buy.B=INDEF=SG house
 ‘I bought a house for myself’.

The genuine interest of the speaker in buying the house in (21a) is expressed by a special pronominal form of first person (*-ke*) together with a relative pronoun.²⁶ The ungrammaticality of (21b) shows that {*N-*} does not have this meaning.²⁷

Apart from the above argument, the middle of intensity of {*N-*} presented in 5.4.1 and the middle of internal energy in 5.4.2 are not found among the extensions undergone by reflexives (Ricardo Maldonado, p.c.). Additionally, Otomi has a substantial number of underived verbs that are otherwise difficult to explain under a restricting reflexive account (see 6 below).

As for derived verbs, consider again *tsiti* ‘weigh something’ in (17f). This verb also means ‘hang something’. The ‘hang’ meaning is more basic than the ‘weigh’ meaning, as the verb came to mean ‘weigh something’ through a bridging context where ‘hanging’ things on a scale was reinterpreted as an act of ‘weighing’ them. In addition, the middle verb *n-tsiti* ‘weigh oneself’ (17f) is homophonous with the transitive middle verb *n-tsiti* ‘wear earrings’ in (7) above. This latter verb derives from the ‘hang’ meaning, but it has NOT come to mean ‘hang oneself’, as would be expected if {*N-*} were a reflexive. Instead, the verb has a clear middle reading pertaining to the grooming semantic type. Semantic idiosyncrasies like these are expected in middle systems. In contrast, Otomi uses another verb for the act of hanging oneself, the

²⁶ This relative pronoun cross-references a participant already established in the clause.

²⁷ The crucial point here is that if {*N-*} were a reflexive, it would be expected to convey the meaning in (21a), despite the fact that other middle systems may use middles to express them too (i.e., Greek).

middle verb *n-zidi*, which derives from the intransitive verb *tsidi* ‘hang’ and not from the transitive *tsiti* ‘hang something’. This verb serves as a reminder that middle verbs in Otomi derive from intransitive sources too.

As we saw in 5.1, middle verbs derived from an intransitive verb represent 30 percent of the sample. If {*N-*} were a reflexive marker, one would expect that transitive source verbs like *tsiti* ‘hang something’ or *khwaʔtsʔi* ‘lean something on somewhere’ would yield derived verbs meaning ‘hang oneself’ or ‘lean oneself on’. However, this is not the case, as has been shown (e.g., see 9 above). Middle systems like the one in SI Otomi are concerned with the degree of involvement of the subject in the situation expressed by the verb. In such systems, source verbs and derived verbs can be, and very frequently are, syntactically intransitive. This is not the case for voice systems that apply diathesis as a means to project arguments to subject position. For example, applying passive voice to an intransitive verb would result in an impersonal verb. In middle voice, the application of middle morphology to an intransitive verb renders an increased degree of involvement or affectedness of the subject, but there is no syntactic adjustment of the arguments involved; that is, both verbs have the same referent as subject. This is again a characteristic typical of middle systems.²⁸

Furthermore, Otomi shows great unpredictability regarding which source verb will function as base for this type of derivation, as illustrated in (22):

- (22a) *dá heka=ma ʔje ko=r kwai*
 1.PST cut.B=1POSS hand/arm with=SG knife
 ‘I got a cut with the knife’. (Lit., ‘I cut my hand with the knife’.)
- (22b) **dá ñ-heki ko=r kwai*
 1.PST M-cut with=SG knife
 ‘I got a cut with the knife’.
- (22c) *dá ñ-heni ko=r kwai*
 1.PST M-chop with=SG knife
 ‘I got a cut/cut myself with the knife’.

(22a) shows the use of the transitive verb *heki* ‘cut’. When a person accidentally cuts himself with a knife, speakers commonly express this self-

²⁸ Reflexive markers are typically pronominal (e.g., Spanish *se* or German *sich*), while {*N-*} is not. Even though middle markers may evolve from reflexive through grammaticalization, it is difficult to discern here any possible historical source for {*N-*}, given that the language has a general liking for transitive constructs when encoding direct reflexive types, and there is little reason to believe that such a constructional pattern is modern. Besides, cases subjected to grammaticalization leave some trace of the change in word order or phonological resemblance. Otomi {*N-*} is a prefix, while object marking in the language is extensively suffixal, as shown in (2) above.

centered situation type with a transitive predication similar to (22a). The verb *hɛki* in (22a) is a generic cutting verb, but it does not derive a middle verb to give an alternative expression to the situation denoted in the example, as the ungrammaticality of (22b) shows. However, another lexical item is used to express such a situation, the verb *ñ-hɛni* ‘cut oneself’ (22c), which derives from *hɛni* ‘chop’, a specific cutting verb that cannot be used to refer to the event in (22c) with the same transitive frame as in (22a) (for more discussion of morphology, see n. 28).

5.4.4. Productivity of middle derivation. Despite the existence of the three major semantic functions for the middle morpheme in Otomi, the derivational process is no longer productive in the language. For example, no loans from Spanish accept {*N*-}. In this sense, derived middle verbs in Otomi form a closed lexical class. Most of the verbs have well-established semantics in the lexicon, a feature that makes them look like old lexemes. One can still discern possible semantic paths leading the verbs to acquire their current meaning. However, if derivation of middles were a productive process, there would be no clues about where the semantic shift of the verb might go. To illustrate this phenomenon, consider the verbs in (23):

Middle	Source
(23a) <i>ñ-handi</i> (in.) ‘face toward a place’	<i>handi</i> (tr.) ‘look/see’
(23b) <i>n-khít’i</i> (in.) ‘struggle to get free’	<i>khít’i</i> (tr.) ‘pull horizontally’
(23c) <i>ñ-hoki</i> (in.) ‘get ready’	<i>hoki</i> (tr.) ‘fix’
(23d) <i>n-thuki</i> (in.) ‘get a spiritual ’ cleansing	<i>thuki</i> (tr.) ‘clean’
(23e) <i>m-ʔii</i> (in.) ‘be born’	<i>ʔbii</i> (tr.) ‘live/be’
(23f) <i>m-bəni</i> (in.) ‘shake (for a turkey)’	<i>bəni</i> (in.) ‘exit’
(23g) <i>m-pɛʔts’i</i> (tr.) ‘mourn a dead person’	<i>pɛʔts’i</i> (tr.) ‘possess’

The middle verb *ñ-handi* ‘face toward a place’ (23a) derives from the verb *handi* ‘see’. The semantic path leading to this extension is similar to the one motivating the use of an English verb like *look* in expressions like *the hotel looks to the ocean*. However, there is no apparent reason the verb did not come to mean ‘see oneself’, ‘see internally, envision, or foresight’, or even ‘see intensively’. All these semantic outcomes could have been possible by virtue of one of the middle meanings in the previous sections. Nevertheless, none of these readings is possible for this verb. The language would need other constructional means to convey these ideas. In a similar vein, the meaning of the middle verb *n-khít’i* (23b) ‘struggle to get free’ was explained by means of the middle of internal energy (treated in 5.4.2), yet the ties the subject of the middle verb is subjected to are not present at all in the semantics of the active source verb. Similar cases are seen in the other verbs in (23). The conventional meanings of the middle verbs *ñ-hoki* (23c), *n-thuki* (23d), and *m-ʔii* (23e) cannot be directly and transparently retrieved

from the lexemes of their morphological sources. The verbs must be learned as independent lexical items. Furthermore, the middle verb *m-bəni* (23f) is a middle verb that describes the sudden body shaking performed by turkeys when they are producing their special mating sound. This verb derives from *bəni* ‘exit’, which is a motion verb that has little to do with the meaning of the middle. Finally, the verb in (23g) represents the most obscure case I have in the sample with regard to semantic transparency. The use of this verb is illustrated in (24):

(24a) *no=r ndə mí pɛʔsa=r bokhɔ*
 DEF=SG man 3.IMP possess.B=SG money

‘The man had money’.

(24b) *no=r ndə mí m-pɛʔsá=r*
 DEF=SG man 3.IMP M-mourn.a.dead.person.3POSS.B=SG

či nono
 DIM mother

‘The man was mourning his (deceased) mother’.

The verb *pɛʔts’i* ‘possess’ is commonly used to render the possession of alienable things, for example, money, as in (24a). The verb *m-pɛʔts’i* in (24b) was originally derived from the verb in (24a), but it is no longer possible to reconstruct the cultural bridge that led the verb in SI Otomi to acquire the meaning ‘mourn a dead person’.²⁹ Important to this notion of unpredictability and non-productivity of the middle morpheme in Otomi are the cases in the data that do not derive from any existing lexical source, reviewed briefly in 6.

6. Underived middles. Across middle systems it is very common to find middle verbs which lack an unmarked lexical counterpart (Kemmer 1994 and Klaiman 1992). Such verbs are often called “deponents” or “*media tantum*.” In the data, 25 verbs have a middle marker but do not derive from any existing lexical item in the language. Such verbs, called underived middle verbs, represent 35 percent of the sample, which is a significant percentage. Among underived verbs, two types may be distinguished.

6.1. Formerly derived middles. A middle verb in Otomi may not have a synchronic lexical correlate in the lexicon, but there may be other pieces of evidence pointing to a historical lexical source that has vanished from the

²⁹ Kinship relations are regarded culturally as a type of possession. However, such a relation is not expressed by *pɛʔts’i* ‘possess’ but by means of a different verb, *ʔnɛhɛ*, which also applies to the possession of domestic animals.

system. The evidence is based on the existence of other verbs with the same morphological root as the middle verb. Consider the verbs in (25):

- (25a) *ñ-hy-i* (tr.) ‘lay eggs’
 (25b) *hy-di* (in.) ‘sit down’
 (25c) *hy-ts’i* (in.) ‘sit down over some place’
 (25d) **hy*
 (25e) **hy-i*

The verbs in (25a)–(25c) share the same morphological root, *hy-*. However, this root is not lexical. In other words, the root does not mean ‘sit’ synchronically, even though it probably had such a meaning historically. The morphs *-i* (25a), *-di* (25b), and *-ts’i* (25c) are semantically obscure; they function nowadays as classificatory formatives, that is, they lack a specific morphemic function. The lexeme in these verbs is realized by the co-occurrence of the root plus the formative, just as in some English verbs like *pre-fer*, *in-fer*, *de-fer*, etc. The verbs in (25a)–(25c) are related because they originated from a common historical source which is no longer available in the language as a verb, as shown in (25d). Crucially, the impossibility of (25e) shows that the middle verb *ñ-hyi* ‘lay eggs’ in (25a) does not derive from any existing verb. Middle verbs like these are called “formerly derived middles.” Table 4 presents a list of such verbs.

Table 4 includes all verbs in the sample that show the same behavior as *ñ-hyi* ‘lay eggs’ (25a). Most of these verbs express actions whose semantics pertain to the situation type of body actions, which is the prototypical semantic realm of middle markers across languages. We also have two reciprocal verbs (see 7 for more details on this category) plus *ñ-hyi* ‘lay eggs’, which has been characterized as a miscellaneous verb. From the related verbs, one may understand that the act of laying eggs was originally construed as an act of sitting. The middle profiled the body posture of this activity performed by hens by emphasizing its internal dynamicity.

6.2. Purely underived middles. In contrast to formerly derived middles, other verbs offer no clues about their lexical source in the derivation. I call such verbs “purely underived middles.” Verbs of this kind are presented in table 5.³⁰

All verbs in table 5 (except the last one) convey actions performed by human beings by default. Most of the actions pertain to physical or psychological domains. In other words, the semantics of the verbs fall into the body function and psychological middle situation types. The other miscellaneous

³⁰ For some of these verbs, related lexical items may exist, but I have not yet been able to find any.

TABLE 4
FORMERLY DERIVED MIDDLES

Middle Type	Middle Verb	Historically Related Verb
Body Action		
Grooming	<i>ñ-hi</i> in. 'wash-one's-head'	~ <i>hi-ti</i> tr. 'wash one's head'
	<i>n-kə</i> in. 'dress below waist'	~ <i>kə-de</i> tr. 'wear below waist'
	<i>n-ča-ha</i> in. 'bathe/take a shower'	~ <i>ša-t'i</i> tr. 'bathe somebody'
	<i>n-čĩ-LEX</i> in. 'wash-one's-LEX (foot/hand/face)'	~ <i>šĩ-ki</i> tr. 'wash one's body part'
		~ <i>šĩ-t'i</i> tr. 'wash up (i.e., dishes)'
Body function	<i>n-čq-hi</i> in. 'itch (internally as a rash)'	~ <i>šq-ki</i> in. 'itch (i.e., mosquito bite)'
	<i>m-pi-t'i</i> in. 'urinate'	~ <i>pi-ts'i</i> tr. 'urinate over someone'
		~ <i>pi</i> in. 'melt'
Reciprocal	<i>n-khy-_{DU}</i> tr. 'chase someone'	~ <i>kų-i</i> tr. 'expel someone'
	<i>n-tų-_{DU}</i> in. 'fight'	~ <i>tų-ki</i> tr. 'push'
Miscellaneous	<i>ñ-hų-i</i> tr. 'lay eggs'	~ <i>hų-di</i> in. 'sit'
		~ <i>hų-ts'i</i> in. 'sit over some place'

TABLE 5
PURELY UNDERIVED MIDDLES

Body Function		
Body function	<i>n-čon-the</i> in. 'sweat' (- <i>the</i> 'water', reduced form)	
Body posture	<i>m-pints'i</i> in. 'turn around as in bed'	
	<i>n-dondi-hmu</i> in. 'kneel' (- <i>hmu</i> 'knee')	
Nontranslational motion	<i>ñ-?ani</i> in. 'duck'	
	<i>ñ-hwqł'i</i> in. 'tremble'	
Translational motion	<i>n-tsoni</i> in. 'fly'	
Psychological		
Emotion	<i>n-sente</i> in. 'be/get sad'	
	<i>n-zqi</i> tr. 'like a place'	
Cognition	<i>n-tsə?mi</i> in. 'worry'	
Miscellaneous	<i>ñ-he</i> in. 'be quiet'	
	<i>n-ti</i> in. 'be/get drunk'	
	<i>n-tsini</i> in. 'be/get covered with thorns'	
	<i>ñ-?eni</i> lab. 'play'	
	<i>n-wphi</i> lab. 'plow'	
	<i>n-the šĩñho</i> in. 'be convenient' (<i>šĩñho</i> 'well/good')	

verbs represent actions that may be construed as happening within the subject's own domain. They can affect him or her directly (e.g., 'be quiet', 'get drunk', 'get covered with thorns') or they represent routine activities (e.g., 'play'). The application of {*N*-} in these verbs is no longer transparent; it has lexicalized, but its semantic pertinence is still justifiable.

7. Other extensions of the middle. According to Kemmer (1993; 1994), middle markers are often found cross-linguistically in verbs depicting reciprocal and spontaneous events. These occur separately in Otomi.

7.1. Reciprocals. Middle morphology has become obligatory for verbs denoting reciprocal actions. In my data, there are eight such verbs, presented in (26):

	Middle		Source
(26a)	<i>n-tu-hni</i> (in.)	'fight (each other)'	~ <i>tu-ki</i> (tr.) 'push'
(26b)	<i>n-khu-DU</i> (tr.)	'chase someone'	~ <i>ku-i</i> (tr.) 'expel'
(26c)	<i>n-tsɔ-DU</i> (in.)	'fight to defend oneself'	<i>tsɔ</i> (in.) 'feel'
(26d)	<i>n-k'ɔt'i</i> (in.)	'see each other'	<i>k'ɔt'i</i> (tr.) 'look at someone'
(26e)	<i>n-təts'e</i> (in.)	'crash each other'	<i>təts'e</i> (tr.) 'stop'
(26f)	<i>n-theš-khi</i> (in.)	'like each other' (<i>khi</i> 'blood')	<i>təts'e</i> (tr.) 'raise'
(26g)	<i>n-thɔti</i> in.(tr.)	'marry'	<i>thɔti</i> (tr.) 'answer someone'
(26h)	<i>m-phɔts'i</i> in.(tr.)	'help each other'	<i>phɔt'i</i> tr. 'help'

The first two verbs in (26) are formerly derived verbs. The verb *n-tuhni* 'fight' (26a) is the only reciprocal verb in the list that cannot take dual morphology. The rest use the dual to mark the involvement of the other party in the reciprocal action. However, the verbs differ in the extent to which its application is obligatory. The verbs in (26b) and (26c) require the dual, while for the remaining ones it is optional. The two last verbs in (26) use this dual morphology to become transitive, as shown in (27):

(27a) *no=r ndə b̄i n-thɔti*
 DEF=SG man 3.PST M-answer someone

'The man got married'.

(27b) *no=r ndə b̄i n-thɔt-wi=r*
 DEF=SG man 3.PST M-answer.someone-DU.INCL=SG

?beh̄no
 woman

'The man married the woman'.

- (27c) *no=r ndə bì n-thqta=r ?behño
 DEF=SG man 3.PST M-answer.someone.B=SG woman
 ‘The man married the woman’.

In (27a), *n-thqti* is used as an intransitive, the basic usage of the form. The verb may additionally receive an object by means of dual morphology, as in (27b).³¹ The impossibility of (27c) illustrates that this morphology is necessary for the verb to be transitive.

7.2. Spontaneous middles. Middle morphology is commonly found in verbs that express spontaneous events. This is the case in Spanish, for example, as shown in (28):

- (28a) *Mario* *abr-i-ó* *la* *puerta*
 (name) open-THV-3SG.PST.IND DEF.SG.FEM door(FEM)
 ‘Mario opened the door’.
- (28b) *la* *puerta* *se=abr-i-ó*
 DEF.SG.FEM door(FEM) 3.REF/M=open-THV-3SG.PST.IND
 ‘The door opened’.

(28a) expresses a causative event where a person called Mario is involved in the volitional action of opening a door. The active verb in (28a) is a transitive verb that lexicalizes this causative scenario. (28b) shows a spontaneous event—the door opens by itself perhaps due to a draft or some internal mechanism. What is crucial here is that there is no perceptually salient agent behind the opening. To express this spontaneous event, Spanish uses middle morphology, the same morphology used for direct reflexives, hence the gloss in (28b). The resulting verbal form is intransitive.

The pattern shown in Spanish (28) is rather common cross-linguistically, according to Kemmer (1993) and Kazenin (2001). However, the pattern is not found in Otomi except for the two verbs in (29):

- | Middle | Source |
|--|---------------------------------------|
| (29a) <i>n-tušni</i> in. ‘be/get tangled’ | <i>tušni</i> tr. ‘entangle something’ |
| (29b) <i>n-k’ěšni</i> in. ‘crack (a bone)’ | <i>k’ěšni</i> tr. ‘crack a bone’ |

These verbs are the only ones in the sample that convey a spontaneous event by means of {*N-*}. They are also the only cases of middle-marked spontaneous events that I have been able to find in the language so far.³²

³¹ The dual here appears to work as an applicative. Study of its distribution and behavior is beyond the scope of this article.

³² Otomi has other nasal morphemes that resemble the middle marker. These two verbs denoting spontaneous events might be related to a different subclass of verbs altogether. This whole area needs further study.

More commonly, Otomi conveys spontaneous events with inherently intransitive or labile verbs (similar to English). The latter case is shown in (30); compare it to the Spanish example in (28):

(30a) *no=r Mario bì šoka=r gɔʃthi*
 DEF=SG (name) 3.PST open.B=SG door
 ‘Mario opened the door’.

(30b) *no=r gɔʃthi bì *({N-})šoki*
 DEG=SG door 3.PST (M-)open
 ‘The door opened’.

In (30a), *šoki* ‘open’ is used in a transitive frame and expresses a transitive event.³³ Without additional morphology, the verb may be used as an intransitive to render a spontaneous event, as in (30b). The use of middle morphology would yield an ungrammatical result, as the impossibility of {N-} in (30b) shows. (I do not explore such cases further because a proper study of this phenomenon is beyond the scope of this article.) Nevertheless, the reason behind the scarcity of spontaneous middles may be found in the way the language treats the phenomenon of transitivity in general. In Palancar (forthcoming), I study how syntactic valence is distributed in a sample of 300 verbs. The figures show that the majority of verbs are transitive, yet the morphology of such verbs reveals that intransitive and equipollent verbal roots were a pervasive pattern historically. Thus the scarcity of derived verbs to express spontaneous events can be explained by the weight labile and intransitive verbs had in the language. This factor could keep the middle voice of Otomi from extending into the syntactic realm of diathesis.

8. Conclusion. My purpose has been to show that Otomi has a nasal morpheme that is an exponent of middle voice and I have provided facts to support such an analysis. My study is based on a sample of 72 verbs in the language that display the morpheme. I have presented a number of verbs that clearly show the outcomes of a transparent derivational process, both at a semantic and at a morphological level (see 5 for discussion). However, my study also reveals that most middle verbs have become lexemes in the language, showing a morphological connection to their source verbs.

Otomi possesses the definitional characteristics of middle systems, which are well attested typologically. On the one hand, the semantics of Otomi middle verbs elaborate on a range of situation types which are treated by Kemmer (1993; 1994) and Maldonado (1999) as typical of middle systems. In contrast to other voice systems that involve diathesis, middle voice is

³³ The form *šoka=* is morphophonemically conditioned.

characterized by increased “involvement” of the subject (Maldonado [forthcoming]). Similarly, my sample also contains cases where the source of the derivation was difficult, if not impossible, to reconstruct (see 6). In this sense, Otomi also has a substantial number of underived middle verbs (also called *media tantum* verbs), which is again typologically to be expected. The application of middle morphology is highly an idiosyncratic phenomenon (Kemmer 1994:185). Otomi middle voice is also typologically consistent in that middle markers often co-occur with both transitive and intransitive bases, even though the middle verb as an outcome is overwhelmingly intransitive (90 percent). Finally and crucially, the morphological exponent of middle voice in SI Otomi is not a reflexive marker for several reasons: (a) true reflexive actions are constructed syntactically by different means; (b) over 90 percent of the verbs in the sample show extensions that go beyond the ones attested for reflexives, displaying the full range of functions typical of middle systems; (c) the process is no longer productive, so it cannot be used to create a reflexive verb. I hope this will contribute to understanding the scope of middle voice in Otomanguan languages. Much research remains to be done, both on other dialects of Otomi and on other languages of the family, to discover what similarities and differences exist.

APPENDIX A
DRESSING VERBS IN OTOMI

Subject's Own Domain		Extended Domain	
in.	tr.	di-tr.	
<i>n̄-he</i>	<i>he</i>	<i>hete</i>	‘wear x above waist’
<i>n-kə</i>	<i>kəde</i>	<i>kəte</i>	‘wear x below waist’
	<i>təde</i>	<i>təte</i>	‘wear x on breast/neck (i.e., necklace)’
	<i>t̄t̄’i</i>	<i>t̄t̄’i</i> -{DAT}	‘wear x on feet (i.e., shoes)’
	<i>phots’i</i>	<i>phots’i</i> -{DAT}	‘wear x through head (i.e., sweater)’
	<i>?bət’i</i>	<i>?bət’i</i> -{DAT}	‘wear x around waist (i.e., belt)’
	<i>h̄uts’i</i>	<i>h̄uts’i</i> -{DAT}	‘wear x on head (i.e., hat)’
	<i>pət’i</i>	<i>pət’i</i> -{DAT}	‘wear x over body (i.e., cloak, robe)’
	<i>kəts’i</i>	<i>kəts’i</i> -{DAT}	‘wear any other x (i.e., sheet, blanket, etc.)’
	<i>n-tsiti</i>	<i>tsiti</i> -{DAT}	‘put earrings on’
<i>m-phits’i</i>		<i>phits’i</i> -{DAT}	‘wear x on top of another piece of clothing’

APPENDIX B
DERIVED VERBS IN THE SAMPLE

	Middle Verb			Source Verb		
BA/BF ³⁴	<i>n-went'i</i>	in.	'swoon'	<i>went'i</i>	in.	'faint'
BA/BF	<i>ñ-?ei</i>	in.	'vomit'	<i>?ei</i>	tr.	'throw'
BA/BF	<i>n-čot'i</i>	in.	'scratch a lot'	<i>šot'i</i>	lab. (MS)	'scratch, peel'
BA/BF	<i>n-ts'oni</i>	in.	'urinate, defecate on'	<i>ts'oni</i>	lab. (MS)	'get dirty'
BA/BP	<i>n-khwats'i</i>	in.	'lean'	<i>khwats'i</i>	in.	'lean (a thing)'
BA/BP	<i>n-čint'i</i>	in.	'turn around'	<i>šint'i</i>	tr.	'turn around'
BA/GR	<i>n-te?is'e</i>	in.	'lick oneself (a cat)'	<i>te?is'e</i>	tr.	'lick'
BA/GR	<i>n-thuki</i>	in.	'have a cleansing'	<i>thuki</i>	tr.	'clean'
BA/GR	<i>m-phi?is'i</i>	lab.	'put one piece of clothing on top of another'	<i>phi?is'i</i>	tr.	'put on top'
BA/GR	<i>n-tšiti</i>	tr.	'put earrings on'	<i>tšiti</i>	tr.	'hang'
BA/GR	<i>ñ-?e-ño³⁵</i>	in.	'comb short hair'	<i>?e-ño</i>	in. (MS)	'comb long hair'
BA/GR	<i>ñ-?eš'i</i>	in.	'cut one's hair'	<i>?eš'i</i>	in. (MS)	'have a hair cut'
BA/GR	<i>ñ-he</i>	in.	'dress above waist'	<i>he</i>	tr. (MS)	'wear above waist'
BA/GR	<i>n-koš'i</i>	tr.	'rub cream'	<i>koš'i</i>	tr. (MS)	'massage'
BA/NTM	<i>n-zinzq³⁶</i>	in.	'swing'	<i>tsidi</i>	in.	'hang'
BA/NTM	<i>m-phelmi</i>	in.	'flutter wings'	<i>phelmi</i>	tr.	'hit (with a stick)'
BA/NTM	<i>m-phe?is'e</i>	in.	'trip'	<i>phe?is'e</i>	tr. (MS)	'trip'
BA/TM	<i>m-phet'i</i>	in.	'swim'	<i>phet'i</i>	tr.	'hit (a drum)'

APPENDIX B—continued

PS/COG	<i>n-ʔəni</i>	in.	'get informed'	<i>ʔəni</i>	tr.	'ask a question'
PS/COG	<i>n-tʃu</i>	in.	'be afraid'	<i>tʃu</i>	in. (MS)	'get frightened'
PS/COG	<i>m-beni</i>	in.	'think'	<i>beni</i>	tr. (MS)	'remember, think of something.'
PS/EM	<i>m-peʔts'i</i>	tr.	'mourn a dead person'	<i>peʔts'i</i>	tr.	'possess'
REC	<i>m-phəts'i</i>	in.	'help each other'	<i>phəts'i</i>	tr.	'help'
REC	<i>n-k'ət'i</i>	in.	'see each other'	<i>k'ət'i</i>	tr.	'look at someone'
REC	<i>n-təts'e</i>	in.	'crash with each other'	<i>təts'e</i>	tr.	'stop'
REC	<i>n-theš-khi-DU</i>	in.	'like each other'	<i>təts'e</i>	tr.	'raise'
REC	<i>n-thəti</i>	in.	'marry'	<i>thəti</i>	tr.	'answer to someone'
REC	<i>n-tʃə-wi</i>	in.	'fight to defend oneself'	<i>tʃə</i>	in. (MS)	'feel'
SP	<i>n-tušni</i>	in.	'get entangled'	<i>tušni</i>	tr.	'entangle'
SP	<i>n-k'əšni</i>	in.	'crack (a bone)'	<i>k'əšni</i>	tr. (MS)	'crack a bone'
MIS	<i>m-bəni</i>	in.	'shake (a turkey)'	<i>bəni</i>	in.	'exit'
MIS	<i>n-tʃi-the³⁷</i>	in.	'drink to get drunk'	<i>tʃi-the</i>	in.	'be drinking'
MIS	<i>m-ʔədi</i>	in.	'be/get lost'	<i>ʔədi</i>	tr.	'lose'
MIS	<i>n-k'hət'i</i>	in.	'struggle to get free'	<i>k'hət'i</i>	tr.	'pull horizontally'
MIS	<i>n-heni</i>	in.	'cut oneself'	<i>heni</i>	tr.	'chop'
MIS	<i>n-handi</i>	in.	'face toward a place'	<i>handi</i>	tr.	'look at/see'
MIS	<i>n-hoki</i>	in.	'get ready'	<i>hoki</i>	tr.	'fix'

MIS	<i>m-ʔi</i>	in.	'be born'	<i>ʔbi</i>	in. (MS)	'live, be'
MIS	<i>ñ-ʔoni</i>	in.	'hide away'	<i>ʔoni</i>	in. (MS)	'hide'
MIS	<i>m-paʔi</i>	in.	'get warm'	<i>paʔi</i>	tr.	'heat'
MIS/S-C	<i>n-zidi</i>	in.	'hang oneself'	<i>tsidi</i>	in.	'hang'
MIS/S-C	<i>n-sy</i>	in.	'take care of oneself'	<i>sy</i>	tr.	'take care'
MIS/S-C	<i>n-tʃiti</i>	in.	'weigh oneself'	<i>tsiti</i>	tr.	'hang'
MIS/S-C	<i>ñ-čoki</i>	in.	'uncover oneself'	<i>čoki</i>	tr.	'uncover someone'
MIS/S-C	<i>ñ-heʔts'e</i>	in.	'cover oneself'	<i>heʔts'e</i>	tr.	'cover someone'
MIS/S-C	<i>ñ-ho-s³⁸</i>	in.	'kill oneself'	<i>ho</i>	tr.	'kill/hit'
MIS/S-C	<i>m-pəte</i>	in.	'harm oneself'	<i>pəte</i>	tr. (MS)	'stumble upon something'

³⁴ Abbreviations used in Appendix B: BA = body action; BF = body function; BP = body posture action; COG = cognition middle; EMO = emotion middle; GR = grooming action; MS = middle semantics; MIS = miscellaneous; NM = nontranslational motion; PS = psychological middle; S-C = self-centered; SP = spontaneous; TM = translational motion.

³⁵ The morph *-ño* means 'head'.

³⁶ This verb is a compound: *-nza* is composed of the interfix *-n-* plus *-za* 'tree' as the second member of the compound.

³⁷ The morph *-the* is the reduced form of *dehe* 'water' and functions here as a liquid classifier.

³⁸ The morph *-s^e* was shown in (20b).

REFERENCES

- ANDREWS, HENRIETTA. 1993. *The Function of Verb Prefixes in Southwestern Otomi*. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- BERNAL PÉREZ, FELIPINO. 1986. *Hñähñu: Recopilación de 505 verbos otomíes conjugados*. Cardonal, Hidalgo, Mexico: Centro Social de Cardonal.
- CONSTABLE, PETER. 1990. Reflexives in Veracruz Huastec. Ms.
- ESTRADA, ZARINA. 2002. *La voz media en pima bajo*. Paper read at the VII Encuentro Internacional de Lingüística del Noroeste, Universidad de Sonora, Sonora, Mexico.
- FOX, BARBARA, AND PAUL J. HOPPER, eds. 1994. *Voice: Form and Function*. Amsterdam: John Benjamins.
- FRIEDRICH, P. 1969. On the meaning of the Tarascan suffixes of locative space. *IJAL Memoir* 23.
- GIVÓN, TALMY. 1981. Typology and functional domain. *Studies in Language* 5:163–93.
- HAIMAN, JOHN. 1983. Iconicity and economic motivation. *Language* 59:781–819.
- HEKKING, EWALD. 1995. *El Otomí de Santiago Mexquititlán: Desplazamiento Lingüístico, Préstamos y Cambios Gramaticales*. Amsterdam: Institute for Functional Research into Language and Language Use.
- HEKING, EWALD, AND SEVERIANO ANDRÉS DE JESÚS. 1984. *Gramática Otomí*. Querétaro, Mexico: Universidad Autónoma de Querétaro.
- HESS, H. H. 1968. *The Syntactic Structure of Mezquital Otomi*. The Hague: Mouton.
- KAZENIN, KONSTANTIN I. 2001. Verbal reflexives and the middle voice. *Language Typology and Language Universals*, ed. Martin Haspelmath et al., vol. 2, pp. 916–27. Berlin: Mouton de Gruyter.
- KEMMER, SUZANNE. 1993. *The Middle Voice*. Amsterdam: John Benjamins.
- _____. 1994. Middle voice, transitivity and the elaboration of events. *Voice: Form and Function*, ed. Barbara Fox and Paul J. Hopper, pp. 179–230. Amsterdam: John Benjamins.
- KLAIMAN, M. H. 1988. Affectedness and control: a typology of voice systems. *Passive and Voice*, ed. Masayoshi Shibatani, pp. 25–84. Amsterdam: John Benjamins.
- _____. 1991. *Grammatical Voice*. Cambridge: Cambridge University Press.
- _____. 1992. Middle verbs, reflexive middle constructions, and middle voice. *Studies in Language* 16:35–61.
- LASTRA, YOLANDA. 1989. *Otomí de San Andrés Cuexcontitlán, Estado de México*. *Archivo de Lenguas Indígenas*, no. 13. Mexico, D.F.: El Colegio de México.
- _____. 1992. *El otomí de Toluca*. Mexico, D.F.: Universidad Nacional Autónoma de México.
- _____. 1997. *El Otomí de Ixtenco*. Mexico, D.F.: Universidad Nacional Autónoma de México.
- _____. 1998. *Ixtenco Otomi*. Munich: Lincom Europa.
- LE CRON FOSTER, MARY. 1969. *The Tarascan Language*. UCPL 56. Berkeley and Los Angeles: University of California Press.
- LYONS, JOHN. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- MALDONADO, RICARDO. 1992. *Middle voice: the case of Spanish 'se'*. Ph.D. dissertation, University of California, San Diego.
- _____. 1999. *A Media Voz: Problemas conceptuales del clítico SE*. Mexico, D.F.: Universidad Nacional Autónoma de México.
- _____. Forthcoming. *Grammatical Voice*. *Handbook of Cognitive Linguistics*, ed. Dirk Geeraerts and Hubert Cuyckens. Oxford: Oxford University Press.
- MEL'ČUK, IGOR A. 1993. The inflectional category of voice: towards a more rigorous definition. *Causatives and Transitivity*, ed. Bernard Comrie and Maria Polinsky, pp. 1–46. Amsterdam: John Benjamins.

- NAVA, FERNANDO, AND RICARDO MALDONADO. Forthcoming. Basic voice patterns in Tarascan. *Language, Culture and Mind*, ed. Michel Archard and Suzanne Kemmer. Los Angeles: CSLI Publications.
- PALANCAR, ENRIQUE. 2003. Middle voice and intransitivity in Otomi. Ms.
- _____. Forthcoming. Transitividad en el Otomí de San Ildefonso: una incursión tipológica en lenguas otomangues. Proceedings of the VII Encuentro Internacional de Lingüística del Noroeste. Sonora, Mexico: Universidad de Sonora.
- SHIBATANI, MASAYOSHI. 1985. Passives and related constructions. *Language* 61:821–48.
- _____. 2002. Synchronic typology and diachronic explanations: The case of middle voice. Paper presented at the VII Encuentro Internacional de Lingüística del Noroeste, Universidad de Sonora, Sonora, Mexico.
- SMYTH, H. W. 1920. *Greek Grammar*. Revised by G. M. Messing. Cambridge, Mass.: Harvard University Press, [Reprint ed., 1956.]
- SOUSTELLE, JACQUES. 1937. *La Famille Otomi-Pame du Mexique Central*. Paris: Institute d'Ethnologie.
- TRASK, R. L. 1993. *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.
- VOIGTLANDER, KATHERINE, AND ARTEMISA ECHEGOYEN. 1985. *Luces Contemporáneas del Otomí: Gramática del otomí de la sierra*. Mexico, D.F.: Instituto Lingüístico de Verano.

3. Travaux sur aspects de la phonologie otomie

- TRAVAIL 1. Palancar, Enrique L. 2013. Aspiration in Northern Otomi 369
and the relic of an old fortis/lenis contrast. Jean Léo
Léonard et Samia Naïm (eds.), *Base Articulatoire Arrière.
Backing and Backness*. Munich : LINCOM Europa.

Preaspiration in Northern Otomi: A relic of an old fortis/lenis contrast

Enrique L. Palancar (SeDyL-CELIA CNRS UMR 8202,
Surrey Morphology Group, University of Surrey)

1. Introduction¹

In this paper I show what can happen to the aspiration feature of a preaspirated series of stops /^hp, ^ht, ^hk/ once such stops are reanalyzed as simple stops /p, t, k/. More specifically, the case study is on the phonetics-phonology interface of Northern Otomi, a Mesoamerican language of the Otomi family (Oto-Pamean, Oto-Manguean). In this respect, I focus on the distribution of an intersyllabic aspiration [h] which occurs in coda of a word-internal syllable, such as for example in the words [7bɛh.pɔ] 'sister-in-law' or [mɔh.k'ɔ] 'priest'. One of the goals of the paper is to provide the possible constraints and rules that determine when this aspiration is phonetic and when it better characterized as phonological, because as innocuous as it may be, it provides important clues as to how to characterize the phonology of the nucleus of a Northern Otomi syllable. It further represents a backing phenomenon that serves as a keystone to understand the history and development of the consonantal system of Otomi. In this way, I interpret intersyllabic aspiration as the only remnant of the aspirated feature of preaspirated plosives in Old Otomi, which got otherwise lost in initial position, so that a word such as [tɪh.ti] 'encourage someone' comes from Old Otomi *^htɪ^hti.

1.1. Introduction to Otomi

Otomi is a group of very closely related languages spoken in Mexico which has emerged as a small linguistic family in recent times. While its internal diversity is uncontroversial (summarized in Lastra 2001), there is no established consensus as to which of the different linguistic subsystems identified so far as Otomi should be treated as independent languages and which as dialects of those languages. This is why Otomi has traditionally been talked about as forming a dialectal continuum rather than a family and why authors in Mexico avoid using terms such as "language" and "dialect" and instead use "variety" as a cover term. Against this inertia, the official stance by the National Institute of Indigenous Languages (INALI) in the CLIN (2008) is to treat Otomi innovatively as forming a "linguistic group"

¹ Acknowledgments: This paper was written under the auspices of the ESRC/AHRC project RES-062-23-3126 "Endangered Complexity: Inflectional classes in Oto-Manguean languages". I heartily thank the editors of this volume for their generous invitation to participate in the International Colloquium 'Backing/Base articulaire arrière' celebrated in Paris the 2-4 May 20012 at the Maison de la Recherche of the New Sorbonne, Paris 3, but especially to Jean-Léo Léonard who has both patiently and diligently provided me with very important clues as to how to make sense of this phenomenon. The paper benefited immensely from the comments, discussions and questions by the participants. All errors and deficiencies remain my responsibility.

with nine different "linguistic varieties". INALI's linguistic varieties are treated as genuine languages for official purposes. This is the stance I take in this paper although with my own understanding of the current dialectal situation.

In this respect, based on previous proposals in Soustelle (1937), Lastra (2001, 2010), Andrews *et al.* (1983 [1978]) and on my own field observations, the different Otomi varieties spoken nowadays could be seen as grouped in four geographical areas representing six different languages: Northern Otomi, Eastern Otomi, Western Otomi and three Southern languages (Ixtenco Otomi, Ocoyoacac Otomi and Tilapa Otomi). While the Southern languages are spoken in small communities with relative uniformity, the other first three constitute dialectal continua which include different linguistic varieties as dialects. Northern Otomi comprises, in this view, the varieties spoken in the central part of Hidalgo, Querétaro, and the north of the State of Mexico, including thus Mezquital Otomi, Santiago Mexquititlán Otomi, Tolimán Otomi, Tierra Blanca Otomi, Guanajuato Otomi, Acambay Otomi and San Ildefonso Tultepec Otomi. The present paper is based on the variety of San Ildefonso Tultepec Otomi.

The consonant chart of phonemes existing in Northern Otomi from the variety of San Ildefonso Tultepec is given in Table 1.² This is a large inventory consisting of 39 units (excluding allophones).

	Lab.	Den.	Alv.	Pal-Alv.	Pal.	Vel.	Glott.
Stop	p	t	d			k	g
Ejective		t'				k'	
Aspirated	p ^h	t ^h				k ^h	
Labial						k ^w	
Aspirated-Labial						k ^{hw}	
Affricates		ts		tʃ			
Ejective		ts'		tʃ'			
Fricatives			s	z	ʃ		h
Flap			r				
Ejective			r'				
Trill			r				
Liquid			l				
Nasals	m ³		n				
Ejective	m ³		n				
Approximants	ã	w		j	j		
Ejective		w'		j	j		

Table 1. Consonant phonemes in San Ildefonso Tultepec Otomi, Northern Otomi.

For the simple stops, there is a contrast based on sonority, i.e. /p, t, k/ vs. /b, d, g/. The different examples in (1) give instances of contrastive structures, sometimes with minimal pairs when they exist, which may additionally include loanwords from Spanish.

² The language has three tones. Two level tones (high, represented by an acute accent and low, not indicated) and an ascending tone.

³ For the notation of voiceless nasals and approximants in Northern Otomi, I follow here Knapp's (2008) treatment of similar segments in Mazahua, which is closely related to Otomi. This treatment follows Maddieson's (1984) typology of sounds in the world's languages. For Otomi, the segments are often articulated as preaspirated nasals, but not always.

(1) /p/		/b/	
pa	'day'	bá	{3.PST.HITHER}
páa	'shovel' (Sp. pala)	báa	'bullet' (Sp. bala)
páda	'culture'	bātsi	'child'
pōnt'i	'blessing'	bəhɔi	'mud'
/t/		/d/	
tó	'mother-in-law'	dō	'stone'
téhé	'cough'	déhe	'water'
tahtá	'father'	dada	'old man'
tɔi	'market'	dɔ	'eye'
/k/		/g/	
ka	'in, on, at'	ga	{1.PRES.IRREALIS}
kūi	'chase'	góné	'dumb'
kít'a	'5'	gō	'twin'
kódo	'turkey'	góhó	'4'

For the purpose of the present study, it is important to elaborate a little on syllabic structure and coda position in Otomi. While it is relatively common that Oto-Manguan languages have a syllabic structure CV, as for Otomi, this is not the only existing structure. The old syllabic structure CV is still recognizable in historical lexical roots in Otomi—an aspect which is only available for language specialist. Word internally, however, open syllables, such as (2a) alternate with close syllables in the making of the sound structure of a word. Close syllables in Otomi can be simple, as in (2b), or they can be complex, as in (2c). These codae emerge from morphotactic adjustments at morpheme boundaries, but for many of the simple codae at least, the morphemes giving rise to them are now unavailable to the common speaker for both segmentation and semantic accessibility, i.e. they have become part of the phonology of the word. In all examples, syllable boundary is indicated by a dot.

(2) a.	/beni/	CV.X	'remember'
	/boʃi/	[be.ni]	'rooster'
	/dese/	[bo.ʃi]	'grasshopper'
b.	/nsabdo/	CVC.X	'Saturday'
	/zeʃk'o/	[nsab.do]	'trousers'
	/mʔep'e/	[zeʃ.k'o]	(a community role)
c.	/handgi/	CVCC.X	'sec me'
	/toʔmhe/	[hand.gi]	'we wait'

Finally, most lexical words in Northern Otomi are disyllabic in prosodic structure. For this type of words, there is a default stress rule by which stress falls on the first syllable, as in (3). In these examples, as elsewhere, stress is indicated with underlining.

(3)	pō.n't'i	'blessing'
	bə.hɔi	'mud'
	dé.he	'water'
	dá.da	'old man'
	ké.t'a	'5'

kó.do 'turkey'
 gó.hó '4'
 etc.

As the rule has a default application across the lexicon, deviating patterns like the one illustrated by the words in (4), should be accounted for as being lexically specified.

(4) pá.da 'vulture'
 bá.ṣá 'child'
 gó.ṣé 'dumb'

Most of the words with more than one syllable which are still morphologically simple are loanwords, and as such they commonly receive the stress of Spanish, like in (5), whose stress pattern is coherent with the native one in (3).

(5) á.bwé.ṣá.tá 'granny'

1.2. The problem

The main object of study in this paper is the distribution of an [h] that commonly occurs at coda of a word internal syllable in the neighbourhood of the voiceless stops in (6). The distribution is at times puzzling, and it constitutes a problem for the linguistic treatment of such segments. For example, this [h] typically occurs before another syllable which has a simple voiceless plosive as its onset, as in the words in (1). In words like these, the presence of [h] is obligatory; its absence is indicated by an underscore. However, it never appears before voiced plosives, as shown in (7).

(6) [ʔbá.po] / *[ʔbɛ.po] 'sister-in-law'
 [dúh.tu] / *[dú.tu] 'clothes'
 [dá.ṣṣá.ká] / *[dá.ṣṣə.ká] 'I arrived (there) myself'

(7) *[háh.bi] / [há.bi] 'where'
 *[k'áh.dí] / [k'á.dí] 'finish'
 *[n.ʔáh.gí] / [n.ʔá.gí] 'once'

However, in some words like in (8a), the [h] cannot occur, and in some others, like in (8b), it may or may not occur; a phenomenon commonly found in the pronunciation of loanwords. Alternating pronunciations are indicated by the tilde.

(8) a. *[ṣáh.tá] / [ṣá.tá] 'saint, god'
 b. [sóh.pa] ~ [só.pa] 'soup' (Sp. *sopa*)
 [kóh.ka] ~ [kó.ka] 'coke' (Sp. *coca*)

In some cases, the [h] appears before the simple plosive at a complex coda within the same syllable, as in (9).

(9) [gá.t'óh.t.ki] 'you tied me up'
 [ṣ'úh.k.dó] '(female) turkey'

Finally, this [h] does not occur before aspirated consonants, as shown in (10). However, while its absence is expected in morphologically complex sequences like (11a), in a number of them it is puzzlingly obligatory, as shown in (11b).

(10) *[jóh.p'ó] / [jó.p'ó] 'harvest maize'
 *[déh.t'ó] / [de.t'ó] 'maize'
 *[bóh.k'ó] / [bó.k'ó] 'money'

(11) a. *[gá.t'óh.k'é] / [gá.t'ó.k'é] 'you let us in'
 *[gá.t'óh.t'á] / [gá.t'ó.t'á] 'you (pl) tied him up'
 b. [dá.t'óh.k'é] / *[dá.t'ó.k'é] 'we let him in'
 [gá.t'óh.t'á] / *[gá.t'ó.t'á] 'you (pl) answered him'

This data are challenging for a phonological account of the rise of aspiration at syllabic boundary, and accounting for them is important at least for the study of Otomi. In this paper, I advance a number of proposals for their analysis, but a comprehensive formal account for them is still needed. The paper is structured as follows. First, in the next section I introduce those cases where I take [h] to be of a phonetic origin, to be contrasted in section 3 with those instances of [h] which I take to be a phonic constituent part of a complex nucleus. In section 4, I study a number of cases where the distribution of [h] proves a little more complex because it involves the morphotactic involvement of object and plural affixation. Section 5 concludes with a brief excursion on the origin of preaspiration.

2. Preaspiration: context where [h] is phonetic

One very simple (and quick) way to account for the distribution of [h] is the examples presented in the previous section is to state that [h] in such a dataset is just another instantiation of the same phoneme /h/ we find elsewhere in the lexicon, such as for example in the word in (12), but in coda position. Under such a treatment, the words in (13) would have a phonological /h/.

(12) [há.ho] /há.ho/ 'vixen'

(13) [k'óh.pi] /k'óh.pi/ 'bless'
 [dúh.tu] /dúh.tu/ 'clothes'
 [éh.ki] /éh.ki/ 'pluck'

Nevertheless, the type of phonological treatment in (13), no matter how simple it may be, overlooks entirely an important factor, that is, that the occurrence of [h] is predictable at least for most of the cases. To such an extent is this predictability expected, that in 1999 the Commission for the Indigenous Peoples of Querétaro – a body in charge of designing a popular orthography for the varieties of Northern Otomi spoken in Querétaro – ruled against giving a written representation to the instances of [h] for such words, under the assumption that its occurrence is predictable from the phonetics of syllabic structure. This is shown, for example in the spelling of the words in (14).

(14) 'bego [ʔbɛ.po] 'sister-in-law'
 dátu [dúh.tu] 'clothes'
 dá tsoka [dá.ṣṣá.ká] 'I arrived (there) myself'

Treating [h] as a phonetic insertion yields a more complicated account, but a simpler phonology. Let us for the sake of argument, pursue the phonetic analysis here. Under this analysis, I will treat many of the instances of [h] as the outcome of a phonetic rule I call "preaspiration", which involves the insertion of the glottal fricative [h] before a voiceless simple stop. There are two contextual rules for preaspiration, which can account for most of the instances for the occurrence of [h] in Northern Otomi. These two rules are stress dependent, very much in the fashion of Verner's law for the stops in Indo-European, which cancelled out the application of Grimm's law, accounting thus for its exceptions.⁴

The first contextual rule for preaspiration, which has a very wide application throughout the lexicon, is spelled out in (15).

(15) **Preaspiration rule 1:** When /p, t, k/ occur as the onset of the weak syllable of a trochaic word, the strong syllable obligatorily receives an aspiration in its coda in the form of an [h]. As an epiphenomenal effect, the inserted glottal fricative [h] precedes the voiceless obstruent, hence the term 'preaspiration'.

The contextual rule in (15) accounts for similar data to the ones given in (16).

(16)	/CV_p.t.kV/	[CVh.p.t.kV]	*[CV_p.t.kV]
/p/	/k'ò.pi/	[k'hò.pi]	/'bless'
	/ʔbɛ.po/	[ʔbɛ.po]	/'sister-in-law'
/t/	/dú.tu/	[dú.tu]	/'clothes'
	/p'ó.to/	[p'hó.to]	/'8'
	/m'á.te/	[m'há.te]	/'behind'
/k/	/t'i.ki/	[t'hi.ki]	/'pluck'
	/ʔo.ki/	[ʔo.ki]	/'hole'

With loanwords like in (17), there are two possibilities: speakers may either filter them through their native phonology and hence, the words receive preaspiration, or they may not, and thus they surface without it, like in Spanish.

(17)		Native ph.	Spanish ph.
/CV_p.t.kV/	[CVh.p.t.kV]	[CV_p.t.kV]	[CV_p.t.kV]
/p/	/só.pa/	[só.pa]	/'soup' (Sp. <i>sopa</i>)
/t/	/á.bwé.lí.tá/	[á.bwé.lí.tá]	/'granny' (Sp. <i>abuelita</i>)
/k/	/kó.ka/	[kó.ka]	/'coke' (Sp. <i>coca</i>)

When the prosodic structure of the word is different and does not serve as an input for the rule, there is no preaspiration. This accounts for most exceptions, as for example (18).

(18)	/CV_p.t.kV/	*[CVh.p.t.kV]	[CV_p.t.kV]
	/sá.tá/	*[sá.h.tá]	/ [sá.tá]
	/nó.tí.sjá/	*[nó.h.tí.sjá]	/ [nó.tí.sjá]
			/'saint, god'
			/'news' (Sp. <i>noticia</i>)

⁴ I want to thank Jean-Léo Léonard for this observation.

Similarly, as expected by rule 1 no preaspiration occurs when the target word is preceded by a proclitic of TAM and person of the subject, like in (19). This is because the stress falls in the first syllable of the verbal stem, which for most cases corresponds to the lexical root.⁵

(19) a.	/ga.tí.ki/	b.	/bi.ká.di/
	[ga.tí.ki]		[bi.ká.di]
	ga = tík-i		bi = kád-i
	1. PRS.IRR=pluck-plant[3OBJ]-F		3. PST=deceive[3OBJ]-F
	'I'll pluck it.'		'He deceived him.'

The second contextual rule for preaspiration has scope over a few structures and it is spelled out in (20).

(20) **Preaspiration rule 2:** When /p, t, k/ occur as the coda of the closed syllable bearing the stress, that syllable receives [h] after the nucleus.

The contextual rule in (20) is illustrated in (21) with the surface form of the nominal compound [tjú+kódó {female+turkey}, which bears the stress on the first syllable and surfaces as /tjúk.dó/ by the reduction of the tonic nucleus of the second member with subsequent resyllabification.

(21)	/CV_p.t.k.V/	[CVhp.t.k.CV]	*[CV_p.t.k.CV]
	/tjúk.dó/	[tjúk.dó]	/'(female) turkey'

3. Phonetic [h] vs. phonological H

In contrast to the phonetic [h] as an outcome of the preaspiration rules presented in the previous section, in Northern Otomi there are also instances of [h] which I take to be realizations of a phonological 'phonological H'. There are two types of phonological H. A first one is a phonemic /h/, as in the words in (22), where it works as a consonantal segment.

(22)	[hó.tí:]	/hó.tí:/	'carry'
	[nhó]	/nhó/	'be good'

The other type of phonological H is when phonetic [h] is a constituent phone of a complex nucleus of the structure /vhV/ (alternatively also as /vh/ at a morphological boundary), as in (23).

(23)	[ʔó.hó]	/ʔóhv/	'sleep'
	[mʔá.há]	/mʔáhV/	'bathe'
	[d'é.hé]	/d'éhV/	'water'
	[mú.hú]	/múhV/	'wolf'
	[tí.h]	/tíhV/	'run'

When the complex nuclei in (23) are syllabified, they surface like words with a phonemic /h/ in internal position, like the ones in (24), and could thus be confused with them. The

⁵ Abbreviations: B bound stem; EMPH emphatic; EXCL exclusive; F free stem (a pausal form or pausal shape); IMPF imperfect; IRR irrealis; OBJ object; PL plural; PREP preposition; PRS present; PST past.

difference between these two types of phonemic H lies in the phonological structure of the word, i.e. complex nuclei always involve a rearticulated vowel, so that the value of 'V' in the structure /vhV/ [v.hV] is always the value of 'v'. This is not the case of words like in (24).

- (24) [nɛ́.hi] /nké.hi/ 'be flavourful'
[há.ho] /há.ho/ 'vixen'

Phonetic /h/ occurs at times in the same contexts as phonological H, and I provide here ways to distinguish them. In principle, a word-internal [h] in coda position should be taken to be a realization of a phonological H if it cannot be accounted for as an outcome of a preaspiration rule.

Phonological H of the phonemic /h/ subtype only occurs in a handful of cases: so far, I have been able to identify only two lexemes. Consider for this purpose the word [tá.h.tá] 'father' in (25a). Due to its stress, if the word had the phonological structure /tá.tá/, its phonetic outcome would be *[tá.tá] like [tsi.tá] 'saint' in (25b) with no [h], for /tá.tá/ like /tsi.tá/ are no inputs to preaspiration rule 1, like for example /dú.tu/ 'clothes' in (25c) is. Instead, the word is pronounced as [tá.h.tá], so the [h] represents a phonemic /h/.

- (25) a. /tá.h.tá/ [tá.h.tá] / [tá.tá] 'father'
b. /tsi.tá/ *[tsi.tá] / [tsi.tá] 'saint'
c. /dú.tu/ [dú.h.tu] / *[dú.tu] 'clothes'

The other word is the lexeme for 'priest' [mɔ́.h.k'ò]. This word also bears a phonological /h/ because its occurrence as a phonetic [h] is not expected because aspirated consonants do not trigger preaspiration. This was already shown in (10) above, more examples appear in (22).

- (26) /p' / /pé.p'h/ * [pé.h.p'h] [pé.p'h] 'work'
/p'o / /p' / [p' /] [p' /] 'guts'
/t' / /tə.t'e / * [tə.h.t'e] [tə.t'e] 'heal'
/d' / /dē.t'ò / * [dē.h.t'ò] [dē.t'ò] 'maize grain'
/gwa.t'ò / * [gwa.h.t'ò] [gwa.t'ò] 'on foot'
/l.t'a / * [l.h.t'a] [l.t'a] 'shoulders'
/k' / /ni.k'ò / * [ni.h.k'ò] [ni.k'ò] 'church'
/ndē.k'ò / * [ndē.h.k'ò] [ndē.k'ò] 'good afternoon'

In this way, the word for 'priest' should have been *[mɔ́.k'ò] had its phonological structure been */mɔ́.k'ò/, but instead it is [mɔ́.h.k'ò] with a phonemic /h/: (i) as the word has the wrong stress, it should have not received a phonetic [h]; and (ii) even with the right stress, [h] should not have occurred before an aspirated stop.⁷

⁶ The word for 'father' in Northern Otomi comes from the Old Otomi word for 'sir, old man' *hɛ́.há; itself a reduplication of the old word for 'father' *hɛ́, still present in more conservative languages.

⁷ The word for 'priest' in Northern Otomi appears to be a lexicalization of an old compound consisting of a particular bound stem for the verb 'to sell' /mɔ́- / (with a complex nucleus) plus the noun for 'god' /k'ò/. As the first morpheme in the compound is no longer accessible, the original phonetic [h] of the complex nucleus was reanalysed as an instance of a phonemic /h/ in coda position.

Other instances of a phonological H involve instances of stems with complex nuclei at morphological boundaries, like in (27), illustrating the verb for 'arriving here' /tsʰhV/, whose bound stem is /tsʰh- /.

- (27) [dá.tsʰh.ká] / * [dá.tsʰh.ká]
dá = tsʰh = ká
I.PST=arrive.there.I.EMPH
'I arrived (there) myself.'

In principle, the occurrence of [h] in instances like (27) could be taken to be the expected and uncontroversial outcome of preaspiration rule 1, so that a structure such as /CV.kV/ is naturally rendered by the phonetics as [CV.h.kV] and not as *[CV.kV]. This is the analysis I adopted in Palancar (2009), but I propose here that this previous analysis can be improved in the form of a phonological H. This new analysis is superior in that it is able to accommodate for a number of exceptions spelled out in Palancar (2009: 44 ff).⁸ Under this analysis, verbs with complex nuclei like 'arriving here' in (27) are taken to have two stems, one free and another bound, like in (28).⁹

- (28) FREE STEM BOUND STEM
tsʰh.hə tsʰh- 'arrive there'
tò.hò tòh- 'win, beat'
já.hə jáh- 'bathe'
hé.hə héh- 'sneeze'

To the verbs in (28) one can add verbs like the ones in (29), whose free stem does not have a complex nucleus.

- (29) FREE STEM BOUND STEM
p'ò.dì p'òh- 'know'
m'ò.dì m'òh- 'love'
ts'ò.gì tòh- 'pass by'
ʔù.gì ʔùh- 'sweeten'
j'ò.p'ò j'òh- 'pick maize'
p'e.p'hì p'eh- 'work at'
ʔb'ðì ʔb'ðh- 'stand'
ʔb'ì ʔb'ìh- 'live'

One further argument that the [h] in such bound stems is a phonological H comes from the fact that the segment is still present before other types of consonants in the onset of the following syllable, which would not trigger preaspiration, like in (30), if compared with (31).

- (30) [mí.ʔb'ìh.ní] / * [mí.ʔb'ì.ní]
mí = ʔb'ìh = ní
3.I.MPF=live.there.B=there
'He was living there.'

⁸ I also thank Jean-Léo Léonard for suggesting to me the complex nucleus hypothesis, which makes much sense to me now.

⁹ The bound stem is further responsible for the unvoicing of the onset of the enclitic. With other stems with simple nuclei, the enclitic = ká in (27) surfaces as =gá, as in dá = ní = gá {I.PST=see.3OBJ=I.EMPH} 'I saw it myself'.

- (31) [mí.pá.ní] / * [mí.páh.ní]
 mí = pa = ni
 3.IMPF=go=there
 'He was going there.'

Furthermore, in the context of word boundary, the glottal component in the complex nuclei of the bound stems of the verbs in (28-29) is realized as [ʔ] by default, like in (32), and only as an [h] when the word following the verb has an onset simple stop, like in (33). In (32), the [h] of complex nucleus shares could be confused as an outcome of preaspiration rule 1, because of its context, when in reality it is not a phonetic [h].

- (32) [mí.tʂóʔ.níʔ]
 mí = tʂóh = níʔ
 3.IMPF=pass.by.B=outside
 'He was passing by from outside.'
- (33) [mí.tʂóh.ka] [gɔʔʰi]
 mí = tʂóh = ka gɔʔʰi
 3.IMPF=pass.by.B=PREP door
 'He was passing by the door.'

In this section, I have shown cases where a given word-internal [h] should be better analysed as instantiating a phonological H and not as the output of the phonetic rule of preaspiration. There remain, however, a number of puzzling forms with the same verb, which carry a semantic contrast and involve the presence or the absence of [h], like the examples in (34) and (35) with the transitive verb /tʰó.ki/ 'let somebody in', which is cognate to the intransitive verb /tʰó.gi/ 'pass by' in (29).

- (34) [gá.tʰó.kʰé] / * [gá.tʰó.h.kʰé]
 gá = tʰó.kʰé
 2.PST=let.in+IOBJ+PL.EXCL
 'You let us in.'
- (35) [dá.tʰó.h.kʰé] / * [dá.tʰó.kʰé]
 dá = tʰó.h.kʰé
 1.PST=let.in[3OBJ]+PL.EXCL
 'We let him in.'

Examples like these result from the morphotactic adjustments involving the hosting of a number enclitic and they will be studied in further detail in the next section. Northern Otomi has two number enclitics, =hú for plural in general and =hé for plural exclusive. However, as these markers can be equally used to have scope over the subject or over the object, they need to be treated separately, for the adjustments they trigger are different in each case.

4. Preaspiration and plural marking

4.1. Plural of subject

Plural of subject in Northern Otomi is realized by means of the enclitics =hú or hé, like in (36). The latter marker is only used to express plural exclusive.

- (36) tsán- 'tell somebody off'
 [gá.tʂá.p.hí]
 gá = tsán = hí
 2.PST=tel.off[3OBJ]=PL
 'You (pl) told him off.'

The bound stem of a given verb may undergo morphotactic adjustments when the verb hosts a plural enclitic. The adjustments involve its coda in contact with the onset consonant of the enclitic. There can simplification of the glottal component of the complex nucleus in order to avoid gemination, as in (37), or there can be anticipatory assimilation of the coda of the stem, whose outcome is given in (38), and then this new outcome undergoes subsequent syllabification.

- (37) a. tóh- 'win/beat'
 [gá.tó.hí]
 gá = tóh = hí
 2.PST=win=PL
 'You (pl) won.'
- b. pch- 'work at'
 [gá.pch.hí]
 gá = pch = hí
 2.PST=work[3OBJ]=PL
 'You (pl) worked at it.'
- (38) a. hwáng- 'fall'
 SYLL. [gá.hwá.nkʰi]
 ASSIM. [gá.hwánk.hí]
 gá = hwáng = hí
 2.PST=fall=PL
 'You (pl) fell.'
- b. hand- 'see'
 [gá.ha.nkʰi]
 [gá.hant.hí]
 gá = hand = hí
 2.PST=see[3OBJ]=PL
 'You (pl) saw it.'

When the stem of the verb ends in a simple stop, it serves as input for the preaspiration rule 2 given in (20), as in (39). The outcome undergoes subsequent syllabification.

- (39) a. tsik- 'get somebody down'
 SYLL. [gá.tʂih.kʰi]
 PREASP. [gá.tʂihk.hí]
 gá = tsik = hí
 2.PST=get.down[3.OBJ]=PL
 'You (pl) got him down.'
- b. tʰót- 'respond to somebody'
 [gá.tʰót.hí]
 [gá.tʰóht.hí]
 gá = tʰót = hí
 2.PST=respond[3.OBJ]=PL
 'You (pl) responded to him.'

4.2. Marking of object and plural of object

The encoding of first person object is relevant for the discussion. The default exponent for this person value is the suffix -g, as in (40).

- (40) a. tsá 'bite'¹⁰
 [gá.tʂa.g-i]
 gá = tsá-g-i
 2.PST=bite-1.OBJ-F
 'You (sg) bit me.'
- b. hand- 'see'
 [gá.hand.g-i]
 gá = hand-g-i
 2.PST=see-1.OBJ-F
 'You (sg) saw me.'

¹⁰ Verbs with monosyllabic /CV/ stems have a derived bound stem with a complex nucleus /CVh-/ serving as a ditransitive stem, i.e. [gá.tʂah.ki] gá = tsah-g-i {2.PST=bite.DTR-1.OBJ-FREE.SHAPE} 'You bit him affecting me.' For another instance of the voiceless assimilation of the object suffix, see example (41).

With other verbs, such as for example verbs with complex nuclei, the suffix undergoes progressive assimilation in contact with the glottal component of the nucleus, as in (41). The phonetic outcome looks like an output of preaspiration rule 1, when in reality it is not.

- (41) tōh- 'win/beat'
 [gá.tōh.ki]
 gá = tōh-g-i
 2.PST=beat-B-1.OBJ-F
 'You beat me.'

On the other hand, the plural number of a first person can be realized by means of adding the appropriate number enclitic (for other encoding options see Palancar, 2009: 260ff). This triggers anticipatory assimilation on the part of suffix in contact with the onset of the enclitic, and both units surface as an aspirated stop in onset. This is shown in (42). Notice that aspirated stops do not serve as inputs for preaspiration, and hence this outcome does not get a phonetic [h].

- (42) tsá 'bite'
 [gá.tsa.k'é] / *[gá.tsa.h.k'é]
 gá = tsa-g = hé
 2.PST=bite-1.OBJ=PL-EXCL
 'You (sg) bit us.'

However, preaspiration occurs with the right input, as in (43), where we have the application of rule 2. For other apparently similar cases, where an [h] is present, this [h] realizes the glottal component of the complex nucleus of a bound stem, i.e. a phonological H, as in (44).

- (43) a. t'ót- 'respond to somebody' b. t'ót- 'respond to somebody'
 PREASP. [gá.t'óht.ki] [gá.t'óht.khé]
 gá = t'ót-g-i
 2.PST=respond-1.OBJ-F
 'You responded to me.'
 gá = t'ót-g = hé
 2.PST=respond-1.OBJ=PL-EXCL
 'You responded to us.'

- (44) tōh- 'win/beat'
 SYLL. [gá.tōh.k'é]
 PREASP. [gá.tōhk.h'é]
 gá = tōh-g = hé
 2.PST=beat-B-1.OBJ=PL-EXCL
 'You beat us.'

With verbs whose bound stem has a velar in coda, like in (45) there is again elision of one of the segments to avoid gemination.

- (45) a. t'ók- 'let somebody in' b. tsóg- 'leave'
 [gá.t'ók-g-i] [gá.tsóh.k'é]
 gá = t'ók-g-i
 2.PST=let.in-1.OBJ-F
 'You let me in.'
 gá = tsóg-g-i
 2.PST=leave-1.OBJ-F
 'You left me.'

When we add a plural enclitic, the outcome of elision and assimilation does not serve in turn as an input for the preaspiration rule because of the aspirated character of the consonant in the onset of the following syllable.

- (46) a. t'ók- 'let somebody in' b. tsóg- 'leave'
 PREASP. _ / *[gá.t'ók.h.k'é] _ / *[gá.tsóh.k'h'é]
 ASSIM. [gá.t'ók.khé] [gá.tsóh.khé]
 ELIS. [gá.t'ók.ghé] [gá.tsóh.ghé]
 gá = t'ók-g = hé gá = tsóg-g = hé
 2.PST=let.in-1.OBJ=PL-EXCL 2.PST=leave-1.OBJ=PL-EXCL
 'You let us in.' 'You left us.'

The outcomes in (46) contrasts neatly with those having plural of subject as presented above, where preaspiration from rule 2 precedes syllabification in the generation of surface structure.

- (47) t'ók- 'let somebody in'
 SYLL. [dá.t'ók.h.khé]
 PREASP. [dá.t'ók.h.k'é] / *[dá.t'ók.h.k'é]
 dá = t'ók-g = hé
 1.PST=let.in[3.OBJ]=PL-EXCL
 'We let him in.'

However, the type of morphotactic adjustments of the stem to a suffix appears to operate at a lower level than those associated to an enclitic. The ones associated with a suffix may serve as inputs for preaspiration, as for example with the verbs whose stem has an ejective stop, as in (48). For such verbs, the ejective stop is first simplified to a simple stop. Then the suffix undergoes assimilation and then we have the application of preaspiration rule 2.

- (48) mīt- 'grab'
 PREASP. [gá.mīht.ki] / *[gá.mīht.ki]
 ASSIM. [gá.mīt.ki]
 SIMPL. [gá.mīt.gi]
 gá = mīt-g-i
 2.PST=grab-1.OBJ-F
 'You grabbed me.'

However, the same adjustment produced by the presence of an enclitic blocks the aspiration rule. This is shown in (49), where the outcome of simplification does not serve as input for preaspiration rule 2, whereas the assimilation output of (48) does. Notice that in (49) the structure should in principle serve as an input for the preaspiration rule, as the same sound structure is found in (50) (or 47, for that matter), where preaspiration occurs.

- (49) mīt- 'grab'
 SYLL. [gá.mī.thfi]
 PREASP. _ / *[gá.mīht.hfi]
 SIMPL. [gá.mīt.hfi]
 gá = mīt-g = hí
 2.PST=grab[3.OBJ]=PL
 'You (pl) grabbed him.'

- (50) hét- 'dress'
 SYLL. [gá.héh.tʰf]
 PREASP. [gá.héh.tʰf] / *[gá.héh.tʰf]
 gá = hét = hí
 2.PST=dress[3.OBJ]=PL
 You (pl) dressed him.

5. Summary of the proposal and a brief excursion into the origins of preaspiration

In this paper, I have proposed that in Northern Otomi a phonetic [h] occurring in coda at word-internal level can be analysed in two ways depending on the structure at hand: it can be an output of a phonetic rule I have called 'preaspiration' or it can be the realization of a more abstract structure of phonological nature representing either a phonemic /h/ or the glottal component of a complex nucleus /vh(V)/.

I have also shown that a simplified treatment as either being phonetic or phonemic for all instances is not a convenient way to account for the complexities of the distribution of the segment in the sound system of Northern Otomi. Such an analysis has consequences for the writing system, because a phonological H should be represented in the orthography when in reality it never is. I have provided a number of important clues as to how to interpret the character of this internal [h], but I admit the distribution remains complex to receive a proper orthographic representation by the lay language user.

In the typology of language change, aspirations similar to the Otomi phenomenon commonly come from historical /s/, such as for example Spanish /kas.ko/ which is pronounced [kah.ko] or even [kax.ko] in certain varieties of Spanish (see Ferguson 1990, for a discussion). The question thus remains, where does Otomi preaspiration come from?

In order to provide an answer to such a question, I suggest that it emerged as the result of the breakdown of the old phonological system still present in Old Otomi, which is still observable in the data of the Colonial grammar finished by Pedro de Cárceres in 1580, only to be published in 1907. In the old system, there was a contrast of preaspirated stops /^hp, ^ht, ^hk/ and simple stops /p, t, k/, in such a way that words such as *^hpa 'day', *^htá 'buy' and *^hkít'a '5' contrasted with words such as *pi '3.pst', *tá 'eye' and *kóhó '4'. The simple stops in the old system were realized by default by voiceless phones [p, t, k], but occasionally, especially in intervocalic context they could also be realized as voiced [b, d, g]. The preaspirated stops were always realized as voiceless. Such a system could in principle be analysed as instantiating a fortis/lenis contrast involving stops for Oto-Pamean, as it has been previously suggested in Gibson (1956), Bartholomew (1960) and Blight and Pike (1976), which has also been proposed for other groups within Oto-Manguan (see DiCancio 2008, for references and a discussion) and for Oto-Manguan itself in Kaufman (1983).

Through time, in the historical dialect that later gave rise to Northern Otomi, which is precisely the one on which the Colonial grammar was based, all simple stops became voiced, so that *pi '3.pst' gave rise to bi, *tá 'eye' to dá or dá and *kóhó '4' to góhó. In contrast, all preaspirated stops lost their preaspiration feature, so that as *^hpa 'day' gave rise to pa, *^htá 'buy' to tá and *^hkít'a '5' to kít'a.

Crucially for our purposes, the preaspiration feature was maintained word-internally, in such a way that this preservation gave rise to the synchronic phonetic phenomenon I have treated here under the rubric of 'preaspiration', e.g. for an old word such as *^hmí.'tí 'lend' we have now its reflex [mí.ɦ.tí], which I have analysed here in phonological terms as /mí.tí/.

The old phonological system has partially survived in Tilapa Otomi, a conservative Southern Otomi language, and it is still at work in Santa Ana Hueytlalpan Otomi, a variety of Otomi which could be characterized as a phonologically conservative dialect of Eastern Otomi, but with reservations. Why the old system became unstable in some varieties of Otomi while it was retained in others remains a mystery, as most sound change still is, but it is undoubtedly connected with the general typological rarity of preaspirated stops cross-linguistically, as surveyed in Silverman (2003).

References

- Andrews, Henrietta. 1993. *The Function of Verb Prefixes in Southwestern Otomi*. [Summer Institute of Linguistics and University of Texas at Arlington Publications in Linguistics 115]. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Andrews, Henrietta, Doris Bartholomew, Richard Blight, Artemisa Echegoyen, Nancy Lanier, Donald Sinclair, Katherine Voigtlander, Steven England, Paul Mellema, and David Persons. 1983 [1978]. Steven England (ed.) in collaboration with Doris Bartholomew and Saúl Cruz Ramos. Otomi. *La inteligibilidad interdialectal en México: resultados de algunos sondeos*. Mexico City: Instituto Lingüístico de Verano. 46–48
- Bartholomew, Doris. 1960. Some Revisions of Proto-Otomi Consonants. *International Journal of American Linguistics* 26-4, 317-329.
- Blight, R. C. and Pike, Eunice. V. 1976. The phonology of Tenango Otomi. *International Journal of American Linguistics* 42-1, 51-57.
- Cárceres, Fray Pedro de. 1907 [1580]. *Arte de la Lengua Othomi*. Edited by Nicolás León. *Boletín del Instituto Bibliográfico Mexicano* 6, 38-155.
- CLIN: Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. 2008. *Diario de la Nación* (14-01-2008). Mexico DF. Accessible at <http://www.inali.gob.mx/pdf/CLIN_completo.pdf>.
- DiCancio, Christian Thomas. 2008. *The Phonetics and Phonology of San Martín Itunyoso Trique*. PhD Dissertation, University of California at Berkeley.
- Ferguson, Charles A. 1990. From essets to aitches: identifying pathways of diachronic change. William Croft, Keith M. Denning and Suzanne Kemmer (eds). *Studies in Topology and Diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday*. Amsterdam: John Benjamins. 61-78.
- Gibson, L. F. 1956. Pame (Otomi) Phonemics and Morphophonemics. *International Journal of American Linguistics* 22-4, 242-265.

- Kaufman, Terrence. 1983. New perspectives on comparative Otomanguean phonology. MS.
- Knapp, Michael. 2008. *Fonología segmental y léxica del mazahua*. Ciudad de México: Instituto Nacional de Antropología e Historia.
- Lastra, Yolanda. 2001. *Unidad y Diversidad de la Lengua: Relatos otomies*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- 2010. Diversidad lingüística: variación dialectal actual. Rebeca Barriga and Pedro Martín Butragueño (eds). *Historia Sociolingüística de México* Volumen II. Mexico City: El Colegio de México. 841-880.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- Palancar, Enrique L. 2009. *Gramática y Textos del Hñiñhö: Otomí de San Ildefonso Tultepec, Querétaro*. [Volumen I: Gramática]. Mexico City: Plaza y Valdés.
- Silverman, Daniel. 2003. On the Rarity of Pre-Aspirated Stops. *Journal of Linguistics* 39-3. 575-598.
- Soustelle, Jacques. 1937. *La Famille Otomi-Pame du Mexique Central*. Paris: Travaux et Mémoires de l'Institut d'Ethnologie.

4. Travaux de typologie

- TRAVAIL 1. Palancar, Enrique L. A paraître. Oto-Pamean. Soeren 381
Wichmann (ed.), *The Languages and Linguistics of Middle
and Central America: A Comprehensive Guide*. Berlin:
Mouton de Gruyter.
- TRAVAIL 2. Palancar, Enrique L. 2016. A typology of tone and 415
inflection: A view from the Oto-Manguean languages of
Mexico. Enrique L. Palancar et Jean Léo Léonard (eds.),
Tone and inflection: New data and new perspectives, pp.
109 – 140. Berlin: De Gruyter Mouton.
- TRAVAIL 3. Palancar, Enrique L. et Roberto Zavala Maldonado. 2013. 447
Clases léxicas, posesión y estructuras de cláusula compleja
en Mesoamérica: Una introducción. Enrique L. Palancar et
Zavala Maldonado (eds.). *Clases léxicas, posesión y
cláusulas complejas en lenguas de Mesoamérica*, pp. 17 –
48. Mexico: CIESAS.
- TRAVAIL 4. Palancar, Enrique L. 2012. A typology of split conjunction. 479
Linguistic Typology 16: 265 – 320.
- TRAVAIL 5. Palancar, Enrique L. 2008. Varieties of ergatives. Andrey 535
Malchukov et Andrew Spencer (eds.), *The Handbook of
Case*, Chapter 37. Oxford: Oxford University Press.

The Oto-Pamean languages

Enrique L. Palancar

Structure et Dynamique des Langues (UMR8202), CNRS

To appear in Wichmann, Søren (ed.),

The Languages and Linguistics of Middle and Central America: A Comprehensive Guide.

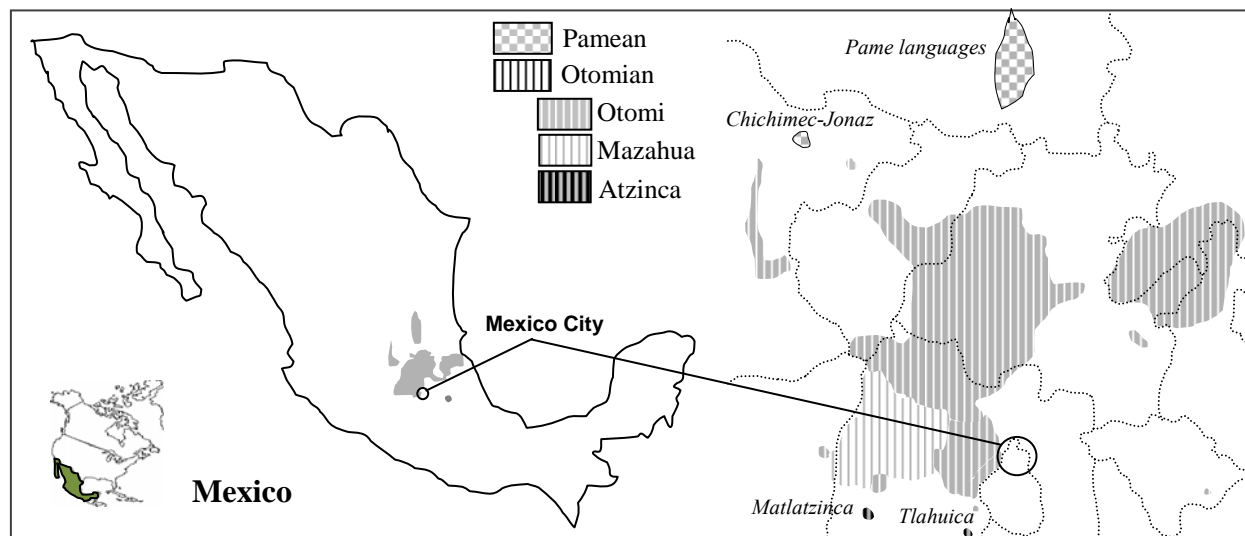
Berlin: de Gruyter Mouton.

The Oto-Pamean languages of Mexico form the northernmost branch of the Oto-Manguean stock (see Kaufmann, this volume). They are currently spoken in isolated villages scattered throughout a very large area of Central Mexico from San Luís Potosí to the north to the far end of the State of Mexico to the south, and to the east in the mountainous ridges of Hidalgo and Puebla. Oto-Pamean languages are neighbours to Mayan Huastec, Totonac-Tepehua languages, Purepecha and Nahuatl.

Bartholomew (1965) proposes that Oto-Pamean is divided into two sub-branches: Pamean and Otomian. Since then, this taxonomy has been accepted by consensus. These sub-branches in turn sub-divide further into two groups each: Pamean into Chichimec and Pame, and Otomian into Atzinca and Otomi-Mazahua. A snapshot of Oto-Pamean is given in Table 1 and in Map 1.

Phylum	Branch	Sub-branch	Group	Language	No. of speakers
Oto-Manguean	Oto-Pamean	Pamean	Chichimec	Chichimeco-Jonaz	1,350
			Pame	Northern Pame	5,600
				Central Pame	7,400
		Southern Pame		†	
		Otomian	Atzinca	Matlatzinca	650
				Tlahuica (Ocuiltec)	400, 100 (2011)
			Otomi-Mazahua	Mazahua	100,000
Otomi languages	241,000				

Table 1. Oto-Pamean at a glance (Estimates of number of speakers are from 2000 except when indicated)



Map 1. Location of Oto-Pamean (following Soustelle 1937).

All these groups except Chichimec consist of various languages. The total number of speakers in the census by the National Institute of Indigenous Languages of Mexico (INALI) from 2000 surpasses 360,000. Three languages are particularly endangered: Tilapa Otomi, Ixtenco Otomi and Tlahuica. For these languages, 15 years makes a fundamental difference regarding number of speakers. For example, the INALI estimated 250 speakers in 2000 for Tilapa Otomi, but at the time of writing there are fewer than a dozen fluent speakers alive.

Within Oto-Pamean, Otomi-Mazahua is the group with the largest number of speakers and the greatest degree of internal diversity. Knapp (2002) proposes that Mazahua consists of at least 12 main dialects as indicated in Table 2. As for Otomi, although at least 22 different linguistic sub-systems have been identified as Otomi, there is no established consensus as to which of them should be treated as independent languages and which as dialects. To circumvent the problem, Otomi has often been talked about as forming a dialectal continuum and the literature has avoided terms such as ‘language’ and ‘dialect’ in favor of a cover term such as ‘variety’. Against this inertia, the current official stance by the INALI in the CLIN (Sp. *Catálogo de Lenguas Indígenas Nacionales*) (2008) is to treat Otomi innovatively as forming a linguistic family with nine different languages. Unfortunately, the methodology behind the classification of the CLIN is far from rigorous and it overrates the weight of sociolinguistic factors, so that closely related dialects are treated as different languages. I use Palancar’s (2013a) taxonomy here, which advances on the taxonomy in Lastra (2001) by proposing the existence of the six main languages given in Table 1.

Language	Speakers	Variety
Mazahua	c. 100,000	Atacomulco Mazahua Temascalcingo Mazahua El Oro Mazahua Jocotitlán Mazahua San Felipe del Progreso Mazahua Villa Victoria Mazahua Zitácuaro Mazahua Ixtlahuaca Mazahua Santa Cruz Tepexpan Mazahua Almoloya Mazahua Donato Guerra Mazahua
Northern Otomi	c. 113,000	Santiago Mexquititlán Otomi San Ildefonso Tultepec Otomi Tolimán Otomi Guanajuato Otomi Acambay Otomi Tierra Blanca Otomi San Ildefonso Chantepec Otomi Mezquital Otomi Ixhuatlán de Madero Otomi
Western Otomi	c. 84,000	Temoaya Otomi San Felipe del Progreso Otomi San Andrés Cuexcontitlán Otomi
Eastern Otomi (Eastern Highlands Otomi)	c. 44,000	San Gregorio Otomi San Antonio Otomi San Felipe Otomi Tenango Otomi

	Texcatepec Otomi Santa Ana Hueytlalpan Otomi
Tilapa Otomi	8 (fluent)
Acazulco Otomi ¹	c. 400
Ixtenco Otomi	c. 450 (2000)

Table 2. Otomi-Mazahua.

Oto-Pamean languages have a clear verb vs. noun word class distinction, but adjectives are rare, if they exist at all (most property concepts are expressed by verbs in small relative clauses or in V+N compounds). All languages are head-marking, and although nouns do not have gender or inflect for case, they fall into different noun classes with interesting consequences for the morphological marking of number (§4.1) and the marking of possession (§4.2). Like nouns, verbs also have intricate morphology. They fall into different conjugation classes (§5.2) and they inflect for the person of subject and for a substantial number of tense/aspect/mood values (§5.1). They may also bear information about spatial deixis (§5.4) and negation (Tlahuica and Chichimeco-Jonaz) and may carry applicative-like morphology for fronted adverbials (Otomi-Mazahua). All languages have semantic alignment (§5.5) at least involving the lexicalization of property concepts by means of verbs. The alignment in ditransitives is probably secundative historically, but the system has been considerably rearranged and it is now indirective in most languages. As for diathesis, transitive verbs have an impersonal-passive voice. Derivational morphology involving causativization, antipassivization and middle voice is now unproductive (§3). There is considerable use of clause-level serialization, which is particularly exploited in complementation contexts. An example of this is shown in (1) in Chichimeco-Jonaz, where a notion such as ‘I want to work’ is conveyed literally as ‘I want I would work’. The meaning of ‘work’ is itself conveyed by the serialization of two synonymous, inflected verbs. Another example of the same structure is given in (2) in Tilapa Otomi. When an example comes from a text, this is indicated as ‘Txt’.

(1) *sá* [é-'i] [ná-tán+nú-tan]
 CHI and 1.PRS-want 1.POT-work+1.POT-work
 ‘And I want to work.’ (Txt) (Angulo 1933: 191)

(2) *porke* [hín=a nde] [ta=tandy]=a
 TIL porque NEG=3.INCPL.R want 3.CPL.IRR=SS/buy.AS=PRTCL
 ‘Because they don’t want to buy it.’ (Txt)

The use of clausal juxtaposition like in (1-2) is typical of Oto-Manguan languages, and goes hand in hand with the lack of non-finite forms in these languages. While there are verb nominalizations in Oto-Pamean languages, the process is commonly unproductive and highly lexical, and no language has a proper inflectional infinitive, except perhaps for Eastern Otomi, which has intransitive infinitives of activity verbs with a limited set of matrix verbs, as for example in (3) (see Palancar, forthcoming).

¹ Acazulco Otomi is spoken in San Jerónimo Acazulco, a village in the municipality of Ocoyoacac, State of Mexico. This village is a neighbor of San Pedro Atlapulco, where a different variety of Otomi is also spoken. But we know very little about this variety, apparently only spoken by a few elderly people. Tilapa Otomi is in turn spoken in Santiago Tilapa, a village that is known to have been founded in 1559 by Otomi people from San Pedro Atlapulco. As a result Atlapulco Otomi and Tilapa Otomi are descended from the same ancestor language.

shown in Table 13, adapted from Knapp (2008), with 15 vowels (oral and nasal) and 53 consonants, including series of simple, aspirated and ejective consonants.²

Vowels			i	ĩ	u			ĩ	ĩ	ũ
			e	ə	o			ẽ		õ
			ɛ	a	ɔ					ã
Stops	Voiceless	Simple	p	t		ts	tʃ	k	k ^w	
		Aspirated	p ^h	t ^h		ts ^h	tʃ ^h	k ^h	k ^{hw}	
		Ejective	-	t'		ts'	tʃ'	k'	k' ^w	ʔ
Fricatives	Voiced	Prenasalized	^m b	ⁿ d		ⁿ dz	ⁿ dʒ	^ŋ g	^ŋ g ^w	
		Ejective	ɸ	ɸ'						
		Simple			s		ʃ			h
Approximants	Voiced	Aspirated			s ^h					
		Ejective			s'					
		Simple	β		z		ʒ		ɣ	
Nasals		Aspirated	w				j			
		Ejective	w'				j'			
		Simple	m	n			ɲ			
Rhotics		Aspirated	m̥	n̥			ɲ̥			
		Ejective	m'	n'			ɲ'			
		Simple		r						
Laterals	-			(r)						
				(l)						

Table 3. The phonemes of Mazahua, adapted from Knapp (2008).

Knapp (2008) treats approximants /w̥, j/ and nasals /m̥, n̥, ɲ̥/ as instantiating an aspirated series to create featural coherence with stops and fricatives. Alternatively, these segments could be treated as voiceless, but in reality they are the only surviving units in Otomi-Mazahua of an old preaspirated series which is still observed in some phonologically conservative Otomi languages, such as Tilapa Otomi, as indicated in Table 4, where they are still found in operation in the sub-system of voiceless stops and affricates, but where the oppositions are considerably weakened elsewhere.

Voiceless	Simple	p	t	k	k ^w	ts	s	ʃ	ʔ	h				
	Preaspirated	^h p	^h t	^h k	^h k ^w	^h ts	-	-	-	-	w̥	j	m̥	n̥
	Aspirated	p ^h	t ^h	k ^h	k ^{hw}	ts ^h	-	ʃ ^h	-	-	-	-	-	-
	Ejective	p'[6]	t'	k'	k' ^w	ts'	-	-	-	-	-	-	-	-
Voiced	Simple	b	d	g	g ^w		z	-	-	-	w	j	m	n
	Prenasalized	^m b	ⁿ d	^ŋ g	^ŋ g ^w	-	-	-	-	-	-	-	-	-
	Ejective	^m b'	ⁿ d'	^ŋ g'	^ŋ g' ^w	-	-	-	-	-	'w	'j	'm	'n

Table 4. The phoneme inventory of consonants of Tilapa Otomi.

In Mazahua and in most Otomi languages, the old series of simple voiceless stops from the common ancestor, which are still attested in Tilapa Otomi, became voiced (i.e. *k>g) while

² In the Oto-Pamean languages, both aspiration and ejection are laryngeal features which are clearly associated with the articulation of consonants. In other words, a model based on the proposal by Golston & Kehrein (2004) where such features are associated with the vocalic nuclei would not be optimal to account for the segmental phonology of these languages.

the preaspirated series lost the preaspiration feature becoming simple voiceless segments (i.e. ^{*h}k>k). Preaspiration survived word internally in unstressed syllables, but it is now a predictable phonetic feature (for details, see Palancar 2013b) (e.g. Old Otomi ^{*/h}kí^hki/ ‘remove’ > Northern Otomi /kiki/ [kí^hki]).

Like other Oto-Manguean languages, Oto-Pamean languages are tonal, but their tonal systems are not elaborate. Pamean languages have a minimal low/high opposition which reflects the old system. For example, in Chichimeco-Jonaz most words are disyllabic and suprasegmentally they are organized attending to the four melodies in (6) (Herrera-Zendejas 2009). Minimal pairs are not common, but they are found; a few are shown in (7).

(6)	/H-H/	<i>émbó</i>	‘woodland, hill’
CHI	/H-L/	<i>súnden</i>	‘scorpion’
	/L-H/	<i>kanzé</i>	‘cicada’
	/L-L/	<i>kithæ</i>	‘oven’

(7)	/H-L/	<i>únhi</i>	‘s/he dropped it’	vs.	/L-H/	<i>unhí</i>	‘s/he lost it’
CHI	/H-H/	<i>sígã</i>	‘magical creature’	vs.	/L-H/	<i>sigã</i>	‘his/her ear’
	/H-L/	<i>úrʔi</i>	‘forest’	vs.	/L-H/	<i>urʔí</i>	‘cloth’

Otomian developed an ascending tone in lexical roots (i.e. it is neither present in affixes nor stem formatives). Again minimal pairs based on tone are rare. The triplet in (8) is exceptional. The element *+ni* in (8) is a toneless stem formative that receives [H] at a prosodic boundary (#), e.g. [L-H]# [khi+ni]# ‘corn-dough’.

(8)	/H/	<i>khí+ni</i>	‘facial hair’
E.OTO	/HL/	<i>khĩ+ni</i>	‘flat stone for grinding’
	/L/	<i>khi+ni</i>	‘lime-soaked maize’

In Oto-Pamean, tone basically works at a lexical level, but it plays an occasional role in inflection. This is briefly discussed below in §5.3.

2. Word internal structure

Lexical roots in Oto-Pamean are /CV/. In Otomian, there are also words with a laryngeal nucleus /CV^hV/, e.g. Eastern Otomi *tuhu* ‘have soup’, *tũhũ* ‘sow’, etc. Bartholomew (1965) has proposed that in Proto-Oto-Pamean, verb roots were compounded with spatial adverbs to express complex meanings, spatial or more abstract, comparable to English phrasal verbs like *take up*, *take in*, *take on*, etc. At a later stage, these spatial satellites (in the sense of Talmy 2000) became bound morphemes and eventually lexicalized. They are now treated as opaque stem formatives. In Pamean, they were reduced to /C/ and then got lost in most cases. They were preserved in Otomian, most notably in Otomi-Mazahua, where they were levelled to the structure /Ci/. The /i/ in such a structure was later reanalyzed as a pausal marker in innovative Otomi languages. The stem formatives in Eastern Otomi are exemplified in (a) in Table 5. The formations in (b) include a nasal extension at root level (data adapted from Echegoyen & Voigtlander, 2006).

(a)			(b)					
+d-	<i>pĩ+d-i</i>	‘scare’	+n-	<i>’da+n-i</i>	‘cross’	-N+n-	<i>tsí-n+n-i</i>	‘leak’
+ph-	<i>pe+ph-i</i>	‘work’		<i>pāh+n-i</i>	‘sniff’	-N+m-	<i>ku-m+m-i</i>	‘grind’
+g-	<i>’a+g-i</i>	‘bury’	+t-	<i>tsoh+t-i</i>	‘spit to’	-N+t’-	<i>wā-n+t’-i</i>	‘stir’
+i	<i>ně+i</i>	‘dance’	+t’-	<i>k’é+t’-i</i>	‘burst’	-N-ts’-	<i>pé-n+ts’-i</i>	‘bend’
	<i>kóh+i</i>	‘stay’	+ts’-	<i>khwá+ts’-i</i>	‘lean’	-N+g-	<i>pú-n+g-i</i>	‘bleach’
+k-	<i>’áh+k-i</i>	‘tilt’	+x-	<i>hu+x-i</i>	‘whistle’	-N+d-	<i>xǎ-n+d-i</i>	‘multiply’
+m-	<i>ná’+m-i</i>	‘scatter’						
	<i>kwäh+m-i</i>	‘stick’						

Table 5. The stem formatives of Eastern Otomi.

Stem formatives are unproductive morphemes, most of them with opaque meanings, but a range of spatial senses can be reconstructed when a great number of lexeme families are studied in detail (see Hernández-Green 2009 and Voigtlander *et al.* 2006). Two such lexeme families appear in (9) from Mezquital Otomi, a variety of Northern Otomi. Some of such families have a monosyllabic root (a); some do not (b), (the *-i* is an exponent of a free form of a verb; a type of pausal form).

- (9) a. *khu* ‘grasp’ b. **hwe* ‘produce light’
 N.OTO *khu+t’-i* ‘drag’ *hwe+t’-i* ‘flash’
khu+ts’-i ‘pull up’ *hwe+ts’-i* ‘shine’
khu+k-i ‘pull’
khu+’m-i ‘pull down’

In Otomi-Mazahua, the lexicalization of stem formatives gave rise to a substantial portion of the verbal lexicon. For example, based on the 1864 verbal entries in Hernández *et al.* (2004) from Mezquital Otomi, Hernández-Green (2009) proposes that at least 1175 are verbs with stem formatives, and only 85 of them are basic monosyllabic roots (the remaining verbs include compounds and other root types).

Suffixation and enclitization occur after stem formatives and produce morphotactic adjustments, which at times display great complexity. But a number of formatives appear to have come in a second wave of compounding resulting in bipartite stems which can still be split. For example, in Atzincan, the formatives *+ts’i*, *+tʃ’i*, *+ti* and *+li* split from the root with the affixation of the applicative $\{-p\}$. This is shown in (10), adapted from Martínez-Ortega (2010, 2012). In Otomi, some verbs may show more than one formative, as in (11).

- (10) TLA *kwäh+ts’i* ‘stick’ > *kwäh-p+ts’i* ‘stick to someone’
wi+li ‘sell’ > *wim-b+li* ‘sell to someone’

- (11) E.OTO *kǫ̃+x+k-e* ‘scrub’ < *^h*kǫ̃+ts’[i]+k-i*

3. Derivation

Derivation, word formation and valence changing mechanisms are unproductive in Oto-Pamean. Nowadays, it tends to be Spanish loans which enrich the lexicon. However, one can identify in many verbs the outcome of old operations. For example, in Otomi-Mazahua many intransitive verbs bear a nasal prefix and have typical semantics associated with the middle voice (reflexive and reciprocal). Most of them come in lexical pairs, as in (12a); a few are deponent (12b). The same nasal prefix is also found in verbs with typical antipassive

semantics involving an inanimate O like in (12c), others with human O's have the suffix *-te* (12d). Examples are from the San Ildefonso Tultepec variety of Northern Otomi.

- (12) a. *m-be+n-i* 'regain consciousness' ~ *be+n-i* 'think, remember'
 N.OTO *m-ʹ[b]e+d-i* 'get lost' ~ *ʹbe+d-i* 'lose'
n-zengwa 'visit each other' ~ *zengwa* 'visit'
- b. *n-k_o* 'dress below waist'
ñ-he 'dress above waist'
n-ti 'get drunk'
n-the 'meet up with someone'
- c. *m-pe+n-i* (intr) 'do the laundry' ~ *pe+n-i* (tr) 'wash clothes'
- d. *tah-te* (intr) 'bite people' (a dog) ~ *tša* (tr) 'bite'

Similarly, a substantial number of intransitive verbs originated from old V+N compounds, like in (13a). Most disyllabic stems come from frozen compounds whose compounding elements have opaque semantics, like in (13b). Examples are from Tilapa Otomi.

- (13) a. ^h*tyü-the* 'be thirsty' < ^h*tyü* 'die' + *t[e]he* 'water'
 TIL ^h*tsi-hme* 'have a meal' < ^h*tsi* 'ingest' + *hme* 'tortilla'
- b. ^h*tša_ya* 'rest' < ^h*tša* '?' + **ya* '?'
ze_ngwa 'visit' < **ze* '?' + **ngwa* '?'

Valence changing mechanisms are only observable in lexical pairs. Many causative verbs were historically formed by a laryngeal nucleus, as in (14a), but the process is unproductive as the same type of nucleus is also observed in many intransitive verbs (14b). There are a myriad of such mechanisms, reflecting a progressive accumulation of old morphological operations (see Palancar 2004, Palacios & Palancar 2010 and Palacios 2011). Examples are from the San Ildefonso Tultepec variety of Northern Otomi.

- (14) a. *ʹo-ʹ+tʹ-i* (tr) 'dry' ~ *ʹo+tʹ-i* (intr) 'dry'
 N.OTO b. *khü-ʹ+tʹ-i* (intr) 'drip'

4. Nominal inflection

As is typical of Oto-Manguéan, Oto-Pamean languages are canonically head-marking, so nouns do not inflect for case. But there is nominal inflection involving number and possession, which is sometimes morphologically intricate in nature because these two dimensions interact with a noun class system, which is vibrant in Pamean, still observable in Atzinca and present, albeit moribund, in Otomi-Mazahua.

4.1. Noun classes

Central Pame is the Oto-Pamean language with the most developed noun class system. It is commonly believed that this is a conservative feature rather than an innovation. In this language, nouns fall into at least 11 different classes attending to the prefixal elements they bear.³ This is shown in Table 6, adapted from Gibson and Bartholomew (1979). The basic stem is used for both singular and dual.

³ To this list one may add a number of other nouns with idiosyncratic behaviour (see Gibson and Bartholomew 1979 for details).

Class	SG/DU	PL	SG/DU	PL	
1	<i>ŋgo-</i>	<i>N-</i>	<i>ŋgo-p^hói</i>	<i>m-p^hói</i>	‘excrement’
2	<i>ko-</i>	<i>N-</i>	<i>ko-póʔ</i>	<i>m-bóʔ</i>	‘land’
3	<i>ni-</i>	<i>li-</i>	<i>ni-k^lwǎŋ</i>	<i>li-k^lwǎŋ</i>	‘capulincillo tree’
4	<i>ní-</i>	<i>rí-</i>	<i>ní-ŋg^jéhe</i>	<i>rí-ŋg^jéhe</i>	‘year’
5	<i>na-</i>	<i>la-</i>	<i>na-tsê</i>	<i>la-tsê</i>	‘plum’
6	<i>na-</i>	<i>ra-</i>	<i>na-ts^héŋ</i>	<i>ra-ts^héŋ</i>	‘key’
7	<i>ma-</i>	<i>wa-</i>	<i>ma-tsì</i>	<i>wa-tsì</i>	‘pitcher’
8	<i>mi-</i>	<i>wi-</i>	<i>mi-tfǎʔ</i>	<i>wi-tfǎʔ</i>	‘corn bin’
9	<i>tsó-</i>	<i>só-</i>	<i>tsó-ndo</i>	<i>só-ndo</i>	‘eggshell’
10	<i>tsa-</i>	<i>sa-</i>	<i>tsa-mp^hǎŋ</i>	<i>sa-mp^hǎŋ</i>	‘pocket’
11	<i>tʃi-</i>	<i>ʃi-</i>	<i>tʃi-hàŋŋ</i>	<i>ʃi-hàŋŋ</i>	‘spoon’

Table 6. Noun classes in Central Pame.

Further research is needed to explore the origins of this system and its development, but nowadays the classes are largely semantically opaque, except for class 11 which includes most nouns denoting tools (but not exclusively). Class 1 serves as a default.⁴ Apart from the plural stems, an animate noun receives additional number marking by means of affixes (at times with additional stem adjustments): *-i* for dual and *-t* for plural. This is illustrated in Table 7, which shows a number of nouns treated as ‘invariable’ (inv.), lying outside the system in Table 6. Such nouns include loanwords (b), agent nouns (c), and a set of nouns referring to animals (d).

	SG	DU	PL	
a. cl.9	<i>tsó-m^hé</i>	<i>tsó-m^hé-i</i>	<i>só-m^hé-t</i>	‘butterfly’
cl.11	<i>tʃi-kíʔ</i>	<i>tʃi-kíʔ^l</i>	<i>ʃi-kíʔ<t>ʔ</i>	‘goat’
b. inv.	<i>pastól</i>	<i>pastó<i>ʔ^l</i>	<i>pastól-t</i>	‘shepherd’ < Sp. <i>pastor</i>
c. inv.	<i>ka-pě</i>	<i>ka-pě-i</i>	<i>ka-pě-t</i>	‘thief’ < <i>pě</i> ‘steal’
d. inv.	<i>ko-pětsʔ</i>	<i>ko-pě<i>tʃʔ</i>	<i>ko-pěs<t>ʔ</i>	‘badger’
inv.	<i>ko-ts^híʔ</i>	<i>ko-ts^híʔ</i>	<i>ko-ts^héʔk^l</i>	‘snake’

Table 7. Number marking of animate nouns in Central Pame.

In Northern Pame, the original system has been considerably reduced, as shown in Table 8. The default class is the only one having a number contrast, where plural is realized by morphophonological adjustments in the stem. Other nouns are invariable. Animate nouns, regardless of class, express dual and plural by means of suffixes, but all nouns of the default class also have a dual. The data are adapted from Berthiaume (2003).

⁴ It is the class with the largest number of members and with the least semantic specification.

Class	SG	DU	PL		
Default	n-m ^h ε	n-m ^h ε-ji	j-m ^h ε	→ m ^h jε	‘tortilla’
	n-t’ùs	n-t’ùs-i	j-t’ùs	→ ts’jùs	‘house’
	n-t ^h úʔ	n-t ^h úʔ-ji	j-t ^h úʔ-ðit	→ tʃ ^h úʔðit	‘armadillo’
inv. <i>gu-</i> (animate)	gu-saʎ’	gu-saʎ’-i	gu-saʎ’-it		‘eagle’
inv. <i>sɬ-</i> (instr.)	sɬ-hăwʔ	–	–		‘spoon’
inv. <i>dɨ-</i> (misc.)	dɨ-nâs	–	–		‘lemon’
inv. <i>kɨ-</i> (misc.)	ki-nte	–	–		‘water’
inv. miscellanea	pi-gas	pi-gas-i	pi-gas-it		‘cow’ < Sp. <i>vacas</i>

Table 8. Noun classes in Northern Pame.

Within Otomian, the most elaborate noun class system is found in Atzacan, where we find the three classes shown in Table 9 in Matlatzinca, adapted from Bartholomew (nd).

Class	SG	DU	PL	
Default	in-fuʔyowi	te-fuʔyowi	ne-fuʔyowi	‘coyote’
	in-tʃuwampa	te-tʃuwampa	ne-tʃuwampa	‘tomato’
	in-buro	te-buro	ne-buro	‘donkey’ < Sp. <i>burro</i>
<i>ni-</i> (misc.)	ni-nseʔe	te-nseʔe	ne-nseʔe	‘big bird’
	ni-nto	–	–	‘hail’
	ni-k ^h we:tʃi	te-k ^h we:tʃi	ne-k ^h we:tʃi	‘broom’ < <i>kwetʃi</i> ‘sweep’
<i>we-</i> (human)	we-benə:bi	te-benə:bi	ne-benə:bi	‘dancer’
	we-to-waʔa	te-to-waʔa	ne-to-waʔa	‘child’

Table 9. Noun classes in Matlatzinca.

Tlahuica is like Matlatzinca, but with additional minor classes; particularly relevant are a *pi*-class which mainly includes body parts (e.g. *pi-ntyɨ* ‘forehead’), and a class with a stem prefix *nu-* (e.g. *nu-tho* ‘roof’, *nu-ye* ‘hand’), a few of whose members are interchangeable with the *ni*-class (e.g. *nu-ha* ~ *ni-ha* ‘earth’).

In Otomi-Mazahua, the nominal class system is in decline. It is still observable in the conservative Otomi languages of Tilapa and Ixtenco, whose classes are attested in the missionary grammar by Pedro de Cárceres, finished in 1580. In all languages of the group, number has been reduced to a singular-plural opposition.

Class	SG	PL	
Default	ar ’yü	i ’yü	‘path’
<i>ru</i> (mainly borrowings)	ru me ^y diko	i me ^y diko	‘doctor’ < Sp. <i>médico</i>
<i>ra</i> (nominalizations and predicative NPs)	ra ^h ně	i ^h ně	‘dance’ < <i>ně</i> ‘dance’
<i>di</i> (nominalizations of Class III verbs)	di ^h pəphi	i ^h pəphi	‘work’ < ^h <i>pəphi</i> ‘work’

Table 10. Noun classes in Tilapa Otomi.

Finally, in both Mazahua and Northern Otomi the old noun class system has disappeared. These two languages are the most innovative within Oto-Pamean and they share many features, probably due to intense contact during Colonial times. In some descriptions (Voigtlander & Echegoyen, 1985; Hess, 1968, etc.), the singular marker *ra/ar* (or in its reduced form =*r*) has been misinterpreted as being a definite article, when it has nothing to do with definiteness. This misconception is partly due to the fact that noun class markers in Otomi-Mazahua are not prefixes, but clitics words, so they look more like

determiners. This is illustrated in the contrast shown in (15) from Tilapa Otomi, where a hesitation pause (indicated by #) can naturally occur between the marker and its head noun.

- (15) a. este#...[**ru**=me^ydiko] re='ëna
 TIL HES SG=doctor 3.INCPL.R=say
 'Eh...the doctor said.' (Txt)
- b. o giti=gwöntsu='mbe[=**ru**#... berdolaga]
 or 1.CPL.IRR=stir.in.AS=PL.EXCL=SG *portulaca.oleracea*
 'Or we stir in the... *verdolaga*.' (Txt)

The behavior of nominal markers in Otomi-Mazahua has important consequences for the way we should interpret nouns and NP structure in the proto language. Following Givón's (1971) aphorism that today's morphology is yesterday's syntax, the wordhood of noun markers as clitics in Otomi-Mazahua reveals an old stage of syntactic compounding. The markers involved were later reanalyzed as prefixes in all other groups, giving the nominal inflectional morphology of these languages a more canonical profile. From this picture, we can imagine that the system of Proto-Oto-Pamean was morphologically isolating.

4.2. Possession

Nouns in Oto-Pamean also fall into different classes attending to how they inflect for the dimension of possession. Again, the Pamean group has the most intricate system while Otomian has the simplest.

In both Pame and Atzacan, possession classes correlate with nominal classes, but at times only tangentially. This means that if a noun belongs to a given noun class, this can often be taken as an indication as to what possession class it will belong to. But in the end, it is a matter of high probability rather than certainty.

4.2.1. Possession in Pamean. In Pamean, there are two major classes of nouns: (i) 'variable nouns', which have internal possession inflection, and (ii) 'invariable nouns', which have invariable stems. The morphology of possession is entrenched in Pamean to such an extent that variable nouns in Chichimeco-Jonaz have been claimed by Angulo (1933) to lack an absolute form, that is, a form to be used when the noun is not possessed. This is illustrated in (16), where the possessed form by a third person plural possessor is also the form used to refer to a generic noun.

- (16) *kúri* 'their water' → 'water' (see example 5b)
 CHI *úvó* 'their land' → 'land, earth, world'

The full paradigm of a possessed noun in Chichimeco-Jonaz is given in Table 11 with the noun *kunú* 'my field', adapted from Angulo (1933) (where low tone, being a default, is not represented). Note that the different forms in the paradigm attend to the number value (SG, DU and PL) of both the possessor and the possessum. The forms in shading represent principal parts, i.e. the forms one would need to know in order to be able to produce the other forms in the paradigm (Finkel and Stump, 2007).

		Number of possessor					
		SG		DU		PL	
Number of possessor	SG	1 st	kunú	EXCL	kunú-m ^p	EXCL	kun-hú
				INCL	kunú-s	INCL	kunú-n
		2 nd	kínu	kínu-s		kínu-n	
	3 rd	kinú	kinú-s		kinhú		
	DU	1 st	kunú-s	EXCL	kunú-vó-s	EXCL	kun-hú-n
				INCL	kunú-s-és	INCL	kunú-s-ín
		2 nd	kínu-s	kínu-s-es		kínu-s-in	
	3 rd	kinú-s	kinú-s-és		kinhú-s		
	PL	1 st	kunú-r	EXCL	kunú-r-úm ^p	EXCL	kunú-r-hú
INCL				kunú-r-és	INCL	kunú-r-ín	
2 nd		kínu-r	kínu-r-es		kínu-r-in		
3 rd	kinú-r	kinú-r-és		kinhú-r			

Table 11. The possession paradigm of a noun in Chichimeco-Jonaz.

In the possession inflection of the noun *kunú* ‘my field’ in Table 11, there are apophonic changes in the stem, i.e. the forms for 1st person have <u> while the other persons have <i>. There are also tonal modifications, and the form for the 3rd person plural possessor has a stem change from *kinú* to *kinhú*. The paradigm in Table 11 represents a specific possession class, but there are three others. All four of them are given in Table 12 with examples.

Cl.	Segments				Examples				Tone changes				Stem changes				
	1SG	2SG	3SG	3PL	1SG	2SG	3SG	3PL	1SG	2SG	3SG	3PL	1SG	2SG	3SG	3PL	
I	-	-	-	-	kanú	kánu	kànu	-	‘nose’	LH	HL	LH	-	-	-	-	
					námen	namén	námen	-	‘knee’	HL	LH	HL	-				
II	<u>	<i>	<i>	<i>	rusé	ríse	risé	-	‘skin’	LH	HL	LH	-				
					kunú	kínu	kinú	kinhú	‘field’	LH	HL	LH	LH	n	n	n	nh
					númbe	nivé	nímbe	úve?	‘hunger’	HL	LH	HL	HL	mb	v	mb	v
					nukhú	níkhú	nikhú	bur’ú	‘child’	LH	HL	LH	LH	kh	kh	kh	r’
					suká	síká	sugá	-	‘ear’	LH	HL	LH	-	k	k	g	(kh)
				tsútse	tjitse	tjitse	-	‘wing’	HL	LH	HL	-	ts	tj	tj	-	
III	na-	u-	u-	u-	nátsa	utsá	úza	úts’a	‘food’	HL	LH	HL	HL	ts	ts	z	ts’
					náku	ukú	úgu	úkhu	‘road’	HL	LH	HL	HL	k	k	g	kh
					nátán	után	úrán	úrhán	‘work’	HH	LH	HH	HH	t	t	r	rh
					námb ⁱ	ungw ⁱ	úm ⁱ	úp ⁱ	‘music’	HL	LH	HL	HL	mb	ngw	m	p
IV	ta-	ki-	ta-	ta-	tásots	kisóts	tásots	tátshots	‘belt’	HL	LH	HL	HL	s	s	s	tsh
				na-	táfin	kifín	táfin	názin	‘axe’	HL	LH	HL	HL	f	f	f	z
				ra-	taté	kíte	taté	raráhé	‘knife’	LH	HL	LH	LH	t	t	t	rh

Table 12. The possession classes of Chichimeco-Jonaz.

Angulo (1933) illustrates the classes in Table 12 with 61 nouns: 12 are of class I; 29 of class II; 14 of class III; and 6 of class IV.⁵ Membership to the classes is arbitrary. In other words, the phonological shape of the stem does not help to predict class membership, and neither

⁵ Angulo’s sample includes c. 30 odd nouns that do not fully abide by the patterns in Table 12. Such nouns include body parts such as ‘hand’, ‘penis’, ‘mouth’, and ‘face’, common kinship terms such as ‘wife’, ‘husband’, ‘mother’, and ‘father’, and other nouns referring to commonly possessed entities such as ‘house’, ‘name’ and ‘money’, but also ‘fire’.

does the semantics of the noun. For example, most nouns of class IV refer to tools—cognate with the possession class in Pame serving instrumental noun classes, see Table 15 below—but nouns for tools are also found in other classes, such as class III, and many are of the invariable type are borrowings. Also in class IV we find the noun for ‘plant’ which is not a tool. Similarly, although half of the 61 nouns in Angulo’s sample are body part nouns—a natural expectation for such a sample—the nouns are found across the classes without an apparent motivation. The same argument can be used with clothing, ‘trousers’ is in class I, ‘clothes’ is in class II, but ‘shoe’ and ‘hat’ are in class III, and many words for clothing are invariable.

On the formal level, the nouns of class I mark possession by a 2nd person by inverting the lexical melody. Class II is characterized by Angulo (1933) as having the apophonic series <u> vs. <i>. There are internal changes at stem level too. Such changes are morphophonologically predictable for some of the cases, for others they are not.

Possession class II has a cognate in Pame. The relevant data is given in Table 13, adapted from Berthiaume (2003) and Gibson and Bartholomew (1979).

		1SG		2SG		3SG		ABSOLUTE	
		SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS
SG	N.PAM	nʔk'wíʃ	n-ʔu-k'íʃ	ntʃ'íʃ	n-j-k'íʃ	nʔtʃ'íʃ	n-ʔj-k'íʃ	nk'íʃ	n-k'íʃ
	C.PAM	nok'wéʃ	no-k'wéʃ	nik'éf	ni-k'éf	nik'éf	ni-k'éf	ngok'éf	ngo-k'éf
PL	N.PAM	riʔk'wíʃ	r-ʔu-k'íʃ	ritʃ'íʃ	r-j-k'íʃ	riʔtʃ'íʃ	r-ʔj-k'íʃ	ts'íʃ	j-k'íʃ
	C.PAM	rok'wéʃ	ro-k'wéʃ	rik'éf	ri-k'éf	rik'éf	ri-k'éf	ng'éf	n-k'éf

Table 13. Possession forms of the word for ‘paper’ in Pame.

Surface forms in Pame pose an analytical problem. For Northern Pame, Berthiaume (2003) posits an underlying representation with double prefixation, i.e. the person of possessor is indicated by ʔu-/j-/ʔj- while the number of the possessum by prefixes *n-* or *r-*.⁶ With this analysis, a surface form such as *nʔk'wíʃ* ‘my paper’ is interpreted as retaining the noun class prefix *n-* characteristic of Northern Pame. In contrast, for Central Pame, Gibson and Bartholomew (1979) propose the prefix series *no-/ni-/ni-* and *ro-/ri-/ri-* that conflate person of possessor and number of possessum. This means that for the cognate form *nok'wéʃ* ‘my paper’ Central Pame is thought to have the prefix *no-* plus a stem alternation proper of 1st person which needs to be accounted for independently. Berthiaume’s analysis of the stem change is accounted for by internal assimilation. The data in Chichimeco-Jonaz do not easily yield to either analysis. Under the prefix analysis, for the 29 verbs of class II in Angulo one would have to propose five prefix series *nu-/ni-*, *ku-/ki-*, *ru-/ri-*, *su-/si-* and *tsu-/tsi-*, while an analysis based on retaining a noun class marker would not hold synchronically because the nouns do not have independent absolute forms.

Possession class III in Chichimeco-Jonaz is cognate with class *na-/ngo-/ngo-* in Central Pame, as shown in Table 14. But in Northern Pame, the cognate morphology has disintegrated.

⁶ A cognate *r-* for plural of possessum is also found in Chichimec, but as a suffix. This can be seen in Table 11.

		1SG		2SG		3SG		ABSOLUTE	
		SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS
SG	N.PAM	nhsíʔ	n-h-tsiʔ	nhsíʔ	n-h-tsiʔ	nsíʔ	n-(n)-tsiʔ	nsíʔ	n-tsiʔ
	C.PAM	nats:iʔ	na-ts:iʔ	ŋgotsíʔ	ŋgo-tsiʔ	ŋgotsíʔ	ŋgo-tsiʔ	ŋgotsíʔ	ŋgo-tsiʔ
PL	N.PAM	htʃíʔ	j-h-tsiʔ	htʃíʔ	j-h-tsiʔ	ntʃíʔ	j-n-tsiʔ	*tʃíʔ	j-tsiʔ
	C.PAM	itʃ:iʔ	i-tʃ:iʔ	ntsíʔ	n-tsiʔ	ntsiʔ	n-tsiʔ	ntsiʔ	n-tsiʔ

Table 14. Possession forms of the word for ‘tooth’ in Pame.

In Northern Pame, the class in Table 13 serves as a default for most nouns of the default noun class and for all nouns of the animate class. But in Central Pame, of the 67 nouns in noun class 1 in Gibson and Bartholomew (1979), 29 are found in the possession class of Table 13 while the 38 remaining ones are of the class in Table 14. On the formal level, surface forms pose further challenges for the morphological analysis.

Finally, Chichimeco-Jonaz class IV is cognate with the possession class in Northern Pame given in (a) in Table 15, which serves the instrumental noun class in Table 8. This class merges two distinct possession classes in Central Pame: the one in (a) serves noun class 6 and that of (b) the instrumental noun class 11.

		1SG		2SG		3SG		ABSOLUTE	
		SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS
a.	N.PAM	stiʔés	ʃtʃʔés	ʃkiʔés	ʃkjʔés	sinʔés	ʃiŋʔés	siʔés	ʃiʔés
	C.PAM	tahéf	ta-héf	kihéf	ki-héf	nahéf	na-héf	*nahéf	*na-héf
b.		staʔéogŋ	sta-ʔéogŋ	skiʔéogŋ	ski-ʔéogŋ	snaʔéogŋ	sna-ʔéogŋ	*tʃiʔéogŋ	*tʃi-ʔéogŋ

Table 15. Possession forms of the words for ‘knife’ (a) and ‘broom’ (b) in Pame.

Invariable nouns lie outside the inflectional sub-system made up by the possession classes. In Pame, they inflect for person-number of the possessor by means of the same suffixes used with the dative paradigm of verbs, as shown in (17), adapted from Berthiaume (2003).

(17)	1.DAT	si-k	‘say to me’	1.POSS	sint’æ-k	‘my waist’
N.PAM	2.DAT	si-k’	‘say to you’	2.POSS	sint’æ-k’	‘your waist’
	3.DAT	si-p	‘say to him/her/them’	3.POSS	sint’æ-p	‘his/her/their waist’

In Chichimeco-Jonaz, invariable nouns enter into a possession relationship by means of phrasal compounding with one of the five possessed nominal elements in (18),⁷ which function as genitive classifiers. The examples are in the plural to additionally show that in such a construction, number of the possessum is only marked on the possessed noun.

(18)	ANIMALS	[ná-mbe	urhě-r]	FOOD	[na-nté ⁸	kindi-r]
CHI		1SG.POSS.animal	horse-PL		1SG.POSS-CL.FOOD	soup-PL
		‘my horses’			‘my soups’	

⁷ The same construction is also found in frozen compounds like *katá ngü* (1SG.POSS.face eye) ‘my eye’.

⁸ The root *-nté* is the only one that is no longer a nominal, but a genuine genitive classifier. Etymologically, it is the word for ‘water’. The word for ‘food’ is nowadays *nátsa* ‘my food’, a nominalization of the verb ‘eat’.

PLANTS	[ta- ngwá 1SG.POSS-plant 'my trees'	rígú-r] tree-PL	OTHER	[ná- mbi 1SG.POSS-thing 'my stones'	kúro-r] stone-PL
CLOTHING	[n<ú> nthü 1SG.POSS/clothes 'my shirts'	mátú-r] shirt-PL			

4.2.2. Possession in Otomian. Like Pamean, Atzincan also has possession classes. In Tlahuica, based on the materials in Muntzel (1984), one can propose the three major classes in Table 16. For alienable nouns, their non-possessed form (i.e. absolute) is abbreviated as ABS.⁹

	cl. Ia		cl. Ib	cl. II	
	'wood'	'husband'	'stone'	'house'	'throat'
1SG	m-(n)tsa	we-p-t'uma	βu-nto	n[i]-a:-tho:	p[i]-a:-hi:
2SG	lí-ntsa	we-lí-t'uma	lí-nto	n[i]-ák-tho:	p[i]-ák-hi:
3SG	l-ntsa	we-l-t'uma	lu-nto	ni-tho:	pi-hi:
1PL	ne-hu-ntsa	ne-hu-t'uma	ne-hu-βu-nto	n[i]-a:-kew-tho	p[i]-a:-kew-hi:
2PL	ne-lo-lí-ntsa	ne-lo-lí-t'uma	ne-lo-lí-nto	n[i]-ák-kuk-tho	p[i]-ák-kuk-hi:
3PL	ne-lo-l-ntsa	ne-lo-l-t'uma	ne-lo-lu-nto	ni-tho:-hnə	pi-hi:-hnə
ABS	ntsa	–	ni+nto	–	–
	cl. IIIa		cl. IIIb		
	'belly'	'forehead'	'father'		
1SG	ni-:-pi	pi-:-t ^y i	w[e]a-ta		
2SG	ni-pi	pi-t ^y i	we-ta		
3SG	ni-m-pi	pi-n-t ^y i	we-n-ta		
1PL	ni-:-pi-hnə	pi-:-t ^y i-hnə	wa-ta-hnə		
2PL	ni-pi-hnə	pi-t ^y i-hnə	we-ta-hnə		
3PL	ni-m-pi-hnə	pi-n-t ^y i-hnə	we-n-ta-hnə		
ABS	–	–	–		

Table 16. Possession classes in Tlahuica.

Like in Northern Pame, noun class prefixes in Tlahuica are also retained and attach to the left edge of the word. Class Ia is a default class. Sub-class Ib has only *ni-* and *nu-* class nouns (but members of such classes are also found in classes II and III). The *we-* class of human role nouns split arbitrarily into two sub-classes according to their possession behavior: some nouns inflect as class I and some as class IIIb. This means that membership in possession classes is lexically specified, e.g., body parts are found in all four classes. The few existing sources on Matlatzinca have scarce information about the encoding of possession but reveal a system very different from that of Tlahuica, as illustrated in Table 17, with tonal lowering for the 2nd person. Data is adapted from Bartholomew (nd).

⁹ Given comparable data in sister language Matlatzinca, it is highly likely that the forms for 2nd person possessor also bear contrastive tone.

	‘stomach’	‘head’	‘face’	‘tooth’
1SG	thé-yá	thé-nú	thé-mhí	thé-sibí
2SG	ní-ʔya	ní-ʔnu	ní-ʔmhi	ní-nsibí
3SG	ní-ʔyá	ní-ʔnú	ní-ʔmhí	ní-nsíbí

Table 17. Possession marking in Matlatzinca.

Finally, in clear contrast to what happens in the other groups, in Otomi-Mazahua the marking of possession is syntactic. This is illustrated in Mazahua in Table 18, adapted from Knapp (2008). The markers *nu* and *k’o* are definite determiners, of singular and plural respectively. Possessive determiners occur closer to the noun. Enclitics mark person and number of possessor and serve to disambiguate between a 1st and a 2nd person possessor.

		Number of possessor					
		SG		DU		PL	
Number of possessum	SG	1 st	nu in nzhini=gɔ	EXCL	nu in nzhini=gɔ=ʃe	EXCL	nu in nzhini=gɔ= ^h me
				INCL	nu in nzhini=gɔ=ʃi	INCL	nu in nzhini=gɔ=hi
		2 nd	nu in nzhini=ge	nu in nzhini=ge=ʃi		nu in nzhini=ge=hi	
		3 rd	nu o nzhini				
Number of possessum	PL	1 st	k’o in nzhini=gɔ	EXCL	k’o in nzhini=gɔ=ʃe	EXCL	k’o in nzhini=gɔ= ^h me
				INCL	k’o in nzhini=gɔ=ʃi	INCL	k’o in nzhini=gɔ=hi
		2 nd	k’o in nzhini=ge	k’o in nzhini=ge=ʃi		k’o in nzhini=ge=hi	
		3 rd	k’o o nzhini				

Table 18. Possession marking of the noun for ‘ox’ in Mazahua.

A handful of nouns in Mazahua require a palatalized stem when possessed by a 1st or a 2nd person. The nouns do not form a coherent natural class (e.g. ‘bench’, ‘blood’, ‘bone’, ‘door’, ‘eye’, ‘lamb’, ‘song’, ‘stone’, ‘woman’, etc.). One example is given in (a) in Table 19 to be compared with an invariable noun in (b), adapted from Knapp (2008). At least five other body part nouns require a prenasalized stem for the 3rd person. The noun *tɔ* ‘eye’ in (c) is one of them.

	(a) ‘stone’	(b) ‘chocolate’	(c) ‘eye’
1 st	ⁿ d → ⁿ dʒ in ⁿ dʒóho=gɔ	ⁿ d in ⁿ dɛk ^h i=gɔ	t → tʃ in tʃɔ=gɔ
2 nd	in ⁿ dʒóho=ge	in ⁿ dɛk ^h i=ge	in tʃɔ=ge
3 rd	ⁿ d o ⁿ dóho	o ⁿ dɛk ^h i	t → ⁿ d o ⁿ dɔ

Table 19. Nouns with and without stem alternations in Mazahua.

Stewart (1966) reports that kinship terms in Mazahua may take the marker *mi* for a 1st person possessor, as in (19). This marker is cognate with the prefix *m-* of possession class Ia in Tlahuica. It is also found in Old Otomi, where it was used for a female speaker. Male speakers used *ma* instead, which later generalized in all Otomi languages as the only possessive marker for 1st person, except in the conservative Otomi languages of Tilapa and Ixtenco, where the old speaker-oriented gender distinction was kept. This is illustrated in (20) in Tilapa Otomi with the markers *mi* and *mu* /mi/ (the reflex of old *ma*).¹⁰

¹⁰ The kinship terms in (20), *hephesita* (from Spanish *jefecita* ‘little boss (female)’) and *tömbè* (‘old lady’), both mean ‘mother’, but the former is a term of male speech, while the latter of female speech.

- (22) ma=dá kwaht-a[=ma beni]=**he**
 N.OTO MOT=1.CPL put.away-B=1POSS suitcase=PL.EXCL
 ‘(So that) we went to put our suitcases away.’ (Txt)

4.2.3. About the marking of possession in Oto-Pamean. Nouns with canonical inalienable semantics such as body parts and kinship terms, together with a handful of cultural artefacts, such as pieces of clothing and food (i.e. tortilla, maize, etc.) are obligatorily possessed in Oto-Pamean, just as they are in most Mesoamerican languages. In Otomi-Mazahua, without possession classes, obligatorily possessed nouns could be said to form the class of inalienable nouns.

But what happens in the other languages with possession classes? Perhaps the best way to tackle possession classes in Pamean and Atzincan as a grammatical phenomenon is to regard them as inflectional classes. In other words, they constitute a case of inflectional allomorphy whose selection is largely determined by the lexicon. Indeed, such classes have little to do with the dimension of alienability.

It is only in Chichimeco-Jonaz where one could argue that the division between variable and invariable nouns reflects a split between inalienable and alienable nouns. This is because variable nouns cannot exist in absolute form and must thus be obligatorily possessed.

However, it is less clear what the status of variable nouns is. From the existing materials we can elucidate that: (a) they include borrowings and old compounds. This means that they are nouns with special stems that lie outside the morphology, which is hardly a property of semantic alienability; and (b) they include nouns which pertain to semantic fields characteristically associated with inalienability, such as body parts, clothing, etc. Again this makes possession classes look like an inflectional phenomenon rather than a semantic property of nouns.

5. Verbal inflection

Verbs in all Oto-Pamean languages are inflected for tense/aspect/mood values by means of markers preposed to the verbal stem, which may be referred to as ‘inflectional formatives’. Often inflectional formatives also encode agreement with subject in person and number in a portmanteau fashion.

In Otomi-Mazahua, inflectional formatives share properties of affixes and properties of clitic words. For example, like prefixes, they always occur preverbally. This is illustrated in (23), where it is shown in Tilapa Otomi that the occurrence of the formative *ra* for the 3rd person incomplete realis is ungrammatical in any position but preverbally.

- (23) (*ra) pwes (*ra) nt’a (*ra) rú (*ra) mbek’o *(ra=)’em-bi (*ra)
 TIL so one SG.3POSS brother.in.law 3.INCPL.R=say-3DAT
 ‘So one of his brothers-in-law is telling him.’ (Txt)

But as clitics, they may be phonologically hosted on a word other than the verb, like on a preverbal adverbial or a pronominal, with subsequent phonological adjustments, as exemplified in (24), which shows two instances of the same formative *ta* for 3rd person completive irrealis.

- (24) **ta**=nde o hin=**da** nde
 TIL 3.CPL.IRR=want or NEG=3.CPL.IRR want
 ‘May he want or may he not want.’ (Txt)

Formatives often depend on a word that has nothing to do with their syntactic domain. This is observed in (25) by the placement of hesitation pauses, indicated by #.

- (25) a. #tú=mbe='mbe=[tú...# tú=^htsu='mbe]=r t'egi=gwa=a#
 TIL 1.CPL.R= SS/go.DU.EXCL=PL.EXCL=1.CPL.R 1.CPL.R=reach=PL.EXCL=SG car=here=PFCL
 'We went to...to take the bus here.' (Txt)
- b. #nthönt'o té='éh=a[=tú... # ^hkundy]=a ni 'i=a#
 soon 1.CPL.R=come.AS=PFCL=1.CPL.R grind.AS=DET.SG DEF.SG chili=PFCL
 'Soon I got here to...grind the chili.' (Txt)

As the entire verbal inflection of Otomi-Mazahua is achieved by such formatives, the verbal inflection in these languages is entirely periphrastic. However, one cannot analyse the formatives as auxiliary verbs, at least not synchronically. In the literature, the formatives of other Oto-Pamean groups have been treated as prefixes, except for Matlatzinca in Bartholomew (nd). In Atzacan, their structure indeed suggests that they are clitic words like in Otomi-Mazahua, so we could think of it as an Otomian phenomenon. In Pamean, however, they behave like canonical prefixes. In Chichimeco-Jonaz they are toneless and monosyllabic. In Pame, as illustrated in Table 20, the prefixes that are not reduced to a single consonant have a centralized vowel as a nucleus and, at least in Berthiaume (2003), they are seen as responsible for producing stem modifications at the surface level.

	INCPL		CPL		POT	
	SURFACE	ANALYSIS	SURFACE	ANALYSIS	SURFACE	ANALYSIS
1SG	ləkáʔs	lə-káʔs	nəkʷáʔs	nə-w-káʔs	nʔkʷáʔs	nʔ-w-káʔs
2SG	k'ətʃáʔs	k'ə-j-káʔs	nətʃáʔs	nə-j-káʔs	nʔkáʔs	nʔ-káʔs
3SG	kwáʔs	w-káʔs	dəkʷáʔs	də-w-káʔs	nʔkáʔs	nʔ-káʔs

Table 20. Partial paradigm of the verb 'soften' in Northern Pame.

5.1 Values of TAM

All systems have at least two sub-paradigms for the basic aspects, incomplete vs. complete (traditionally called 'present' and 'past') and one or various sub-paradigms for the irrealis (often called 'future' and/or 'potential'). Chichimeco-Jonaz has a rich system; it is described by Angulo (1933) and Lastra (1984) as having three different past tenses and a sequential tense used in clause combining constructions. Examples are presented in Table 21.

PRS	<i>e-nú</i>	's/he's seeing it'
IMM.PST	<i>zu-nú</i>	's/he's just seen it'
REC.PST	<i>ku-nú</i>	's/he's seen it'
ANT.PST	<i>u-nú</i>	's/he saw it'
FUT	<i>ga-nú</i>	's/he will see it'
POT	<i>mu-nú</i>	's/he'd see it'
SEQ	<i>ru-nú</i>	'and she see it'

Table 21. TAM forms (3rd person singular) of the verb 'see' in Chichimeco-Jonaz

But it is in the conservative Otomi languages, like Tilapa Otomi, where one observes the most developed TAM system within Oto-Pamean with at least 12 values. This is shown in Table 22 with the forms of the verb for ‘fall’.

	REALIS		IRREALIS	
INCPL	ra ^h tagi	‘s/he’s falling’	tra tagi	‘s/he’ll be falling’
HAB	ru ^h tagi	‘s/he falls’		
IMPF	ramá ^h tagi	‘s/he was falling’		
IMPF.HAB	múru ^h tagi	‘s/he used to fall’		
CPL	bi tagi	‘s/he fell’	ta tagi	‘s/he’ll fall’
PST			tigi tagi	‘s/he would’ve fallen’
PRF	xún tagi	‘s/he’s fallen’	xti tagi	‘s/he’ll have fallen’
PPRF	xkán tagi	‘s/he’d fallen’		
PROSP			xtá tagi	‘s/he’s about to fall’

Table 22. TAM forms (3rd person singular) of the verb ‘fall’ in Tilapa Otomi.

5.2 Conjugation classes

Oto-Pamean languages have conjugation classes whose membership is listed in the lexicon. This means that membership is not accounted for by properties of the lexeme, whether they be phonological, semantic or valence oriented. In this light, there are two main ways in which verbs are classified in Oto-Pamean for inflectional purposes: (i) ‘formative-based classes’, which are based on specific different sets of inflectional formatives which verbs may select; and (ii) ‘stem-based classes’, which are based on different stem alternation patterns different verbs may display in their paradigm. Let us see them in order.

5.2.1. Formative-based classes. Chichimeco-Jonaz has up to seven formative-based conjugation classes. Such classes display up to five different ways of partitioning the paradigm regarding the expression of number of the subject. The different forms are given in Table 23, adapted from Angulo (1933).

	1SG	1DU	1PL	2	3SG/DU	3PL	1SG	1DU	1PL	2	3SG/DU	3PL	1SG	1DU	1PL	2	3SG/DU	3PL
	cl. I						cl. II						cl. III					
PRS	e		ki		e		tu		su		u		tu		su		u	e
IMM.PST	u		i		zu		u		i		zu		u		i		zu	zu
REC.PST	ku		ki		ku		ku		ki		ku		ku		ki		ku	ku
ANT.PST	tu		ki		u		tu		ki		u		tu		su		u	e
FUT	ga		ki		ga		gu		ki		ga		gu		ki		ga	ga
POT	nu		mi		mu		nu		mi		mu		nu		mi		mu	mi
SEQ	ra		gi		ru		ra		gi		ru		ra		gi		ru	ru
	cl. IV						cl. V						cl. VI					
PRS	tu		ka		u		e	e	u	ki	e	e	e	e	u	ki	e	e
IMM.PST	u		i		zu		ka	ki	ki	ki	ku	ku	sa	si	si	sa	sa	sa
REC.PST	ku		e		ku		ka	ki	ki	ki	ku	ku	sa	si	si	sa	sa	sa
ANT.PST	tu		ka		u		ta	ti	ti	ki	u	u	ta	ti	ti	sa	ta	ta
FUT	ga		ka		ga		ta	ti	gu	ki	ga	ga	ta	ti	gi	sa	ta	ta
POT	mu		ma		ma		ma	ma	mu	mi	ma	mi	na	ni	ni	za	na	na
SEQ	ra		ga		ru		na	na	nu	mi	ra	ra	na	ni	ni	za	na	na

cl. VII				
PRS	ti	ti	si	i
IMM.PST	i	i	i	i
REC.PST	sa	si	sa	sa
ANT.PST	ta	ti	sa	ta
FUT	ta	ti	sa	ta
POT	na	ni	za	na
SEQ	na	ni	za	na

Table 23. The formative-based conjugation classes of Chichimeco-Jonaz.

While membership to the classes is arbitrary, verbs with middle semantics (i.e. reciprocal, reflexive, posture, grooming, etc.) are found in class VII, as shown in (26). This is a characteristic feature of Oto-Pamean, where languages have active verbs in one class and their middle voice counterparts are inflected in another class.

(26) *e-nú* (I) ‘I see it’ ~ *ti-nú* (VII) ‘I see myself’

Atzincan has six formative-based classes, with slight differences between Tlahuica and Matlatzinca. They are shown in Table 24, adapted from Carranza-Martínez (2013), (dual forms are not included; the 1st plural form is the inclusive; and ‘Σ’ represents the stem).

	PRS	PRF	POT	PRS	PRF	POT	PRS	PRF	POT
	cl. I (tr)			cl. II (tr)			cl. III (tr)		
1SG	tuh-	to-	ru-	tu-tú-	to-	ru-	tuht-	to-	ru-
2SG	ʔih-	ʔo-	ri-	ʔi-tú-	ʔo-	ri-	ʔiht-	ʔo-	ri-
3SG	kuh-	tu-	tátu-	ku-tú-	tu-	tá-tu-	kuht-	tu-	tátu-
1PL	khwen-	kho-	ru-Σ-he	khwen-tú-	kho-	ru-Σ-he	khwent-	kho-	ru-Σ-he
2PL	čhen-	čo-	ri-Σ-he	čhen-tú-	čo-	ri-Σ-he	čhent-	čo-	ri-Σ-he
3PL	ron-	tu-Σ-he	tátu-Σ-he	ron-tú-	tu-Σ-he	tá-tu-Σ-he	ront-	tu-Σ-he	tátu-Σ-he
	cl. IV (intr)			cl. V (intr)			cl. VI (intr)		
1SG	tu-	taʔ-	ruh-	tu-te-	tah-te-	ruh-te-	tun-	tan-	run-
2SG	ʔi-	ʔiʔ-	riʔiʔ-	ʔi-te-	ʔeʔ-	riʔe-	ʔin-	ʔín-	riʔin-
3SG	ku-	kaʔ-	taʔ-	ku-te-	re-	táʔ-re-	kun-	kan-	tan-
1PL	khwen-	khweʔ-	rukhwēʔ-	khwen-te-	khweʔ-	rukhwēʔ-	khwen-	khwen-	rukhwēn-
2PL	čhen-	čeʔ-	ričheʔ-	čhen-te-	čeʔ-	ričheʔ-	čhen-	čhen-	ričhen-
3PL	ron-	roʔ-	taroʔ-	ron-te-	roʔ-re-	tároʔ-re-	ron-	ron-	taroʔn-

Table 24. The formative-based conjugation classes of Matlatzinca.

Membership to classes in Atzincan is more clearly associated with transitivity. The transitive verbs in the language fall into classes I-III; intransitives into classes IV-VI. Transitive class I and intransitive class IV serve as morphological defaults (and most likely point to the primary verbs of the proto-language). The distinctions making up the other classes are formally derived from these two classes, although the derivation is only historical. For example, classes II and V originate, respectively, from historical causative verbs (derived by prefix {*tu-*}) and historical middle voice verbs (derived by prefix {*te-*}). For more details see Palancar and Carranza (forthcoming).

Otomi-Mazahua has considerable internal diversity regarding conjugation classes. The conservative Otomi languages like Tilapa Otomi, Acazulco Otomi and Ixtenco Otomi have preserved the three main classes from Old Otomi in Cárceres (1907/1580). In other varieties, the system has undergone considerable change. The verbs in Table 25 illustrate the classes of Tilapa Otomi.

		‘water (cornfield)’ cl. I (tr/intr)		‘cut up wood’ cl. II (intr)		‘pray’ cl. III (tr/intr)	
		REALIS	IRR	REALIS	IRR	REALIS	IRR
INCPL	1 st	trá xithe	gra xithe	trá xu	gratu xu	trátí xadi	grati xadi
	2 nd	grá xithe	gra xithe	grá xu	gragu xu	grátí xadi	grati xadi
	3 rd	ra xithe	tra xithe	ra xu	tra xu	rati xadi	trati xadi
CPL	1 st	tú xithe	gi xithe	túdú xu	gutu xu	tú xadi	giti xadi
	2 nd	gú xithe	gi xithe	gúgú xu	gugu xu	gú xadi	giti xadi
	3 rd	bi xithe	ta xithe	bi xu	ti xu	bi xadi	ti xadi
PRF	1 st	xtú xithe	xkugu xithe	xtúdú xu	xtigutu xu	xtú xadi	xkugu xadi
	2 nd	xkú xithe	xkugu xithe	xkúgú xu	xtigugu xu	xkú xadi	xkigi xadi
	3 rd	xún xithe	xti xithe	xpi xu	xtigi xu	xpi xadi	xtigi xadi

Table 25. The formative-based conjugation classes of Tilapa Otomi.

Class III results from a merger of two distinct classes in Otomian, whose reflexes survive in Atzincan classes II and V. As a result of this merger, Otomi class III contains both highly transitive verbs and intransitive verbs with middle voice semantics. For such a historical merger to be able to happen, the morphology involved had to be perceived by speakers at the time as functionally opaque. Otomi class II contains intransitive verbs only, and most of its members refer to activities performed by animates, mostly humans. In Eastern Otomi, class II has an innovative nasal marker distributed in the paradigm, most probably as an extension of the middle voice marker, which is also a nasal (Palancar 2006). This language is also innovative in having developed an intransitive class IV, as a heteroclite of classes II and III. These facts are summarized in Table 26, adapted from Voigtlander & Echegoyen (1985).

		‘gather’ cl. Ia (tr/intr)	‘save’ cl. Ib (tr)	‘walk’ cl. II (intr)	‘fix’ cl. III (tr/intr)	‘hurry’ cl. IV (intr)
INCPL	1 st	dí kho	dín yäni	dí ’yo	dídí ho ^h ki	dídí xoni
	2 nd	gí kho	gín yäni	gí ’yo	gídí ho ^h ki	gídí xoni
	3 rd	(i) kho	in yäni	(i) ’yo	(i)dí ho ^h ki	(i)dí xoni
CPL	1 st	dá kho	dá yäni	dán ’yo	dá ho ^h ki	dán xoni
	2 nd	gá kho	gá yäni	gán ’yo	gá ho ^h ki	gán xoni
	3 rd	bi go	bi yäni	bin ’yo	bi ho ^h ki	bin xoni
IRR	1 st	ga kho	gan yäni	dan ’yo	ga ho ^h ki	dan xoni
	2 nd	gi kho	gin yäni	gan ’yo	gi ho ^h ki	gan xoni
	3 rd	da go	da yäni	din ’yo	di ho ^h ki	din xoni

Table 26. The formative-based conjugation classes of Eastern Otomi.

5.2.2. Stem-based classes. Apart from classes defined by sets of inflectional formatives, Oto-Pamean verbs also fall into classes attending to the stem alternation patterns they display. This happens in all groups except in Atzinca where the historical allomorphy of

stems was effaced. The simple system is found in Otomi-Mazahua, the most complex in Pamean.

In the Otomi languages with formative-based classes, class I verbs have two stems: a basic ‘primary stem’ and a derived ‘secondary stem’, e.g. E.OTO *kho* vs. *go* ‘gather’. In Mezquital Otomi, a variety of Northern Otomi, the nasal marker characteristic of class II in Eastern Otomi merged with the stem as a typical prefix. In this innovative language, all formative-based classes were also levelled to class I, but the old stem alternation distinctions were retained, preserving old class contrasts, but this time realized by stem alternations only. This is shown in Table 27.

		‘gather’ cl. I		‘walk’ cl. II		‘fix’ cl. III	
INCPL	1 st	dí	kho	dí	’yo	dí	ho ^h ki
	2 nd	gí	kho	gí	’yo	gí	ho ^h ki
	3 rd	(i)	kho	(i)	’yo	(i)	ho ^h ki
CPL	1 st	dá	kho	dá	’<ñ>o	dá	ho ^h ki
	2 nd	gá	kho	gá	’<ñ>o	gá	ho ^h ki
	3 rd	bi	go	bi	’<ñ>o	bi	ho ^h ki
IRR	1 st	ga	kho	ga	’<ñ>o	ga	ho ^h ki
	2 nd	gi	kho	gi	’<ñ>o	gi	ho ^h ki
	3 rd	da	go	da	’<ñ>o	da	ho ^h ki

Table 27. The stem-based classes of Mezquital Otomi

In Pamean the situation is complex, at least in Chichimeco-Jonaz. Based on a sample of 170 verbs from Angulo (1933), Palancar and Avelino (forthcoming) propose that verbs fall into at least 14 different stem patterns. In Table 28, I present seven of the most representative ones. The different stems involved in the paradigms are labelled A, B, C, etc. attending to how they are distributed in the paradigm. Stem pattern 2 could be considered a default paradigm. This pattern serves as a base for the structure of patterns 3 to 6.

	1 st	2 nd	3SG/DU	3PL	1 st	2 nd d	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
	Stem pattern 1				Stem pattern 2				Stem pattern 3				Stem pattern 4			
PRS	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
IMM.PST	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
REC.PST	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
ANT.PST	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
FUT	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
POT	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
SEQ	A	A	A	A	A	A	A	B	A	A	C	B	A	D	A	B
	‘laugh’				‘see’				‘heat’				‘do’			
PRS	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
IMM.PST	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
REC.PST	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
ANT.PST	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
FUT	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
POT	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá
SEQ	tér	tér	tér	tér	nú	nú	nú	nhú	pan	pan	mban	phan	tsá	tcá	tsá	tshá

	Stem pattern 5				Stem pattern 6				Stem pattern 7			
PRS	A	A	A	B	A	A	A	B	A	A	A	A
IMM.PST	A	A	C	B	A	D	A	B	D	D	D	A
REC.PST	E	A	C	B	A	D	C	B	E	D	C	A
ANT.PST	A	A	C	B	A	D	C	B	A	D	C	B
FUT	A	A	C	B	A	D	C	B	A	D	C	B
POT	A	A	A	B	A	D	A	B	D	D	D	A
SEQ	A	A	A	B	A	D	A	B	A	D	A	A

	'learn'				'want'				'narrate'			
PRS	<i>pen</i>	<i>pen</i>	<i>pen</i>	<i>phen</i>	<i>ʔi</i>	<i>ʔi</i>	<i>ʔi</i>	<i>rʔi</i>	<i>pín</i>	<i>pín</i>	<i>pín</i>	<i>pín</i>
IMM.PST	<i>pen</i>	<i>pen</i>	<i>ven</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ʔi</i>	<i>rʔi</i>	<i>ngwín</i>	<i>ngwín</i>	<i>ngwín</i>	<i>pín</i>
REC.PST	<i>mben</i>	<i>pen</i>	<i>ven</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ndí</i>	<i>rʔi</i>	<i>mbín</i>	<i>ngwín</i>	<i>mín</i>	<i>pín</i>
ANT.PST	<i>pen</i>	<i>pen</i>	<i>ven</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ndí</i>	<i>rʔi</i>	<i>pín</i>	<i>ngwín</i>	<i>mín</i>	<i>mbín</i>
FUT	<i>pen</i>	<i>pen</i>	<i>ven</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ndí</i>	<i>rʔi</i>	<i>pín</i>	<i>ngwín</i>	<i>mín</i>	<i>mbín</i>
POT	<i>pen</i>	<i>pen</i>	<i>pen</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ʔi</i>	<i>rʔi</i>	<i>ngwín</i>	<i>ngwín</i>	<i>ngwín</i>	<i>pín</i>
SEQ	<i>pen</i>	<i>pen</i>	<i>pen</i>	<i>phen</i>	<i>ʔi</i>	<i>tí</i>	<i>ʔi</i>	<i>rʔi</i>	<i>pín</i>	<i>ngwín</i>	<i>pín</i>	<i>pín</i>

Table 28. The main stem-based classes of Chichimeco-Jonaz.

5.3. Tone and inflection

In general, compared to other Oto-Manguen languages, tone in Oto-Pamean does not play a major role in verbal inflection, but tonal alternations are nonetheless found in paradigms. In Pamean, they are restricted to a few verbs and they co-occur with the various stem patterns. This is illustrated in Table 29, where the focus is on the forms for the 2nd person as they all display the inverse tone of (most of) the forms for the 1st person. This marking mechanism is reminiscent of possession class I in Table 12.

	'win' (pattern 2 or 4)				'work' (subtype of pattern 5)				'hurt' (subtype of pattern 6)			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tan</i>	<i>tán</i>	<i>tan</i>	<i>rhán</i>	<i>tsa</i>	<i>fa</i>	<i>tsa</i>	<i>tsha</i>
IMM.PST	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tan</i>	<i>tán</i>	<i>tan</i>	<i>rhán</i>	<i>tsá</i>	<i>fa</i>	<i>tsá</i>	<i>tshá</i>
REC.PST	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>ndan</i>	<i>tán</i>	<i>rán</i>	<i>rhán</i>	<i>tsa</i>	<i>fa</i>	<i>za</i>	<i>tsha</i>
ANT.PST	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tan</i>	<i>tán</i>	<i>rán</i>	<i>rhán</i>	<i>tsá</i>	<i>fa</i>	<i>zá</i>	<i>tshá</i>
FUT	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tan</i>	<i>tán</i>	<i>rán</i>	<i>rhán</i>	<i>tsá</i>	<i>fa</i>	<i>za</i>	<i>tshá</i>
POT	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tán</i>	<i>tán</i>	<i>tán</i>	<i>rhán</i>	<i>tsa</i>	<i>fa</i>	<i>tsa</i>	<i>tsha</i>
SEQ	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	<i>tan</i>	<i>tán</i>	<i>tan</i>	<i>rhán</i>	<i>tsá</i>	<i>fa</i>	<i>tsá</i>	<i>tshá</i>

Table 29. Cases of 2nd person marking indicated by tone inversion in Chichimeco-Jonaz.

A rare case of tone inversion has been reported in Acazolco Otomi by Hernández-Green (forthcoming) where it applies to stems of class I and II verbs for the building of a few specific tenses, for example the habitual realis. This is shown in (27) when compared with the lexical tone in the incomplete realis. If the verb has a stem formative (27a), or the human antipassive *-te* (27b) or is a compound (27c), the root receives a high tone. If the verb has a monosyllabic root (27d) or has the enclitic *=tshé* 'alone' (27e), the root receives a low tone.

(27)	INCPL.REALIS				HAB.REALIS	
a.	<i>ra=zo+ni</i>	‘s/he's weeping’	/L/	> /H/	<i>an=zó+ni</i>	‘s/he weeps’
	<i>ra=hú+xi</i>	‘s/he's whistling’	/H/	> /H/	<i>an=hí+xi</i>	‘s/he whistles’
b.	<i>ra=tsah-te</i>	‘it's biting (people)’	/L/	> /H/	<i>an=tsáh-te</i>	‘it bites (people)’
	<i>ra=núh-te</i>	‘s/he's seeing (people)’	/H/	> /H/	<i>an=núh-te</i>	‘s/he sees (people)’
c.	<i>ra=tü-the</i>	‘s/he's thirsty’ (die-water)	/L/	> /H/	<i>an=tü-the</i>	‘s/he's often thirsty’
d.	<i>ra=nú</i>	‘s/he's seeing it’	/H/	> /L/	<i>an=nü</i>	‘s/he often sees it’
	<i>ra=tsa</i>	‘it's biting it’	/L/	> /L/	<i>an=tsa</i>	‘it bites it’
e.	<i>ra=hú+x=tshe</i>	‘s/he's whistling alone’	/H/	> /L/	<i>an=hú+x=tshe</i>	‘s/he whistles alone’
	<i>ra=zo+n=tshe</i>	‘s/he's weeping alone’	/L/	> /L/	<i>an=zo+n=tshe</i>	‘s/he weeps alone’

Oto-Pamean inflectional tone is typologically atypical, and thus theoretically challenging for a number of reasons. For example, a specific tone value, whether it be a high or a low tone, is very rarely associated with a specific grammatical function; the only exception being the high tone marking 3rd person possessor in Otomi. Also, the tonal alternations in inflection involve some type of tone inversion. In Otomi, the inversion is contextual and depends on word structure. Finally, the distribution of tone is partly lexical and partly morphological. In the Otomi case, tone inversion is restricted to certain cells in the paradigm of a verb of a certain class. Once all conditions are met, it applies regularly. In Chichimeco-Jonaz, however, all changes are lexically stored. Tone inversion does not have a clear-cut grammatical function in either system.

5.4. More about verbal inflection

Deixis and motion. Verbs in Oto-Pamean languages often inflect for deictic information. Otomi has again the most developed system. An example is given in Table 30.

	REALIS		IRR
AMBUL	a ^h tagi	‘s/he falls all about the place’	–
CISLOC	bá ^h tagi	‘s/he's coming and falling’	–
PRF.TRANSLOC	xpán tagi	‘s/he's just fallen away somewhere’	–
PPRF.TRANSLOC	xkwá tagi	‘s/he'd just fallen away somewhere’	–
CPL.TRANSLOC	bwu ^h tagi	‘s/he fell away somewhere’	tü tagi ‘s/he'll fall away somewhere’
ANDAT	ar ^h tagi	‘s/he passes by there and fall’	tí tagi ‘s/he'll pass by over there and fall’
ADLAT	–		taga tagi ‘s/he'll be going there and fall’

Table 30. Deictic paradigm of the verb for ‘fall’ in Tilapa Otomi.

Adverbial inflection. The verbal inflection in Otomi-Mazahua includes formatives which have the function of registering a focused adverbial in the clause. This is shown in the contrast in (28), adapted from Stewart (1966: 131, 50). The morphosyntax of focus adverbials is restricted to a closed class with a few members; the one in (28c) does not require special inflection, i.e. the verb selects the same inflectional formatives of the basic paradigm, as in (28a).

(28)	a. ró=ñá=gö	b. [naho] rvá=ñá=gö	c. [khá] ró=khwari=gö
MAZ	1.PST=speak=1 ‘I spoke.’	well 1.PST.ADV=speak=1 ‘I spoke well.’	almost 1.PST=finish=1 ‘I'm almost finished.’

Mazahua has three different sub-paradigms of adverbial inflection depending on three classes of adverbs. This is shown in Table 31.

		Basic	Adverbial				
			cl. I	cl. II	cl. III		
INCPL	1 st	rí	rgá	rgá	rrí/rrú	cl. I	<i>na ho('o)</i> 'well'
	2 nd	ín	gí	ní	ní		<i>na nihi</i> 'fast'
	3 rd	Ø	ga	ní	ní		<i>hamnch'a</i> 'slowly'
CPL	1 st	ró	rvá	rvá	rvá		<i>ts'ë ts'ë</i> 'little by little'
	2 nd	ín	ví	ví	ví		<i>ndahmetho</i> 'in/for a while'
	3 rd	ó	vá	vá	ví		<i>hã (c'o)</i> 'how'
FUT	1 st	rá	rga	rga	rga	cl. II	<i>ts'i</i> 'early'
	2 nd	rí	rgi	rgi	rgi	cl. III	<i>ya</i> 'just'
	3 rd	ra	rgá/rgí	rgí	rgí		<i>dya be</i> 'still'
POT	1 st	ri	rva	rva	rva		<i>ya xo</i> 'also'
	2 nd	ri	rvi	rvi	rvi		
	3 rd	ri	rva	rva	rví		

Table 31. Adverbial inflection in Mazahua.

In Otomi languages, adverbial inflection registers a wider range of adverbial phrases including PPs, which may not be fronted to initial position (Hernández-Green 2015: 98).

- (29) **dá=k'úhki=ga** [ko yá-m ts'í]
 ACA 1.CPL.GEN=tear[3.OBJ]=1 with DEM.PL-1POSS tooth
 'I tore it with my teeth.'

But when the construction involves an adverb, fronting is obligatory (Echegoyen & Voigtlander 2008: 4).

- (30) maske [asta nikheya] dá=n=nü=wi ya (*asta nikheya)
 E.OTO but until next.year 1.IRR.ADV=MID=see=DU PTCL
 'But we won't see each other again until next year.'

The most natural context for adverbial inflection to occur is when the participant bearing adverbial meaning is in focus, as in the wh-words in (31b) or the clefts in (31b-c).

- (31) a. ['bē='a] ge ga=tsi?
 E.OTO what=DEM.SG COP 1.IRR=ingest[3.OBJ]
 'What is it that I'm going to eat?'
 b. ['bē='a] ge dá=xah=ma thä ya?
 what=DEM.SG COP 1.IRR.ADV=peel.corncob.AS/B[3.OBJ]=1POSS maize PTCL
 Instrument: 'What is it that I shall harvest my corn with?'
 Reason: 'For what reason is it that I shall harvest my corn?' (E&V 2008: 320)

- (32) a. [ge='a] dí-m=män=ga='a=ya
 E.OTO COP=DEM.SG 1.INCPL-INFL=say.AS[3.OBJ]=1EMPH=DEM.SG=PTCL
 'That is what I mean.' (E&V 2008: 115)

b. [ge='a] g-dí=pa'm-a=kaphe
 COP=DEM.SG 2-IRR.ADV=heat[3.OBJ]-B=coffee
 Instrument: 'It's that that you may heat coffee with.' (V&E 1985: 223)

c. [ge='a] í-n=hah-te
 COP= DEM.SG 3.CPL.ADV-INFL=cheat-HUM.ANTIP
 Manner: 'It's so that he cheated.' (V&E 1985: 222)

Bare NPs in focus can also trigger adverbial inflection, but only in Eastern Otomi, which suggests that in this Otomi language the syntax of adverbial registration has been realized as the syntax of oblique applicatives. This is illustrated in (33), adapted from Voigtlander and Echegoyen (1985: 220, 306).

(33) a. [ra 'ye] na=nte yu paxi
 E.OTO SG rain 3.INCPL.ADV=grow PL plant
 'WITH THE RAIN, plants are alive.'
 'IT'S WITH RAIN that plants are alive.'

b. [ra koni] í=ts'>ix=ra phani
 SG hackberry.rope CPL.ADV=<PASS>carry.AS/B[3.OBJ]=SG mule
 'WITH A HACKBERRY ROPE, the mule was brought.'
 'IT'S WITH HACKBERRY ROPE that the mule was brought.'

In Acazolco Otomi there are three different paradigms to register a focused adverbial phrase (also sensitive to the conjugation class of the verb in question). The basics of the system are given in Table 32, from Hernández-Green (2015: 409 ff). The locative paradigm is used with a focused locative phrase, the adverbial paradigm is used when a given adverbial phrase of manner, time or duration is in focus; and the general one is used for anything else (means, instrument, allative, ablative, etc.).

Class		LOC		ADVERBIAL				GENERAL			
		REALIS		IRR		REALIS		IRR			
		INCPL	CPL	CPL	PST	INCPL	CPL	CPL	PST		
I	1 st	dí	dra	dá	dá	ngá	drá	dá	dá	ngá	
	2 nd	gí	gru	gí	dí	ngí	grí	gí	dí	ngí	
	3 rd	bí/rá	ga	gá	ká	ngá	rí	í	dí	ngí	
II	1 st	dí	dradi	dádí	dádí	ngádí	drádí	dádí	dádí	ngádí	
	2 nd	gí	gradi	gádí	dádí	ngádí	grádí	gádí	dádí	ngádí	
	3 rd	bí/rá	ga	gá	ká	ngá	rá	á	dá	ngá	
III	1 st	dídi	dra	dá	dá	ngá	drá	dá	dá	ngá	
	2 nd	gídi	gru	gí	dí	ngí	grí	gí	dí	ngí	
	3 rd	bídi	gi	gí	kí	ngí	rí	í	dí	ngí	

Table 32. Paradigms for adverbial registration in Acazolco Otomi.

Besides adverbial phrases in focus in a given clause, in Otomi discourse the adverbial participant to which this special inflection makes reference is commonly an established topic. This can be seen in (34) where the reference is anaphoric, or in (35) where it is cataphoric.

(34) 'neh=a mbaha g<w>u=^htyü=hu [yo sku-za],
 TIL and=PFCL SS/go.PL.INCL <CISLOC>1.CPL.IRR=bring[3.OBJ]=PL.INCL two DIM-stick

para gu=^hka'tsu=hu,
 PURP 1.CPL.IRR=burn.AS[3.OBJ]=PL.INCL

para gata-tu=^hpotu=hu
 PURP 1.IRR.ADV-1=make.tortillas.AS=PL.INCL

'And now we're going to bring two sticks to burn them, to make tortillas that way/with them.'

(35) ya ná-m='buh=t='a na ra xitsu:
 E.OTO PTCL 3.INCPL.ADV-INFL=live.AS=DEL.B=DEF.SG DEM.SG SG woman

[nge hingi zám=mi n'da ra n'yohu]
 that NEG [3.INCPL]persist.AS=DU one SG man

'In that way lives this woman, in that she never stays with one man.' (E&V 2008: 48)

5.5. Argument encoding

We have already seen that information about the person of the subject is encoded by way of distinctions in the inflectional formatives. Object of 1st and 2nd person is encoded by means of person suffixes that, when affixed, generate a string of intricate morphophonological adjustments on both simple and complex stems. Object anaphora of a 3rd person is realized by a bare stem (or a zero suffix). Examples are given in Table 33 from Tilapa Otomi.

	'pick up'	'frighten'	'pull out'
1OBJ	tsi+n-gi	^h pih-ki	khuy-gi
2OBJ	tsi+n-k'i	^h pi?-k'i	khuy-k'i
3OBJ	tsi+ni-Ø	^h pi+di-Ø	khuh+ki-Ø

Table 33. Person marking of object in Tilapa Otomi.

In Atzincan, transitive verbs may receive benefactive applicative {-pi} where the beneficiary is encoded as the primary object. The case is illustrated in Table 34, adapted from Martínez-Ortega (2012). The table also shows that verbs change valence by changing inflectional class. A verb such as 'wake up' can be inflected both as intransitive or as transitive, but in different classes, hence the selection of different inflectional formatives.

	'wake up'		'wake sb. up'		'wake sb. up for sb.'	
1 st	tu-nũ	'I wake up'	kitu-nũ-khi	'you wake me'	kitu-nũ-pi-khi	'you wake him for me'
2 nd	ki-nũ	'you wake up'	tatu-nũ-k'i	'I wake you'	tatu-nũ-pi-k'i	'I wake him for you'
3 rd	mu-nũ	'he wakes up'	ntu-nũ-Ø	'he wakes him'	ntu-nũ-pi-Ø	'he wakes him for him'

Table 34. Three related verbs in Tlahuica with increased valence.

The Atzincan applicative {-pi} comes from a pronominal marker of benefactive of the 3rd person which is still found in the other groups. In Otomi-Mazahua for example, transitive verbs become ditransitive by predictable stem changes; some are shown in Table 35. Hernández-Green (2015) proposes that in ditransitive constructions, recipient and benefactive arguments are treated by the syntax as primary objects in the conservative

language of Acazolco Otomi, while in innovative languages like Northern Otomi, Palancar (2009) claims that the alignment is indirective.

	‘give.as.gift’		‘sell’		‘frighten’	
1OBJ	hwě ^y -gi	give.as.gift-1OBJ	pa-gi	sell.tr-1OBJ	^h pih-ki	frighten.tr-1OBJ
2OBJ	hwě ^y -k’i	give.as.gift-2OBJ	pa-k’i	sell.tr-2OBJ	^h pi’-k’i	frighten.tr-2OBJ
3OBJ	hwěh+ki-Ø	give.as.gift[3OBJ]	pa-Ø	sell.tr[3OBJ]	^h pi+di-Ø	frighten.tr[3OBJ]
1DAT	hwě ^y -gi	give.as.gift-1DAT	pah-ki	sell.dtr-1DAT	^h pih+ti-ki	frighten.dtr-1DAT
2DAT	hwě ^y -k’i	give.as.gift-2DAT	pa’-k’i	sell.dtr-2DAT	^h pih+ti-k’i	frighten.dtr-2DAT
3DAT	hwěh+ki-bi	give.as.gift-3DAT	pah-pi	sell.dtr-3DAT	^h pih+ti-bi	frighten.dtr-3DAT

Table 35. Object and dative marking of three verbs in Tilapa Otomi.

Oto-Pamean also has semantic alignment. In Pamean, it appears restricted to stative predication, but in Otomian it is fully fledged. The most sophisticated system is observed in Mazahua (Vargas-Bernal 2012) and in Northern Otomi (San Ildefonso Otomi) with three classes of intransitive verbs: (i) a patientive class including inchoative verbs and all passive verbs, like in (36) from the variety of San Ildefonso Tultepec (other verbs include *no^hki* ‘get fat’, *pat’i* ‘get warm’, *thengi* ‘redden, blush’, *tset’i* ‘get cold’, *tits’i* ‘get well’, etc.); (ii) a patientive class including stative verbs (37) (e.g. *k’a* ‘be wet’, *kuhi* ‘be tasty’, *ʔi* ‘be spicy’, *ʔu* ‘be painful’, etc.); and (iii) an active class including all other intransitive verbs (38). Inactive predicates receive for subject the same person indexes that are used to index the object with transitive verbs.

(36) a. hi=mí t’axk-’-a=’i
 N.OTONEG=IMPF get.pale-2S_O-B=2S_O
 ‘You weren’t pale.’

b. bi=ts<’>ix-k-a=gi
 CPL=<PASS>take.AS-1S_O-B=1S_O
 ‘I was taken.’

(37) hi=már n-t’axi=’i
 N.OTO NEG=IMPF.ST ST-be.white=2S_O
 ‘You weren’t pale-skinned.’

(38) hi=ngí tag-i
 N.OTO NEG=2.IMPF fall-F
 ‘You weren’t falling.’

6. Final remarks

Being geographically isolated in Central Mexico, the Oto-Pamean languages are notably different in spirit from the rest of the Oto-Manguen languages, which are mostly spoken in the southern states of Oaxaca and Guerrero. They have completely renovated the old verbal inflectional system by making it fully periphrastic. This system is at a cross roads in Otomian, while it has become synthetic again in Pamean. The nominal class system is among the most developed one in the phylum and its correlations with the marking of number and possession are unique. The Oto-Pamean languages are also languages with a plural and a dual number in both verbs and nouns (except for Otomi-Mazahua where nominal dual has been lost – for more details see Palancar 2013a). Such a number system is

not common in Mesoamerica. Number agreement on verbs is also used in interesting ways not observed elsewhere. For example, in Otomian one finds a rare conjunctive strategy to express comitative semantics that is based on having a split subject by way of two NPs occupying two different positions in the clause with a verb that agrees in number with them as a totality.¹¹ This is illustrated by Eastern Otomi in (39), from the variety of Santa Ana Hueytlalpan, where the topical NPs occur thematized to the front of the clause (Palancar 2012: 267).

- (39) [ma mbe] nuya bá=n-tsi-hme=**wi** [a mbahä]
 ANA 1POSS mother today 3.CPL.TRANSLOC=MID-ingest-tortilla=DU DEF.SG priest
 ‘My mother had lunch today with the priest.’
 (Lit. ‘My mother today they two had lunch the priest.’)

But despite their idiosyncrasies, the Oto-Pamean languages have some other properties also shared by other Oto-Manguen languages. This can be seen in the fact that they are tonal, that they have intricate inflection and unproductive derivational morphology, and in that they have a tendency towards clausal juxtaposition as a syntactic device for clause combining and linking.

I want to finish this chapter with some prospects. Considerable progress has been made in bringing the Oto-Pamean languages to light since the pioneer approach by Soustelle (1937). But many areas remain in the shade escaping our understanding. Thorough studies on tone are much desired, but of those whose goal is to go beyond the lexical phonology which try to give answers to how the morphology works. For this, one needs to get one's hand in the dirt, as it were, for such studies would require an engagement in the inflection or the derivation that is challenging as it would go beyond a list of lexical items repeated three times. In this respect, as in others, we still have a very poor understanding of how tone inversion works in the different languages of the branch, and that, it seems to me, can be a very interesting area of study for tonal morphophonology, regardless of the theoretical approach.

Verbal and nominal inflection needs to be studied with rigor and methodology in Pamean. For Pame, the studies have focused on the phonology, but little on the morphology. As for Chichimec, the readily available data are becoming dated. In this respect, we need to know what has changed and how. Considering how complex the social history of the XX c. has been in Mexico, we expect much to have changed since the 1930's. This could shed light onto how complex morphological systems change, but also where they may remain unaltered. Noun classes and possession classes are in want of high-quality lexical studies based on samples of hundreds of entries. Such studies could also shed light onto how the alienable/inalienable distinction is handled by the inflection in these languages. We know practically nothing of the prosodic and syntactic characteristics of the inflectional formatives of Atzinca: are they prefixes like in Pamean or are they clitic words like in Otomi-Mazahua? etc.

But the area by far where research is most needed is syntax without any doubt. For example, we know nothing of the syntax of Pamean languages. This is unfortunate, but the only knowledge we have comes from the glimpses we can get from the old textual materials in Angulo and from its complex inflectional morphology. But if our knowledge of the syntax of clause structure, relative clauses, clause linkage, and subordination is still poor

¹¹ As a way to get to grips with the intricate syntax of the construction, some authors like Ecker (1952) and Hess (1968) concluded that the marker of dual in an example like (34) works as a comitative applicative, when in reality it is just a morpho-syntactic token to indicate that the verb has a dual subject.

outside Otomi-Mazahua, total darkness falls on us when it comes to the syntax and pragmatics of topic structures or focus structures such as clefts.

All these lagoons are understandable, because the phonology and the morphology of the Oto-Pamean languages is not an easy bone to chew on. But as our knowledge of these areas has been increasing through the years, any efforts oriented towards giving answers to some of these questions would not only be of help to understand Oto-Pamean better, but would be directed towards gaining better knowledge of Mesoamerican languages in general.

References

- Angulo, Jaime de. 1933. The Chichimeco language, *International Journal of American Linguistics* 7: 152-194.
- Avelino, Heriberto. 1997. Fonología y morfofonología del Pame Norte. BA thesis, Escuela Nacional de Antropología e Historia.
- Bartholomew, Doris. 1965. A reconstruction of Otopamean (Mexico). PhD dissertation, University of Chicago.
- nd. Observations about Matlatzinca Grammar: Based on informant work done in May, 1966. Language Associate: Ezequiel Hernández. MS
- Berthiaume, Scott. 2003. A phonological grammar of Pame. Ph.D. dissertation, University of Texas at Arlington.
- Campbell, Lyle, Terrence Kaufman, and Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62: 530–70.
- Cárceles, Fray Pedro de. 1907 [1580]. Arte de la lengua otomí. In Nicolás León (ed.), *Boletín del Instituto Bibliográfico Mexicano* 6: 38-155. Mexico City: Instituto Bibliográfico Mexicano.
- Carranza-Martínez, Leonardo. 2013. Flexión verbal del matlatzinca: Marcación de persona-número y de tiempo-aspecto-modo. MA thesis. Universidad Autónoma Metropolitana.
- CLIN: Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. 2008. *Diario de la Nación* (14-01-2008), México D.F.
- Echegoyen, Artemisa and Katherine Voigtlander (E&V). 2007. *Diccionario yuhú: Otomí de la Sierra Madre Oriental (Estados de Hidalgo, Puebla y Veracruz, México)* Mexico City: Instituto Lingüístico de Verano.
- Ecker, Lawrence. 1952. Compendio de gramática otomí: Introducción a un diccionario otomí-español. *Anales del Instituto Nacional de Antropología e Historia* IV-32: 121-174.
- Finkel, Raphael and Gregory T. Stump. 2007. Principal Parts and Morphological Typology. *Morphology* 17-1: 39-75.
- Gibson, Lorna and Doris Bartholomew. 1979. Pame noun inflection. *International Journal of American Linguistics* 45(4): 309-322.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's fieldtrip. *The Proceedings of the Seventh Regional Meeting of the Chicago Linguistic Society*: 394-415.
- Golston, Chris and Wolfgang Kehrein. 2004. A prosodic theory of laryngeal contrasts, *Phonology* 21, 1-33.
- Hernández Cruz, Luis, Moisés Victoria Torquemada and Donald Sinclair Crawford (eds). *Diccionario del Hñähñu (Otomí) del Valle de Mezquital, Estado de Hidalgo*. [Vocabularios Indígenas, 45]. México D.F.: Instituto Lingüístico de Verano.
- Hernández-Gómez, Selene. In preparation. Otomí de Santa Ana Hueytlalpan: Descripción gramatical y documentación lingüística. PhD dissertation MS, Universidad Nacional Autónoma de México.

- Hernández-Green, Néstor. 2012. Construcciones aplicativas de registro en otomí de Acapulco. Paper read at the Coloquio de lenguas otomangués y sus vecinas, COLOV 5. Universidad Benito Juárez de Oaxaca, Oaxaca, 20-22 April 2012.
- 2009. Semántica de los formativos temáticos en otomí: Un estudio de corpus en la variante del Valle del Mezquital. Unpublished MA thesis, Universidad Autónoma de Querétaro.
- Forthcoming. Inflectional verb classes in Acapulco Otomi. In Enrique L. Palancar, Matthew Baerman and Timothy Feist (eds.), *Inflectional Complexity and Verb Classes: A view from the Oto-Manguéan languages of Mexico*. Language Science Press.
- Herrera-Zendejas, Esther. 2009. *Formas sonoras: mapa fónico de las lenguas mexicanas*. México D.F.: El Colegio de México.
- Hess, H. Harwood. 1968. *The Syntactic Structure of Mezquital Otomi*. [Janua Linguarum. Series Practica 43]. The Hague: Mouton.
- Knapp Ring, Michael H. 2002. Elementos de dialectología mazahua. In Paulette Levy (ed.), *Del cora al maya yucateco. Estudios lingüísticos sobre algunas lenguas indígenas mexicanas*, 59-80. México D.F.: Instituto Nacional de Antropología e Historia & UNAM.
- Lastra, Yolanda. 1984. Chichimeco Jonaz, in Munro S. Edmonson S. (ed.), *Supplement to the Handbook of Middle American Languages: Linguistics*, 20-42. Austin: University of Texas Press.
- Martínez-Ortega, Aileen. 2012. Clases verbales, transitividad y valencia verbal en el p'jyekakj'ó, tlahuica de San Juan Atzingo. Unpublished PhD dissertation, El Colegio de México.
- 2010. Variaciones en los sufijos verbales de plural y de aplicativo en el tlahuica (*p'jyí kakj'ó*). Paper read at the Coloquio de Lingüística de la ENAH.
- Palacios, Patricia. 2011. Mecanismos de cambio de valencia en el otomí de la Sierra. Unpublished MA thesis. Universidad Autónoma de Querétaro.
- Palacios, Patricia and Enrique L. Palancar. 2010. *Mecanismos de cambio de valencia en el otomí de la Sierra*. Paper read at the X Encuentro Internacional de Lingüística en el Noroeste, Sonora: Universidad de Sonora. 17-19 Nov.
- Palancar, Enrique L. 2004. Transitividad en el otomí de San Ildefonso: una incursión tipológica en lenguas otomangués. In Isabel Barreras and Mirna Castro (eds.), *Memoria del VII Encuentro Internacional de Lingüística en el Noroeste* 7(I), 173-196. Sonora: UniSon
- 2006. Middle voice and intransitivity in Otomi. *Linguistics* 43(1): 613-43.
- 2008. *Gramática y textos del Hñõñhõ: Otomí de San Ildefonso Tultepec, Querétaro. Volumen I: Gramática*. México D.F.: Plaza y Valdés.
- 2012. A typology of split conjunction. *Linguistic Typology* 16: 265-320.
- 2013a. The evolution of number in Otomi: The many faces of the dual. *Studies in Language* 37(1): 94-142.
- 2013b. Preaspiration in Northern Otomi. In Samia Naim and Jean-Léo Léonard (eds.), *Backing and Backness*, pp. 205-220. LINCOM Europa.
- Forthcoming. On the tightrope between infinitives and action nouns: The case of Otomi nominalizations. In Claudine Chamoreau and Zarina Estrada (eds.), *Finiteness and Nominalization*. Amsterdam: John Benjamins.
- Palancar, Enrique L. and Heriberto Avelino. Forthcoming. Inflectional complexity and verb classes in Chichimec. In Enrique L. Palancar, Matthew Baerman and Timothy Feist (eds.), *Inflectional Complexity and Verb Classes: A view from the Oto-Manguéan languages of Mexico*. Language Science Press.

- Palancar, Enrique L. and Leonardo Carranza. Forthcoming. Inflectional class interaction in Matlatzinca. In Enrique L. Palancar, Matthew Baerman and Timothy Feist (eds.), *Inflectional Complexity and Verb Classes: A view from the Oto-Manguean languages of Mexico*. Language Science Press.
- Soustelle, Jacques. 1937. *La famille Otomi-Pame du Mexique Central*. Paris: Institut d'ethnologie.
- Stewart, Donald. 1966. *Gramática del mazahua*. Instituto Lingüístico de Verano, México. MS.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics. Volume II: Typology and process in concept structuring*. Cambridge, MA: MIT press.
- Vargas-Bernal, Gloria. 2012. *Alineamiento e intransitividad escindida en mazahua*. Unpublished MA thesis. CIESAS, México.
- Voigtlander, Katherine and Artemisa Echegoyen (V&E). 1985. *Luces contemporáneas del otomí: Gramática del otomí de la Sierra*. México D.F.: Instituto Lingüístico de Verano.
- Voigtlander, Katherine, Artemisa Echegoyen, and Doris Bartholomew. 2006. La semántica de los sufijos temáticos en yuhu (otomí de la Sierra). *Estudios de Cultura Otopame* 5: 279-302.

Acknowledgements:

This research was partially supported by the ESRC/AHRC project RES-062-23-3126 “Endangered complexity: Inflectional classes in Oto-Manguean languages” and by the PICS 2015-2017 “Méso-Amérique et la syntaxe de la proposition relative”. I wish to thank Søren Wichmann for inviting me to participate in this handbook and for his comments to the chapter as editor. A version of this chapter was presented at the conference titled “State of the art in Mesoamerican Linguistics” celebrated at the Max Planck Institute for Evolutionary Anthropology in Leipzig in December 2014. I want to thank the conference participants for valuable comments received after the talk during breaks and especially to Roberto Zavala and Claudine Chamoreau for useful exchanges about the Oto-Pamean languages. I am very grateful to Timothy Feist for his insightful comments and for proof-reading my English. All errors and deficiencies remain my responsibility.

Abbreviations

ABS	absolute	INFL	inflectional morpheme
AS	adjusted stem	INV	invariable
ADLAT	adlative	IRR	irrealis
ADV	adverbial	MID	middle
AMBUL	ambulative	MOT	directed motion
ANDAT	andative	NEG	negation
ANT	anterior	NMLZ	nominalization
B	bound	OBJ	object
CISLOC	cislocative	PRTCL	particle
CL	class	PASS	passive
CPL	completive	PFCL	phrase final clitic
COP	copula	PST	past
DAT	dative	PRF	perfect
DEF	definite	PPRF	plu-perfect
DEM	demonstrative	PL	plural
DU	dual	POSS	possessive
EMPH	emphatic	POT	potential
EXCL	exclusive	PRS	present
F	free	PROSP	prospective
FUT	future	PURP	purposive
GEN	general adverbial	QUOT	quotative
HAB	habitual	R	realis
HES	hesitation	REC	recent
HUM.ANTIP	human antipassive	S _o	inactive subject
IMM	immediate	SS	secondary stem
IMPF	imperfect	SEQ	sequential
INCL	inclusive	SG	singular
INCPL	incompletive	ST	stative
INDF	indefinite	TRANSLOC	translocative
INF	infinitive		

Languages

ACA	Acazulco Otomi
ANA	Santa Ana Otomi
C.PAM	Central Pame
CHI	Chichimeco-Jonaz
E.OTO	Eastern Otomi
MAZ	Mazahua
N.PAM	Northern Pame
N.OTO	Northern Otomi
TIL	Tilapa Otomi
TLA	Tlahuica

Enrique L. Palancar

6 A typology of tone and inflection: A view from the Oto-Manguean languages of Mexico

1 Introduction

Tone is well-known for playing a role in the makeup of inflection, but the literature has not yet produced a comprehensive typology of the relation between tone and inflection from a morphological perspective. There are various reasons for this. On the one hand, the interest in the morphological side (i.e. the form) of inflection has lain outside the general spotlight of linguistic typology, which, except for the Surrey Morphology Group, has been mainly concerned with trying to make sense of linguistic structure in functional terms. I show here that in most situations where tone is involved in inflection it is difficult, if not impossible, to give it a clear-cut morphosyntactic function.

On the other hand, our understanding of how tone works is relatively recent and the emphasis of its study has been on its phonological nature and on how it contributes to the creation and maintenance of lexical meaning. This understanding has been made possible thanks to the substantial body of knowledge generated through the laborious work of many field linguists describing tonal languages with complex inflection. In this respect, the contribution to the field has mainly come from the languages of Africa. We know less of languages from other areas, especially of those from the Americas because it has taken the field more time to gather the basic relevant information. This is somewhat ironic because Pike's 1948 seminal work on tone was precisely inspired by the Oto-Manguean languages; a large phylum of tonal languages spoken in Mexico.

In this light, I outline here a typology of the ways we find tone involved in the realization of inflectional categories. For this, I take a morphological perspective on tone. This also means that I have consciously avoided using jargon from the discipline of phonology since its technicalities often hinder the understanding of a broader readership. This also means that the emphasis is not on what a given tone sounds like or how it stands, works or should be characterized with respect to other suprasegmental units in the system, but rather on what tone is used for in the formal makeup of inflection and on how it is used for that purpose.

In this respect, because inflection is the part of syntax that is relevant to the morphology, the emphasis will be on the mapping between linguistic form and grammatical meaning. To understand the role of tone in this context, we could

imagine a continuum of two ideal systems. At the one end, tones have no role in inflection. They are regarded as a lexical property of an inflected stem or of a grammatical affix. At this end, tones are bits of the lexical phonology. This type of tone is presented in §2. At the other end, tones are treated as discrete supra-segmental affixes (commonly called ‘auto-segments’) working as genuine exponents of morphosyntactic feature values. Such type of inflectional tone is briefly presented in §3. Languages vary considerably along this logical continuum and the present study is an attempt to pinpoint the most relevant types that can be observed focusing on morphological tone which is a type of inflectional tone used as a building block of the grammar of a language for the construction of forms in inflectional paradigms. In this respect, in §4 I introduce the four main subtypes of morphological tone. Besides morphological tone, we can also find tonal allomorphy organized in inflectional classes. This type is presented in §5.

To illustrate this typology I mainly focus on the Oto-Manguean languages of Mexico (although not exclusively). There are various reasons for this. Within Oto-Manguean, one finds the most complex morphological systems we know of (only comparable to some found among the Nilotic languages of Africa, notably Dinka and Nuer) and a great part of that complexity is due to tone. Also their internal diversity displays a great wealth of possible types. I also concentrate on verbal inflection, not only because verbs are by far the word class in these languages where the morphology is more complex, but also because we know less about the nominal inflection, whose study would no doubt have some bearing on the present proposal.

1.1 The Oto-Manguean languages of Mexico

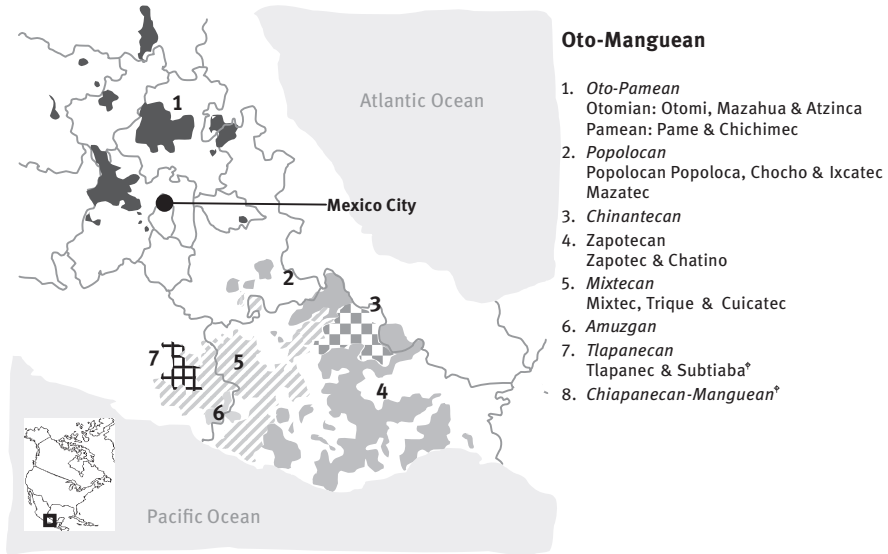
Oto-Manguean (also ‘Otomanguean’) is a linguistic phylum of Amerindian languages spoken in Central and Southern Mexico.¹ Oto-Manguean is typologically unique in the Americas because it is the only phylum where all its languages are tonal.

The 2005 census by the INALI² states there are 1,769,971 speakers of Oto-Manguean languages, of which 218,679 are monolingual. Oto-Manguean

¹ The three Oto-Manguean languages once spoken outside the current territory of Mexico are Tlapanecan Subtiaba† (Nicaragua) and Manguean Monimbot† (Nicaragua) and Chorotega† (Costa Rica).

² National Institute for Indigenous Languages.

consists of the eight different linguistic groups in Map 1, which for the present purposes I treat as branches of the phylum.³



Map 1: The Oto-Manguean languages of Mexico

While there is agreement among scholars as to which languages are Oto-Manguean and which are not, there is currently no established consensus regarding the number of languages involved. This is because Oto-Manguean generally involves important diasystems or dialectal continua within each of its different branches, making it difficult to define categorically where a given dialect finishes and a different language starts. Figures vary from source to source: while the Ethnologue and the Glottolog propose 176 and 179 different languages, respectively, the INALI proposes 220 giving more weight to sociolinguistic factors (in much the same

³ The traditional taxonomy found in Campbell (1997), which is in turn based on unpublished materials by Terence Kaufman, proposes four major branches organized into two large subgroups: (i) Western Oto-Manguean consisting of Oto-Pamean-Chinantecan and Tlapanecan-Manguean; and (ii) Eastern Oto-Manguean formed of Amuzgo-Mixtecan and Popolocan-Zapotecan. I prefer to take here a flat approach to the phylum since this taxonomy has not yet been discussed at length. Coincidentally, Cecil H. Brown and Eric Holman supplied to me in a personal communication a MDS plot using data available at the Automated Similarity Judgment Program database (<http://asjp.cld.org/>). This plot shows that the superordinate groups of the traditional taxonomy lack a centered core and Western Oto-Manguean has no integration.

vein as Danish and Norwegian Bokmål or Serbian and Croatian are treated as different languages).

All Oto-Manguean languages are tonal. But the tonal systems within the phylum vary enormously; from languages with an inventory of two tones (e.g. Chichimec and Pame from the Oto-Pamean branch) to languages with an inventory of ten or more tones (e.g. the Eastern Chatino languages of the Zapotecan branch). Apart from tone, the internal diversity of the phylum makes it difficult to come up with an inventory of features that is shared by all languages. There is an overwhelming tendency towards open syllables, especially at root level. Languages often have nasal vowels, complex phonation and voice quality. All languages are head-marking. Nouns do not inflect for case, although they may receive classifiers and their possession inflection can be at times challenging. In some branches, the animacy value or the discourse status of a noun is also relevant for the inflection. Oto-Manguean languages lack non-finite forms such as infinitives, and this has consequences for the syntax of clause chaining. In general, inflectional morphology is very intricate. This is partly due to the interaction of multiple systems acting independently of each other, resulting in cross-classification: e.g. a verb may simultaneously belong to various inflectional classes: one for its endings, another for its stem changes and a third for its tonal changes. This situation produces in turn the great deal of irregularity in verbal paradigms that is characteristic of some branches (Levy 1987). At the level of word order syntax, all branches can be characterized as verb initial (also a Mesoamerican feature according to Campbell *et al.* 1986); only recently some have become predominantly SV. Oto-Manguean languages are largely fusional; they do not have noun incorporation except for the basic type. Their derivational morphology is largely unproductive. New words are often borrowed and often lie outside the scope of the more morphological aspects of the inflection.

2 Tone as a phonological property of grammatical markers

Inflectional information is often, although not exclusively, conveyed by affixes. In many tone languages with relatively complex inflectional morphology, tone may be simply a property of the lexical phonology of the inflectional affixes, just as it is associated with the phonology of other words with lexical content.

For example, Otomi languages (Oto-Pamean) have three tones, but only two are relevant for our purposes: high (represented with an acute accent) and low (not represented). A contrast based on the high vs. the low tone operates as a

useful mechanism to keep grammatical markers phonologically distinctive, which would otherwise be segmentally homophonous. This is illustrated in (1) from the variety of Northern Otomi spoken in San Ildefonso Tultepec in the state of Querétaro. Here the tonal contrast of high vs. low that serves to create minimal pairs for lexical words (e.g. 'b^óts'e (H) 'religious offering' vs. 'b_óts'e (L) 'basket, ribcage'; d^ó'yo (H) 'bone' vs. d_ó'yo (L) 'hotplate to heat tortillas', etc.) is exploited in inflectional clitics *dá* vs. *da* or *gí* vs. *gi*.⁴

- | | | |
|-----|-----------------------------------|--------------------------------|
| (1) | Tone /H/ | Tone /L/ |
| | a. dá =ts _ó t'e | da =z _ó t'e |
| | 1.CPL=arrive.there | 3.IRR=SS/arrive.there |
| | 'I arrived there.' | 'S/he'll arrive there.' |
| | b. gí =ts _ó t'e | gi =ts _ó t'e |
| | 2.INCPL=arrive.there | 2.IRR=arrive.there |
| | 'You arrive there.' | 'You'll arrive there.' |

The Otomi case is commonly found across languages, and it is the most common source for inflectional tonogenesis. In this respect, Palancar *et al.* (this volume) and Feist and Palancar (this volume) show cases in Mixtecan languages where a given grammatical prefix consisting of segments and tone loses its segmental part and all that remains of it is its tone, which then becomes the exponent of the grammatical information. Such tones need to be linked to a segment in the stem, and often invade its lexical phonology giving rise over time to complex tonal patterns (such as for example the tonal classes in §5).⁵

3 Morphosyntactic tone

Tones can do things other than sit inert in the phonology. In Otomi languages, a third person possessor is encoded by a floating high tone. This happens without exception every time the information is required by the syntax. The high tone

⁴ The third tone is a raising tone that is not found in inflectional formatives, but only found in lexical roots. This contour tone emerged from the monosyllabification of historical disyllabic stems bearing the melody /L-H/. This is a well-known source of lexical tonogenesis.

⁵ Inflectional tonogenesis can also arise in unpredictable ways, see for example Fedden (this volume) for a very interesting case study of how a particular tone value in Mian, a Papuan language with a rich tonal system, serves an inflectional function in a very restricted context.

lands on the mora immediately preceding the possessed noun.⁶ The landing site may be a determiner as in (2a) or a preposition as in (2b) (in the examples, the singular enclitic =*r* does not bear tone).⁷ The Otomi case also serves as an illustration that within the same inflectional system one may also find different situations at the same time.

(2)	Possessed		Non-possessed
a.	nó=r	bötsi	no=r
	<3.POSS>DEF.SG=SG	child	DEF.SG=SG
	‘His/her/their child.’		‘The child.’
b.	ká=r	méxá	ka=r
	<3.POSS>on=SG	table	on=SG
	‘On his/her/their table.’		‘On the table.’

The tone inventory of Otomi languages used for situations like this is restricted to a binary contrast of high vs. low. A more interesting case is found in Yoloxóchitl Mixtec, a language with more tones.

Yoloxóchitl Mixtec (Mixtecan) has nine basic tones in lexical contrasts (cf. Vietnamese has six according to the common view): four level tones (written with superscript numbers from /1/ for the low tone through /4/ for the high tone); three rising tones (13, 14, 24) and two falling tones (32, 42) (DiCano *et al.* 2012). Out of these nine tones, at least three are fully exploited in the inflection. Table 1 illustrates the forms of different verbs for the irrealis, the negative irrealis and the completive. For this language, the tones in the irrealis can be taken as basic; that is, as provided by the lexicon (see Palancar *et al.*, this volume).

In Table 1, we have verbs with different melodies in the base form: /3-2/, /3-3/, /3-3-3/, /3-3-4/ and /3-3-4-4/. If this form is compared with the tonal melodies of the other two forms, we observe there is a tone change affecting the first mora, i.e. instead of /3/, we find /14/ in the negative irrealis and /13/ in the completive. The mapping of form (i.e., a specific tone value) and grammatical meaning (i.e., a specific inflectional value) is constant and there are no exceptions. This remarkable systematicity leaves us with an inflectional system operating with flawless morphosyntactic rules that involve tone. In such a system, tone could be said to

⁶ If the mora on which the floating tone lands has high tone (for independent reasons it is not possible that it has rising tone), the word is pronounced with a rising tone that starts very high.

⁷ Abbreviations: ANT anterior; AOR aorist; CPL completive; DEF definite; FIN final vowel; FUT future; HAB habitual; IMM immediate; IMPER imperative; IMPF imperfect; INCL incomplete; INF infinitive; IRR irrealis; NEG negative; POSS possessive; POT potential; PROG progressive; PROSP prospective; PRS present; PST past; SUB subject; REC recent; SEQ sequential; SS secondary stem.

work as a regular suprasegmental affix: the affix being suppletive for the negative irrealis (i.e. /14/ replaces /3/ on the first mora of the stem), or additive for the completive (i.e. /1/ is inserted before the lexical /3/ on the first mora of the stem producing a raising tone /13/).

Table 1: Tone and verbal inflection in Yoloxóchitl Mixtec

IRR	NEG. IRR	CPL		
nu ³ mi ³	nu ¹⁴ mi ³	nu ¹³ mi ³	‘hug’	(tr)
ku ³ u ³	ku ¹⁴ u ³	ku ¹³ u ³	‘happen’	(intr)
nda ³ kin ²	nda ¹⁴ kin ²	nda ¹³ kin ²	‘sharpen’	(tr)
ku ³ ya ³ tin ³	ku ¹⁴ ya ³ tin ³	ku ¹³ ya ³ tin ³	‘get closer’	(intr)
ndu ³ toʔ ³ ni ³	ndu ¹⁴ toʔ ³ ni ³	ndu ¹³ toʔ ³ ni ³	‘open eyes’	(intr)
nda ³ xiʔ ³ i ⁴	nda ¹⁴ xiʔ ³ i ⁴	nda ¹³ xiʔ ³ i ⁴	‘open’	(tr)
cha ³ ka ³ nduʔ ⁴ u ⁴	cha ¹⁴ ka ³ nduʔ ⁴ u ⁴	cha ¹³ ka ³ nduʔ ⁴ u ⁴	‘lay down’	(tr)

This regularity makes Yoloxóchitl Mixtec an ideal language for syntactic approaches to inflection that do not particularly appreciate the ‘messy bits’ of the morphology. But those bits exist, nonetheless; they sometimes mess about the mapping of form and meaning in the inflectional realization or they serve as mere building blocks in the formation of grammatical structure. Tone plays a fundamental part in those messy bits. In the following section, I outline the main types of purely morphological tone.

4 Morphological tone: Tone linked to stems as a building block of paradigms

To introduce morphological tone, let us first depart from a situation where tones are just an inherent property of the phonology of the stem, just as they may be of affixes (as in the Otomi case presented in §2). In such a situation, stems surface with tones associated with them by the lexicon, just as they have specific vowels or consonants.⁸ In an inflectional system involving verbs, tone may sit inert on

⁸ It is more convenient here to talk about stems rather than roots because the stem is the fundamental formal unit that represents the lexeme in an inflected form. Also, while some stems consist of a root only, some consist of other material besides the root.

verbal stems. This may happen to all verbs, to only some or to just a handful of them. A particularly interesting case to start with is Tswana, a Bantu language spoken mainly in South Africa.

Tswana has two tones: high and low. The tones are lexically associated both with stems and with affixes. Stems can have inherent high tone or be toneless. If toneless, they are subject to complex prosodic rules at the word level (Creissels *et al.* 1997). Take for example the verb *t^halovχaŋ-* ‘understand’ which is a toneless stem that surfaces as /HLL/ in (3) as a result of a regular spreading rule by means of which the high tone of the preceding potential prefix *ka^H-* spreads on to the first syllable of the stem; the other two syllables receive low tone by a prosodic default.

- (3) *ba^H-ka^H-t^ha^Hl^Lovχa^Lŋ-a^L*
 3.SUB-POT-understand-FIN
 ‘They can understand.’

The issue at hand here is what Creissels (2006:10) treats as ‘grammatical high tone’. In contrast to (3), in (4) the same toneless stem now surfaces as /LHH/ in the negative of the same mood. Here the high tones in the stem cannot be explained as the result of any prosodic rule, i.e. we might have expected the structure to have been /LLL/.

- (4) *ba^H-ka^H-s^Ll^L-t^ha^Ll^Lovχa^Hŋ-a^L*
 3.SUB-POT-NEG-understand-FIN
 ‘They cannot understand.’

Creissels concludes that the distribution of the high tone in (4) is purely morphological in nature (in the sense of Aronoff 1994), arguing that it does not respond to any semantically coherent bundle of morphosyntactic features values.

Verbal inflection in Tswana is rich. A verb can inflect for a large number of morphosyntactic values, including mood (indicative, circumstantial, subjunctive, imperative, potential, etc.), tense-aspect (present, future, perfect, etc.), polarity (positive vs. negative), disjunction (conjunct vs. disjunct), as well as non-finite forms such as the infinitive, the relative and the sequential. In this context, the grammatical high tone of Tswana is found in all negative forms, so in principle, one could take it to be an exponent of negation alongside other segmental ones (namely the prefix *sl^L-*), but it is also found in other forms that have little to do with negation, such as the positive indicative perfect conjunct, the positive circumstantial perfect, the imperative and the sequential. This heterogeneity of meanings indicates a lack of morphosyntactic function, but it does not mean it lacks a grammatical function altogether.

In our view, the grammatical tone of Tswana serves as a building block in inflected forms in paradigms. It works, as it were, like thematic vowels in languages like Latin or like stress patterns in Greek or Russian. We treat this type of tone as ‘morphological tone’.

A very similar morphological tone system, but with a more complex distribution that requires an intricate morphophonological apparatus is found in Kabiye, a Gur language of Togo that belongs to the North-Volta Congo subgroup of the Volta-Congo branch of Atlantic-Congo.

4.1 Morphological tone conditioned by phonological shape

Like Tswana, Kabiye also has two tones: high (marked by an acute accent) and low (not represented). Verbs in Kabiye have a rich array of inflectional possibilities, but to illustrate the point we will only concentrate on six basic forms built by suffixation (including zero suffixes, i.e. bare stems): imperfective present, imperfective past, past, aorist, infinitive and imperative (with a bare stem). There are verbs with inert tone across the paradigm, i.e. stem tone does not change, like for example *kpéyu* ‘forgive’ in Table 2.

Table 2: The Kabiye verb *kpéyu* ‘forgive’ has a stem with invariable tone

	‘forgive’	tone
IMPF.PRS	-kpéy-u	H
IMPF.PST	-kpéy-aɣ	H
PST	-kpéy-a	H
AOR	-kpéy	H
INF	kpéy-u	H
IMPER	kpéy	H

But not all verbs in Kabiye are like *kpéyu*. For our analysis, we rely on the sample of 800 verbs in Roberts (2013). Out of them, only 165 behave like this; the rest have tonal alternations involving stems that receive a low tone in certain cells of the paradigm. What is interesting is that not all such verbs alternate in the same way.

All verbs considered, including the ones that do not alternate, display five patterns of tonal behavior. Table 3 shows two such patterns. Shading highlights

stems that receive a low tone. Besides tone, stems may be also display other segmental changes that are independent of tone alternations.

Table 3: Two verbs in Kabiye showing different patterns of tone alternations

	'sweep'	tone	segm.	'cut'	tone	segm.
IMPF.PRS	-há-z-ɪʔ	H	haz	-cék-ɪ	H	cɛk
IMPF.PST	-haz-aʔ	L	haz	-cɛk-aʔ	L	cɛk
PST	-haz-aá	L	haz	-cɛb-á	L	cɛb
AOR	-há-z-ɪ	H	haz	-cék	H	cɛ
INF	haz-óu	L	haz	cék-u	H	cɛb
IMPER	haz-ɪ	L	haz	cɛ	L	cɛ

Here we are interested in two important questions: (i) regarding the form: for the inflection of a given verb, can we predict when the stem will have a high or a low tone? and if so, how can we predict it? and (ii) regarding the function: what is the relation of tone to grammatical function in such a system? Let us address the first question.

The study of inflectional morphology is a quest to find grammatical structure and analyses are just possible paths to it. The standard procedure with tone is to take a given form as base, whose tone is considered lexical (i.e. stable and memorized), from which other forms are derived through morphophonological rules. This means that tone allomorphy may be predicted if one finds the key to the system that can make generalizations. Kabiye is a good example of this.

Roberts (2013) divides Kabiye verbs into 23 different conjugation models attending to other aspects of inflection, but not tone.⁹ The exemplar verbs for each one of Roberts' conjugations are given in Table 4 together with an indication of their population size in his sample. To facilitate understanding, the forms in the table have been arranged in a slightly different way to the way they appear in tables 2 and 3. Shading applies to cells with low tone stems. The different patterns of tonal behavior are indicated by roman numerals.

⁹ Roberts (2013) includes four other paradigms for another ten verbs with motion components and complex stems which we exclude here.

Table 4: Distribution of high vs. low tone in stems in Kabiye verbal paradigms

	IMPF.PRS	AOR	IMPER	INF	IMPF.PST	PST		#	Conj	
I	a.	-há-z-ɪx	-há-z-ɪ	haz-ɪ	haz-úu	-haz-aʔ	-haz-aá	‘sweep’	92	9
	b.	-súl-uu	-súl-ɪ	sul-ɪ	sul-úu	-sul-aʔ	-sul-aá	‘cover up’	52	10
	c.	-kélés-ɪx	-kélés-ɪ	keles-ɪ	keles-úu	-keles-aʔ	-keles-aá	‘rinse’	48	17
	d.	-lúlús-uu	-lúlús-ɪ	lulus-ɪ	lulus-úu	-lulus-aʔ	-lulus-aá	‘pour’	32	18
II	e.	-páx	-páa	paa	pá-ú	-pax-aʔ	-pa(w)-á	‘dance’	36	13
	f.	-yó-u	-yóo	yoo	yó-ú	-yow-aʔ	-yow-á	‘quarrel’	19	14
	g.	-dʒng	-dʒ	dʒ	dʒm	-dʒng-aʔ	-dʒm-á	‘walk’	31	16
	h.	-sók-ɪ	-súu	suu	súu-u	-sok-aʔ	-suw-aá	‘load’	9	12
	i.	-cÉk-ɪ	-cÉ	cɛ	cÉb-u	-cɛk-aʔ	-cɛb-á	‘cut’	28	15
	j.	-cák-ɪ	-cáx	cax	caʔ-u	-cak-aʔ	-cax-á	‘sit’	12	11
III	k.	-kpéx-u	-kpéx	kpéx	kpéx-u	-kpéx-aʔ	-kpéx-a	‘forgive’	2	21
	l.	-kát-ɪx	-kát-ɪ	kát-ɪ	kát-uu	-kát-aʔ	-kát-aa	‘meet’	111	19
	m.	-dʒk-uu	-dʒk-ɪ	dʒk-ɪ	dʒk-uu	-dʒk-aʔ	-dʒk-aa	‘hold’	52	20
IV	n.	-háx	-há	há	há-u	-háx-aʔ	-ha(w)-á	‘give’	17	5
	o.	-lú-u	-lú	lú	lú-u	-lúw-aʔ	-luw-á	‘take out’	17	6
V	p.	-tók-úu	-túu	túu	túu-u	-tuk-aʔ	-tu-áa	‘decrease’	5	4
	q.	-kpák-íʔ	-kpáʔ	kpáʔ	kpaʔ-u	-kpak-aʔ	-kpaʔ-a	‘take’	10	3
VI	r.	-cél-íʔ	-cél-í	cél-í	cɛl-úu	-cɛl-aʔ	-cɛl-áa	‘give back’	70	1
	s.	-kót-úu	-kót-í	kót-í	kot-úu	-kot-aʔ	-kot-áa	‘fold’	28	2
	t.	-kúúd-úu	-kúúd-í	kúúd-í	kuud-úu	-kuud-aʔ	-kuud-áa	‘twist’	7	8
	u.	-wélés-íʔ	-wélés-í	wélés-í	weles-úu	-weles-aʔ	-weles-áa	‘listen’	31	7
	v.	-cíkíl-ɪx	-cíkíl-ɪ	cíkíl-ɪ	cíkíl-uu	-cíkíl-aʔ	-cíkíl-aa	‘tickle’	66	22
	w.	-húlút-uu	-húlút-ɪ	húlút-ɪ	hulút-uu	-hulút-aʔ	-hulút-aa	‘shout (ritual)’	25	23

We observe that a high tone is always used for the imperfective present and the aorist, and this can be easily spelled out as a rule. The question is when we can predict low tone. For the building of the tonal structure of remaining forms, we propose that the stem of the imperative is informative, i.e. a principal part (Finkel and Stump 2007). In this we follow Roberts, but we reach different conclusions.

If a verb has a low tone in the stem used for the imperative form, then it will have low tone in all other forms (pattern I), except when the imperative stem is monosyllabic, in which case the infinitive will have high tone (pattern II). When the verb has high tone in the imperative, things become more complex as conditions get more specific. With many options, one commonly resorts to the existence of a default rule that applies elsewhere. The default for a verb with a high tone in the imperative is to also have high tone for all other forms, i.e. the invariable pattern III grouping verbs (k-m). In this, it behaves like the default of low tone. Exceptions can be explained by appealing to phonological properties of the stem, but the rules need to be stored somewhere under the control of the morphology. First, if the segmental stem is of the type /CVɣ-/ or /CVw-/ as in (n-o), then the past has low tone, despite all other forms having high tone (pattern IV). If the imperative has the shape /-(H-)H+H/ (+ indicates the mora of the suffix -i/i), then the shape for the infinitive, the imperfective past, and the past will be /-(L-)L+H/ making pattern VI as in (r-u). But if the shape is /H-H+L/, for those same forms, the shape will be /L-H+L/ as in (v-w). Finally, verbs in (p-q) have two segmental stems in alternation (CVV- vs. CVk-); for such verbs with a high tone imperative, the stem for the past tenses always carries low tone (pattern V).

This analysis could perhaps be rephrased in a more sophisticated way, but it illustrates a situation where the distribution of tonal changes in paradigms can be accounted for as attending to other aspects of the form without having to appeal to the lexicon.

Having established the rules behind the making of forms in the paradigm, we are left with the second question: what is the function of tone in Kabiye? Our view is that we are facing a similar type of morphological tone as seen in Tswana. In other words, tone in Kabiye does not have a morphosyntactic function. Neither low nor high tone is used as an inflectional exponent for any given inflectional category. The fact that high tone is always used for the stem of the imperfective present and the aorist could cause one to ponder whether there is a grammatical connection between high tone and such values. If we were to accept such a view, we would then be confronted with two further issues: (i) what is the grammatical connection between those two tenses which would enable us to make sense, in functional terms, of the fact that the same form (i.e., high tone) encodes both; but most significantly (ii) why do other stems have high tone for other cells of the paradigm?

While morphological tone in Kabiye and Tswana serves the same purpose, it works in typologically different ways in each language. In Tswana, the rules of high tone assignment apply to all verbs and are aimed at the same cells of

the paradigm.¹⁰ In Kabiyè, however, the assignment of a low tone to a stem can vary from verb to verb and the cells involved also vary. This assignment depends on the application of a complex set of abstract rules linked to properties of the phonological shape of the lexical stem. Verbs which have lexical stems with the identical key property will behave in the same way, otherwise they won't.

4.2 Morphological tone conditioned by inflectional class

We have seen two situations where tone has a purely morphological role in the inflection. In the first situation, found in Tswana, it serves as a building block that applies to all verbs and to the same paradigmatic cells. The second situation, found in Kabiyè, is a possible deviation from the first one. Inflectional tone does not involve all verbs, and does not apply to the same cells. A second deviation is found in Acazolco Otomi, where we see that the application of morphological tone is conditioned by inflectional class membership of the verb.

Acazolco Otomi (Oto-Pamean) also has two tones involved in inflection: high (represented with an acute accent) and low (not represented). Hernández Green (2015a, 2015b, forthcoming), proposes that verbs in Acazolco Otomi fall into three main inflectional classes, defined by the specific set of inflectional proclitic markers they select.¹¹ Table 5 illustrates the basic paradigms of the three classes.¹² Class I verbs have two different paradigms depending on whether they are transitive or intransitive. Shading indicates slabs in the paradigm where the lexical tone of the stem is altered, in this case changed to a low tone.

Three things make morphological tone in Acazolco Otomi remarkable typologically:

To begin with, it has an unusual realization. The stem receives a low tone as a default, but it may also receive a high tone under specific circumstances: (i) when the stem is not monosyllabic but consists of a root plus a stem formative (indicated with the + sign) and has no further suffixation; (ii) when the verb is a V+N compound; and (iii) when the stem bears the human antipassive suffix *-te*, the enclitic =*'mbé* for plural exclusive, the delimitative =*thoho* 'just' or the counterfactual =*möhö*. These are illustrated in (5) as (a), (b) and (c), respectively.

¹⁰ When the stem already has an inherent high tone, grammatical high tone is not observable.

¹¹ There is an additional fourth class including patientive verbs which we will not further discuss here because it is mainly based on class III.

¹² Apart from the basic paradigm, Hernández Green (2015a, 2015b) shows that verbs also have two other paradigms called deictic and of adjunct registration.

Table 5: The three main verbal inflectional classes of Acazulco Otomi

			Class I		Class II		Class III			
			(tr) nũ 'see'	(intr) mĩhĩ 'sit'	ʔõhõ 'sleep'		ʔũni 'give'			
Realis	CPL	1 st	dí	nũ=ga	dí	mĩh=ka	dídí	ʔõh=ka	dí	ʔũn=ga
		2 nd	gí	nũ	gí	mĩh=a	gídí	ʔõh=a	gí	ʔũni
		3 rd	bi	nũ	bi	mĩh=a	bi	ʔõh=a	bi	ʔũni
	INCPL	1 st	drá	nũ=ga	drá	mĩh=ka	drá	ʔõh=ka	drádí	ʔũn=ga
		2 nd	grá	nũ	grá	mĩh=a	grá	ʔõh=a	grádí	ʔũni
		3 rd	ra	nũ	ra	mĩh=a	ra	ʔõh=a	radi	ʔũni
	HAB	1 st	dádí	nũ=ga	dán	mĩh=ka	dán	ʔõh=ka	dádí	ʔũn=ga
		2 nd	gádí	nũ	gán	mĩh=a	gán	ʔõh=a	gádí	ʔũni
		3 rd	an	nũ	an	mĩh=a	an	ʔõh=a	adi	ʔũni
PROSP	1 st	xtá	nũ=ga	xtá	mĩh=ka	xtá	ʔõh=ka	xtá	ʔũn=ga	
	2 nd	xtágí	nũ	xtágí	mĩh=a	xtágí	ʔõh=a	xtágí	ʔũni	
	3 rd	xta	nũ	xta	mĩh=a	xtagi	ʔõh=a	xtagi	ʔũni	
Irrealis	CPL	1 st	gũ	nũ=ga	gũ	mĩh=ka	gidi	ʔõh=ka	gidi	ʔũn=ga
		2 nd	gi	nũ	gi	mĩh=a	gidi	ʔõh=a	gidi	ʔũni
		3 rd	da	nũ	da	mĩh=a	di	ʔõh=a	di	ʔũni
	INCPL	1 st	gra	nũ=ga	gra	mĩh=ka	gra	ʔõh=ka	gradi	ʔũn=ga
		2 nd	gra	nũ	gra	mĩh=a	gra	ʔõh=a	gradi	ʔũni
		3 rd	dra	nũ	dra	mĩh=a	dra	ʔõh=a	dradi	ʔũni
	HAB	1 st	gadi	nũ=ga	gan	mĩhka	gan	ʔõh=ka	gadi	ʔũn=ga
		2 nd	gadi	nũ	gan	mĩh=a	gan	ʔõh=a	gadi	ʔũni
		3 rd	dan	nũ	dan	mĩh=a	dan	ʔõh=a	dadi	ʔũni

(5) INCPL.REALIS

- a. *ra=zo+ni* 'S/he's weeping'
 b. *ra=tũ-the* 'S/he's thirsty'
 (die-water)
 c. *ra=tsah-te* 'It's biting (people)'
 (e.g. a dog)

HAB.REALIS

- an=zó+ni* 'S/he weeps'
an=tũ-the 'S/he's often
 thirsty'
an=tsáh-te 'It bites
 (people)'
 (e.g. a dog)

If a verb has a lexical tone which is identical to that of the output of the rules, we cannot observe the effects of morphological tone.

- | | | | |
|-----|------------------|--------------------------|---|
| (6) | INCPL.REALIS | | HAB.REALIS |
| a. | <i>ra=tsa</i> | 'It's biting it' | /L/ > /L/ <i>an=tsà</i> 'it bites it' |
| b. | <i>ra=hí+xi</i> | 's/he's whistling' | /H/ > /H/ <i>an=hí+xi</i> 's/he whistles' |
| c. | <i>ra=núh-te</i> | 's/he's seeing (people)' | /H/ > /H/ <i>an=núh-te</i> 's/he sees (people)' |

The distribution of this grammatical tone in Acazolco Otomi is morphological. Like in Kabiye and Tswana, different cells in the paradigm that have little in common in functional terms require the same tone. However, the conditioning is different. Unlike in Tswana, but like in Kabiye, morphological tone on Acazolco Otomi is not found in all verbs. But while in Kabiye, its distribution is linked to the morphophonology, in Acazolco Otomi it is linked to the inflectional class of the verb. This means that it is a property of the paradigm of a certain class of verbs, which in itself is defined by the lexicon. Once speakers know to which inflectional class a given verb belongs, then it is predictable whether that verb will receive morphological tone or not, and if it does, which specific cells will be affected.

4.3 Morphological tone linked to the lexicon

A final type of morphological tone is characterized by a distribution bearing all the typical idiosyncrasies of the lexicon. This is found in Chichimec (Oto-Pamean). As shown in Table 6, in Chichimec there are verbs like *-pór* 'remove' and *-ʔor* 'appear' which retain their lexical tone throughout the paradigm.¹³

Table 6 further shows that verbs in Chichimec inflect by means of prefixes for the person of subject (third person has an additional number distinction) and for various TAM values, namely, present, future, anterior past, recent past, immediate past, potential and sequential (a sort of neutral grammatical tense used in the dependent clause in clause combining contexts). These prefixes do not carry lexical tone. In this they are different from the grammatical markers of Tswana or the Otomi clitics in (2). The tone they receive is provided by the word's prosodic phonology and it is manifested by inverting the lexical tone of the stem. The inverse happens in Tswana.

¹³ Chichimec data come from Angulo (1933).

Table 6: The Chichimec verbs *-pór* ‘remove’ and *-ʔor* ‘appear’

	Forms				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	<i>tu-pór</i>	<i>su-pór</i>	<i>u-pór</i>	<i>e-pór</i>	L-H	L-H	L-H	L-H
FUT	<i>gu-pór</i>	<i>ki-pór</i>	<i>ga-pór</i>	<i>ga-pór</i>	L-H	L-H	L-H	L-H
ANT.PST	<i>tu-pór</i>	<i>su-pór</i>	<i>u-pór</i>	<i>e-pór</i>	L-H	L-H	L-H	L-H
REC.PST	<i>ku-pór</i>	<i>ki-pór</i>	<i>ku-pór</i>	<i>ku-pór</i>	L-H	L-H	L-H	L-H
IMM.PST	<i>u-pór</i>	<i>i-pór</i>	<i>zu-pór</i>	<i>zu-pór</i>	L-H	L-H	L-H	L-H
POT	<i>nu-pór</i>	<i>mi-pór</i>	<i>mu-pór</i>	<i>mi-pór</i>	L-H	L-H	L-H	L-H
SEQ	<i>ra-pór</i>	<i>gi-pór</i>	<i>ru-pór</i>	<i>ru-pór</i>	L-H	L-H	L-H	L-H
PRS	<i>tú-ʔor</i>	<i>sú-ʔor</i>	<i>ú-ʔor</i>	<i>é-ʔor</i>	H-L	H-L	H-L	H-L
FUT	<i>gú-ʔor</i>	<i>kí-ʔor</i>	<i>gá-ʔor</i>	<i>gá-ʔor</i>	H-L	H-L	H-L	H-L
ANT.PST	<i>tú-ʔor</i>	<i>sú-ʔor</i>	<i>ú-ʔor</i>	<i>é-ʔor</i>	H-L	H-L	H-L	H-L
REC.PST	<i>kú-ʔor</i>	<i>kí-ʔor</i>	<i>kú-ʔor</i>	<i>kú-ʔor</i>	H-L	H-L	H-L	H-L
IMM.PST	<i>ú-ʔor</i>	<i>í-ʔor</i>	<i>zú-ʔor</i>	<i>zú-ʔor</i>	H-L	H-L	H-L	H-L
POT	<i>nú-ʔor</i>	<i>mí-ʔor</i>	<i>mú-ʔor</i>	<i>mí-ʔor</i>	H-L	H-L	H-L	H-L
SEQ	<i>rá-ʔor</i>	<i>gí-ʔor</i>	<i>rú-ʔor</i>	<i>rú-ʔor</i>	H-L	H-L	H-L	H-L

However, unlike the verbs in Table 6, most verbs in Chichimec have more than one stem: the majority have only two, while some have up to five. When this happens, the verbs organize these different stems into specific patterns (i.e. the distribution is not entirely chaotic). Following Palancar and Avelino (in preparation) there are at least 14 such patterns. One of those patterns is illustrated by the verbs *-tsá* ‘do’ in Table 7. As the tone value of affixes appears to be predictable, we shall only be concerned with the stems. To aid visualization of the patterns, each different stem has been assigned a different capital letter and a different type of shading.

The verb *-tsá* ‘do’ requires three stems, annotated here for convenience as ‘A’, ‘B’ and ‘D’. The A-stem could be taken as a lexical default, while the B-stem and the D-stem are used for a specific person value, e.g. the D-stem is used for the second person. Now take the verb *-sá* ‘win’ in table 8. This verb lacks a segmentally different D-stem, but we observe a contrast in tone for the same cells in which we have the D-stem of *-tsá* ‘do’.

Table 7: The stem pattern of the Chichimec verb *-tsá* ‘do’

	Forms				Stems				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
FUT	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
ANT.PST	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
REC.PST	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
IMM.PST	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
POT	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H
SEQ	-tsá	-fá	-tsá	-tjá	A	D	A	B	H	H	H	H

Table 8: The stem pattern of the Chichimec verb *-sá* ‘win’

	Forms				Stems				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
FUT	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
ANT.PST	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
REC.PST	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
IMM.PST	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
POT	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H
SEQ	-sá	-sa	-sá	-tjá	A	A	A	B	H	L	H	H

For such a case, one could in principle posit that the verb *-sá* ‘win’ follows the same stem alternation pattern as *-tsá* ‘do’, and that the low tone is just a phonological property of the D-stem of *-sá* to distinguish it from its A-stem (i.e., the stems would otherwise be segmentally homophonous). If this view is correct, tone would bear no inflectional function here; instead being a lexical property of stems. Similarly, if we pursue this analysis, we would have to conclude that the D-stem of the verb *-tsá* ‘do’ just happens to have the same lexical tone as the A-stem.

A more complex pattern of stem alternations is given in Table 9. Here things start getting more interesting since we observe a high tone in the potential form of the verbs *-ʔi* ‘want’ and *-hun* ‘throw’.

Table 9: The stem pattern of the Chichimec verbs *-ʔa* ‘bury’, *-ʔi* ‘want’ and *-hun* ‘throw’

	Forms				Stems				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	<i>-ʔa</i>	<i>-ʔa</i>	<i>-ʔa</i>	<i>-rʔa</i>	A	A	A	B	L	L	L	L
FUT	<i>-ʔa</i>	<i>-ta</i>	<i>-nda</i>	<i>-rʔa</i>	A	D	C	B	L	L	L	L
ANT.PST	<i>-ʔa</i>	<i>-ta</i>	<i>-nda</i>	<i>-rʔa</i>	A	D	C	B	L	L	L	L
REC.PST	<i>-ʔa</i>	<i>-ta</i>	<i>-nda</i>	<i>-rʔa</i>	A	D	C	B	L	L	L	L
IMM.PST	<i>-ʔa</i>	<i>-ta</i>	<i>-ʔa</i>	<i>-rʔa</i>	A	D	A	B	L	L	L	L
POT	<i>-ʔa</i>	<i>-ta</i>	<i>-ʔa</i>	<i>-rʔa</i>	A	D	A	B	L	L	L	L
SEQ	<i>-ʔa</i>	<i>-ta</i>	<i>-ʔa</i>	<i>-rʔa</i>	A	D	A	B	L	L	L	L
PRS	<i>-ʔi</i>	<i>-ʔi</i>	<i>-ʔi</i>	<i>-rʔí</i>	A	A	A	B	L	L	L	H
FUT	<i>-ʔi</i>	<i>-tí</i>	<i>-ndí</i>	<i>-rʔí</i>	A	D	C	B	L	H	H	H
ANT.PST	<i>-ʔi</i>	<i>-tí</i>	<i>-ndí</i>	<i>-rʔí</i>	A	D	C	B	L	H	H	H
REC.PST	<i>-ʔi</i>	<i>-tí</i>	<i>-ndí</i>	<i>-rʔí</i>	A	D	C	B	L	H	H	H
IMM.PST	<i>-ʔi</i>	<i>-tí</i>	<i>-ʔi</i>	<i>-rʔí</i>	A	D	A	B	L	H	L	H
POT	<i>-ʔí</i>	<i>-tí</i>	<i>-ʔí</i>	<i>-rʔí</i>	A	D	A	B	H	H	H	H
SEQ	<i>-ʔí</i>	<i>-tí</i>	<i>-ʔí</i>	<i>-rʔí</i>	A	D	A	B	L	H	H	H
PRS	<i>-hun</i>	<i>-hun</i>	<i>-hun</i>	<i>-hún</i>	A	A	A	B?	L	L	L	H
FUT	<i>-hun</i>	<i>-nhín</i>	<i>-nhín</i>	<i>-hún</i>	A	D	C	B?	L	H	L	H
ANT.PST	<i>-hun</i>	<i>-nhín</i>	<i>-nhín</i>	<i>-hún</i>	A	D	C	B?	L	H	L	H
REC.PST	<i>-hun</i>	<i>-nhín</i>	<i>-nhín</i>	<i>-hún</i>	A	D	C	B?	L	H	L	H
IMM.PST	<i>-hun</i>	<i>-nhín</i>	<i>-hun</i>	<i>-hún</i>	A	D	A	B?	L	H	L	H
POT	<i>-hún</i>	<i>-nhín</i>	<i>-hún</i>	<i>-hún</i>	A	D	A	B?	H	H	H	H
SEQ	<i>-hun</i>	<i>-nhín</i>	<i>-hun</i>	<i>-hún</i>	A	D	A	B?	L	H	L	H

Verbs like *-ʔi* ‘want’ and *-hun* ‘throw’ could be said to have the same pattern as *-ʔa* ‘bury’, only that their B-stem and D-stem would have a high tone.¹⁴ But the high tone in the A-stem of the verbs *-ʔi* and *-hun* for the potential of the first person and

¹⁴ We could assume that the B-stem of *-hun* ‘throw’ is built by a tone contrast like the D-stem of *-sá* ‘win’ in Table 8.

the third person singular and dual is surprising. One does not expect a different tone here because at the segmental level there is no comparable stem change in other verbs for the same cells. We could take this to be an incipient inflectional use of tone, in this case to mark the potential mood, but its further idiosyncrasies would perhaps let us down.

For example, in other verbs, instead of a high tone we find a low tone, like in the verb *-mê* ‘defend’ in Table 10. As the A-stem has a high tone, one could think of a mechanism of tonal inversion. Notice that a low tone also applies to the second person, suggesting that we are facing a verb with a stem pattern somehow similar to that of Table 8, but with a D-stem like in Table 9.

Table 10: The stem pattern of the Chichimec verb *-mê* ‘defend’

	Forms				Stems				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	A	A	B	H	H	H	H
FUT	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	H	L	H	H
ANT.PST	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	H	L	H	H
REC.PST	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	H	L	H	H
IMM.PST	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	H	L	H	H
POT	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	L	L	L	H
SEQ	<i>-mê</i>	<i>-mê</i>	<i>-mê</i>	<i>-mhê</i>	A	D?	A	B	H	L	H	H

Table 11: The stem pattern of the Chichimec verb *-tsa* ‘hurt’

	Forms				Stems				Tones			
	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL	1 st	2 nd	3SG/DU	3PL
PRS	<i>-tsa</i>	<i>-fa</i>	<i>-tsa</i>	<i>-tfa</i>	A	D	A	B	L	L	L	L
FUT	<i>-tsá</i>	<i>-fa</i>	<i>-za</i>	<i>-tfa</i>	A	D	C	B	H	L	H	H
ANT.PST	<i>-tsá</i>	<i>-fa</i>	<i>-zá</i>	<i>-tfa</i>	A	D	C	B	H	L	H	H
REC.PST	<i>-tsa</i>	<i>-fa</i>	<i>-za</i>	<i>-tfa</i>	A	D	C	B	L	L	L	L
IMM.PST	<i>-tsá</i>	<i>-fa</i>	<i>-tsá</i>	<i>-tfa</i>	A	D	A	B	H	L	H	H
POT	<i>-tsa</i>	<i>-fa</i>	<i>-tsa</i>	<i>-tfa</i>	A	D	A	B	L	L	L	L
SEQ	<i>-tsá</i>	<i>-fa</i>	<i>-tsá</i>	<i>-tfa</i>	A	D	A	B	H	L	H	H

To complicate things further, the verb *-tsá* ‘hurt’ in Table 11 shows yet another pattern. Here the potential is encoded throughout all persons by a low tone.¹⁵ But most importantly, this pattern is not exclusive to the potential as it also applies to the forms of the present and the recent past.

The tonal patterns increase in numbers as we go from one verb to the next. Our understanding of Chichimec verbs is based on the behavior of only 168 verbs from Angulo (1933) and thus remains limited. Of these, only 31 display some sort of tonal perturbations, but only a handful are of the sort described for the verbs in tables 9–11. This suggests that morphological tone in Chichimec remains tightly and intricately linked to the lexicon. The Chichimec case reminds us of the stem tonal stem patterns in Khaling whose distribution Jacques (this volume) accounts for appealing to the rules of a historical sound change. However, for Chichimec the key to the system appears to have been lost. This means that there is no indication from the morphophonology or the morphology as to what verbs have morphological tone and which do not. Similarly, once we know that a given verb has tone perturbations, we are not in a position to predict in what cells this will happen. Speakers need to store this information as part of the lexical entry.

In the following sections, I present a final type of inflectional tone which is different from all others presented so far. In this type, there is a great degree of tonal allomorphy in the realization of morphosyntactic values. The optimal way to account for this type of allomorphy is by means of inflectional classes defined by tone. To show this type, let us start with Tlatepuzco Chinantec, another Oto-Manguean language.

5 Tonal classes: Inflectional classes made by tone

Tlatepuzco Chinantec has six tones involved in lexical contrasts (cf. Mandarin Chinese has only four): three level tones and two rising contours (low-mid /12/ and low-high /13/). Out of these six tones, five are involved in verbal inflection in rather intricate ways. Verbs in Chinantec inflect for TAM by means of prefixes (including zero suffixes or bare stems). These prefixes, in turn, are associated with stems inflected for aspect/mood by tone, e.g. the past prefix *ca*¹ and the hodiernal past prefix *na*² are always affixed to the completive stem. The same tonally inflected stems also indicate distinctions involving the person of the

¹⁵ Whether the verbs *-ʔi* ‘want’ and *-hun* ‘throw’ in Table 9 have the potential encoded in the same way as *-mĕ* ‘defend’ is unclear when it comes to the cells for the second person and for the third person singular and dual.

subject. The relevant person values for the inflection are first person singular and dual, second person and third person (see for example Merrifield 1968, Pace 1990, Foris 2000, etc.).

Table 12: The twelve inflected stems of six verbs in Tlatepuzco Chinantec

		1SG		1PL		2		3	
CPL-STEM	a. 'bend'	húʔ ¹	1	húʔ ¹³	13	húʔ ¹	1	húʔ ²	2
	b. 'gnaw'	tsóʔ ¹	1	tsóʔ ¹³	13	tsóʔ ¹	1	tsóʔ ¹	1
	c. 'bring'	quián ¹	1	quián ¹³	13	quián ¹²	12	quián ²	2
	d. 'call'	tóʔ ¹	1	tóʔ ¹³	13	tóʔ ¹²	12	tóʔ ¹	1
	e. 'cut'	tiu ¹	1	tiu ³	3	tiu ³²	32	tiu ¹	1
	f. 'scare'	ʔán ²	2	ʔán ¹³	13	ʔán ¹	1	ʔán ²	2
INCL-STEM	a. 'bend'	húʔ ¹²	12	húʔ ¹²	12	húʔ ¹²	12	húʔ ²	2
	b. 'gnaw'	tsóʔ ¹²	12	tsóʔ ¹²	12	tsóʔ ¹²	12	tsóʔ ¹²	12
	c. 'bring'	quián ¹²	12	quián ¹²	12	quián ¹²	12	quián ²	2
	d. 'call'	tóʔ ¹²	12	tóʔ ¹²	12	tóʔ ¹²	12	tóʔ ²	2
	e. 'cut'	tiu ¹²	12	tiu ²	2	tiu ²	2	tiu ¹²	12
	f. 'scare'	ʔán ²	2	ʔán ¹²	12	ʔán	2	ʔán ²	2
IRR-STEM	a. 'bend'	húʔ ¹³	13	húʔ ¹³	13	húʔ ¹³	13	húʔ ²	2
	b. 'gnaw'	tsóʔ ¹³	13	tsóʔ ¹³	13	tsóʔ ¹³	13	tsóʔ ¹	1
	c. 'bring'	quián	13	quián ¹³	13	quián ¹³	13	quián ²	2
	d. 'call'	tóʔ ¹³	13	tóʔ ¹³	13	tóʔ ¹³	13	tóʔ ²	2
	e. 'cut'	tiu ¹³	13	tiu ³	3	tiu ³	3	tiu ¹	1
	f. 'scare'	ʔán ³	3	ʔán ¹³	13	ʔán ³	3	ʔán ²	2

The sources propose that the paradigm of a verb can be reconstructed from three principal stems: the completive, the incompletive and the irrealis stem. In Table 12, we observe the inflected stems of six different verbs (a-f).¹⁶

¹⁶ The data of Tlatepuzco Chinantec come from Merrifield and Anderson (2007).

A glance at the distribution of tones in the inflected forms in Table 12 suffices to show that tone in Tlatepuzco Chinantec verbal inflection does not work in the same predictable ways as it does in languages like Yoloxóchitl Mixtec in §3. In Chinantec there is no evident consistent mapping between form and meaning. For example, for the encoding of the first person singular (the first column in Table 12), one could say that tone values /1/, /12/ and /13/ encode the bundles {1SG, CPL}, {1SG, INCPL} and {1SG, IRR}, respectively. While this would already have to exclude the pattern in the verb in (f), to complete this picture one would also have to account for the many other instances where exactly the same tone values are found in other cells of the paradigm to encode other, very different grammatical information, as indicated by the shading. This suggests that the distribution of tone values is, if anything, morphological.

This type of tonal inflection provides further challenges for a theory of tone, because contrary to morphosyntactic tone, identifying a form of a verb in Chinantec that may serve as a lexical base from which the tone values of other forms can be derived is, in our humble opinion, a futile task (for a similar position, but stated differently, see Finkel & Stump, 2009). However, *qui quaerit invenit...* and Baerman & Palancar (2014) propose that there is still some structure beyond the apparent chaos.

In this regard, Baerman & Palancar (2014) propose that one way to start finding some paradigmatic structure can be accomplished by first isolating third person forms from other person forms. If we then concentrate on the forms involving other persons, we start to see some patterns emerging. For example, one possible analysis of the data in Table 12 is that the verbs in (a-d) follow the same pattern: incomplete is marked by /12/ and irrealis by /13/ without person distinctions. But the verbs then fall into two inflectional classes attending to how they mark the complete, all encode first person singular and plural by /1/ and /13/, respectively, whereas for the second person, verbs in (a-b) use /1/ while the verbs in (c-d) use /2/. Verbs in (e) and (f) would have to be accounted for as being deviant or irregular. All such verbs would then have to be further analyzed as belonging to other independent classes attending to the ways they encode the third person. In reality, things become more complex when we move from a handful of verbs to hundreds of them. As a result of this, a typical Chinantecan language is often analyzed as consisting of 70 or more such classes (see Palancar 2014).

A very similar system to Chinantec is found in Yaitepec Chatino, an Eastern Chatino language of the Zapotecan branch with ten contrastive tones.¹⁷ Verbs in Yaitepec Chatino inflect for four aspect-mood values: complete, potential,

¹⁷ According to Rasch (2015), there are three level tones: /1/ High with a slight rising; /2/ High, but slightly lower than /1/; /3/ Mid; and seven contours: /31/ Mid-rising; /21/ High-rising, rising

habitual and progressive (also called ‘continuous’ or ‘continuative’). Verbs have different stems for each value, and tone plays a major role in the marking of aspect-mood distinctions as well as distinctions of person of subject. Examples of three different paradigms are given in Table 13.

Table 13: The Yaitepec Chatino verbs *xnu* ‘leave behind’, *swi* ‘choose’ and *jwi* ‘kill’

	<i>xnu</i> ‘leave behind’			<i>swi</i> ‘choose’			<i>jwi</i> ‘kill’		
	1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd
CPL	nwxnu ²¹	nwxnu ³¹	nwxnu ¹	nswi ³	nswi ³¹	nswi ¹	yjwi ³	yjwi ²	yjwi ¹²
POT	xnu ²¹	xnu ³¹	xnu ¹	swi ³	swi ³¹	swi ⁴³	kjwi ³	kjwi ³¹	kjwi ²⁴
HAB	nxnu ²¹	nxnu ³¹	nxnu ¹	nswi ³	nswi ³¹	nswi ⁴³	ntyjwi ³	ntyjwi ³¹	ntyjwi ²⁴
PROG	nxnu ²¹	nxnu ³¹	nxnu ¹	nswi ²	nswi ³¹	nswi ¹	ntjwi ²	ntjwi ²	ntjwi ²

The verb *xnu* ‘leave behind’ has contrastive tone for person, but not for aspect-mood. The verb *swi* ‘choose’ has invariant tone for the second person. The forms for potential and habitual receive the same tone; a pattern followed by many verbs, but not all. The verb *jwi* ‘kill’ is similar to *swi* ‘choose’ in many respects, but we have different tones involved in the marking of second person. The dataset in Rasch (2015) includes 324 inflected verbs and they could be said to instantiate almost 90 different tonal paradigms. A handful of such paradigms have an alternative tone in one or more cells, for example the transitive verb *xkwi* ‘boil’ inflects like *xnu* ‘leave behind’, but as it can also have a tone /31/ for first person complete, it is treated as belonging to a different class. If we discard such cases, and a handful of inanimate verbs that cannot be inflected in all persons for semantic reasons, we are left with the 75 paradigms given in Table 14 in the Appendix.¹⁸

more recognizably than /1/; /14/ High and sharply falling, starting at /1/ and continuing to fall; /12/ and /23/ High and falling; /24/ High and sharply falling; and /43/ Low and rising.

18 The most complex tonal system in Oto-Manguean, and possibly in the world’s languages, is found in Tlapanec (Tlapanecan). In his grammar of the Tlapanec language of Malinaltepec, Suárez (1983) provides information about the tones in the inflectional paradigms of about 160 verbs. Out of these, 135 fall into at least 18 different tonal classes, but 24 of them have more complex tonal paradigms. Such verbs are the type of transitive verbs that express actions that may involve two persons in an actor-undergoer sort of relation (e.g., 1>2, 1pl>2, 2>3, etc.). Each of these types of interrelation constitutes a cell in the paradigm of these verbs. There are 30 cells in total and tone plays a fundamental role in the making of such paradigms. These 24 verbs, which were selected at random by Suárez to register tone, already display 20 different tonal classes!

All ten lexical tones are also found at least once in the 75 different paradigms. A correlation of tone value and person is given in Table 15 (no relevant correlations are found involving aspect-mood). The figures indicate the number of times a given tone is found in a given cell across the paradigms of the classes.

Table 15: Tone values used in the tonal classes of Yaitepec Chatino per person value

Tones	1 st	2 nd	3 rd	Total	1 st	2 nd	3 rd	Total
/1/	19	47	25	91	21 %	52 %	27 %	100 %
/2/	28	48	10	86	32 %	57 %	11 %	100 %
/3/	29	0	35	64	45 %	0 %	55 %	100 %
/12/	22	0	38	60	36 %	0 %	64 %	100 %
/14/	0	0	3	3	0 %	0 %	100 %	100 %
/21/	41	0	24	65	63 %	0 %	37 %	100 %
/23/	5	0	15	20	25 %	0 %	75 %	100 %
/24/	0	0	55	55	0 %	0 %	100 %	100 %
/31/	7	130	14	151	4 %	87 %	9 %	100 %
/43/	74	0	6	80	93 %	0 %	7 %	100 %

Yaitepec Chatino is a mixed system regarding the mapping of a specific form to a specific morphosyntactic value. There are, for example, tones which are associated with a specific person: tone /24/ only occurs in third person forms; tone /43/ could be treated as an exponent of first person; and we could practically say the same about tone /31/ but for second person. Other tones are used for more than one person so a clear mapping between form and meaning breaks down here. Also, while all ten tones are found in third person forms and most of them in first person forms, only three of them are used to encode second person. This suggests that the marking of second person attends to other sorts of rules. We can explore this possibility further.

For example, in complex systems like Yaitepec Chatino, it often happens that tone assignment for a specific value may be more predictable for some areas of the paradigm than for other areas. This is the view taken by implicative approaches to inflectional morphology in works such as Bonami & Boyé (2007), Montermini & Bonami (2013), and Bonami & Luís (2014), which are inspired by Morin (1987). Such works explore the degree of predictability that exists among the cells of a given paradigm. As for Yaitepec Chatino, it is possible to find instances where the tone of the second person can be predicted from the marking of the third person

(the same cannot be said of the marking for the first person, which is far more chaotic, i.e. unpredictable). This is shown in Table 16, where we compare the correlations between the tone used for the {3, CPL} with the one used for the {2, CPL}.

Table 16: Tone correlations between 3.CPL and 2.CPL in Yaitepec Chatino

{3, CPL}	{2, CPL}				{2, CPL}		
	/1/	/2/	/31/		/1/	/2/	/31/
/1/	0	0	17	17	–	–	100 %
/3/	0	0	8	8	–	–	100 %
/21/	0	0	2	2	–	–	100 %
/24/	0	0	1	1	–	–	100 %
/14/	0	0	1	1	–	–	100 %
/31/	10	0	0	10	100 %	–	–
/23/	8	1	0	9	88 %	12 %	–
/12/	3	18	1	22	14 %	82 %	4 %
/2/	0	3	2	5	–	60 %	40 %

At least six tones (1, 3, 21, 24, 14 and 31) when found in the third person complete render a predictability value of 100 % with regard to the tone value we will find in the second person complete. This actually happens regardless of aspect-mood, because the same correlation is found in the potential and the progressive (although not shown in the table). But beyond those tones, certainty of outcomes cannot be guaranteed. The next best correlation is /23/, which will most likely render /1/ for second person, followed by /12/, etc. This suggests that for a substantial part of the 75 patterns, the assignment of tone to the second person is subject to a morphophonological operation. But in situations like this, when some tone value is predictable from another tone value, one still needs to appeal to the morphology, because the specific tone value we are to find for the second person is not in itself given by the phonology. In other words, there is nothing inherently phonological that would make a tone /31/ become /1/ or vice versa. Yaitepec Chatino evolved from an ancestor system which might have been very similar to conservative Zenzontepec Chatino as described in Campbell (this volume) where the tonal marking for second person is to a large extent still predictable.

Whether speakers of Chinantec or Yaitepec Chatino learn their paradigms in the way we analyze them or whether they learn them on a cell-to-cell basis as if each form were an independent word remains an open question, which we must

leave to the experts of child language acquisition. For the moment, the structure in inflectional classes appears to be the most convenient way to deal with such systems.

6 Final remarks

I have outlined here a typology of the different situations where tone is involved in inflection taking a morphological perspective. I have illustrated most of the types using Oto-Manguean languages, as the internal linguistic diversity of this phylum from the Americas displays almost all types. From the lexical end of the continuum, where grammatical affixes receive their own lexical tone, to the morphosyntactic end, where tone works as a regular exponent of a given grammatical category, there are at least five other types; each one with its own logic.

We are at the dawn of understanding how complex systems work. But to do better we need help. We need full paradigms in grammars of tonal languages, not just rules, abstract representations or examples of how a given form bearing tone is used in a natural context. This is a cordial invitation to descriptive linguists to enrich the field with new data on inflection. It matters. It matters in a time when most languages with complex morphology are dying. By doing so, we will be paying tribute both to the languages and to the field of linguistics, because in a hundred years from now, when all of us are gone, it will only be our data that shall remain for future linguists to continue increasing our understanding of our human languages.

Acknowledgements

This paper was supported by the ESRC/AHRC project ES/I029621/1 “Endangered Complexity: Inflectional classes in Oto-Manguean languages”. I thank Jean Léo Léonard for comments on an earlier version of this typology. I am heartily grateful to Timothy Feist for accepting reviewing this paper with a short notice and for all the comments and suggestions that have contributed a lot to its improvement. Thanks to the anonymous referee for suggesting that a paper like this should be included in the volume.

Appendix

Table 14: Tonal paradigms in Yaitepec Chatino

		CPL			POT			HAB			PROG		
		1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd
Class 1	1,1,1A	21	31	1	21	31	1	21	31	1	21	31	1
Class 2	1,1,1C	3	31	1	3	31	1	3	31	1	2	31	1
Class 3	1,1,1D	43	31	1	43	31	1	43	31	1	43	31	1
Class 4	1,1,1E	21	31	1	3	31	1	3	31	1	21	31	1
Class 5	1,21,1A	3	31	1	3	31	21	2	31	21	2	31	1
Class 6	1,21,1B	21	31	1	21	31	21	21	31	21	21	31	1
Class 7	1,21,1E	1	31	1	21	31	21	21	31	21	1	31	1
Class 8	1,21,1H	21	31	1	21	31	21	21	31	21	1	31	1
Class 9	1,21,1I	21	31	1	21	31	21	21	31	21	2	31	1
Class 10	1,21,1J	21	31	1	21	31	21	21	31	21	43	31	1
Class 11	1,21,2A	3	31	1	3	31	21	3	31	21	21	31	2
Class 12	1,3,12A	43	2	12	3	31	3	3	31	3	43	2	12
Class 13	1,3,2B	21	31	1	3	31	3	3	31	3	2	31	2
Class 14	1,43,1A	3	31	1	3	31	43	3	31	43	2	31	1
Class 15	1,43,1B	3	31	1	3	31	43	3	31	43	1	31	1
Class 16	1,43,1C	43	31	1	43	31	43	43	31	43	43	31	1
Class 17	12,12,12A	12	2	12	12	2	12	12	2	12	12	2	12
Class 18	12,12,12B	12	1	12	12	2	12	12	2	12	12	2	12
Class 19	12,12,12C	43	1	12	43	1	12	43	1	12	43	1	12
Class 20	12,12,12D	43	2	12	43	2	12	43	2	12	43	1	12
Class 21	12,12,12E	12	1	12	12	1	12	12	1	12	12	1	12
Class 22	12,12,2A	43	2	12	43	2	12	43	2	12	43	2	2
Class 23	12,12,2B	12	2	12	12	2	12	12	2	12	12	2	2
Class 24	12,12,2C	12	2	12	12	2	12	12	2	12	2	2	2
Class 25	12,24,12B	43	2	12	43	31	24	43	31	24	43	2	12
Class 26	12,24,14A	12	2	12	43	31	24	43	31	24	12	2	14

Table 14: (continued)

		CPL			POT			HAB			PROG		
		1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd
Class 27	12,24,14B	1	2	12	43	31	24	43	31	24	1	2	14
Class 28	12,24,24A	2	2	12	43	31	24	43	31	24	43	31	24
Class 29	12,24,2A	43	2	12	43	31	24	43	31	24	43	2	2
Class 30	12,24,2B	3	2	12	3	31	24	3	31	24	2	2	2
Class 31	12,24,2C	1	2	12	21	31	24	21	31	24	1	2	2
Class 32	12,24,2D	12	2	12	43	31	24	43	31	24	2	2	2
Class 33	12,24,2E	3	31	12	43	31	24	43	31	24	2	31	2
Class 34	12,24,3A	2	2	12	43	31	24	21	31	3	43	31	3
Class 35	12,3,2A	12	2	12	3	31	3	2	31	3	2	31	2
Class 36	12,3,2B	12	2	12	3	31	3	3	31	3	2	2	2
Class 37	12,3,2C	43	2	12	3	31	3	3	31	3	2	31	2
Class 38	14,14,14A	43	31	14	43	31	14	43	31	14	43	31	14
Class 39	2,2,2A	2	31	2	2	31	2	2	31	2	2	31	2
Class 40	2,2,2B	2	2	2	2	2	2	2	2	2	2	2	2
Class 41	2,24,14A	2	2	2	43	31	24	43	31	24	43	2	14
Class 42	2,24,14B	1	31	2	43	31	24	43	31	24	43	31	14
Class 43	2,3,14A	2	2	2	43	31	3	43	31	24	2	2	14
Class 44	21,21,1A	21	31	21	21	31	21	21	31	21	21	31	1
Class 45	21,21,21A	21	31	21	21	31	21	21	31	21	21	31	21
Class 46	23,23,12A	23	1	23	23	1	23	23	1	23	12	1	12
Class 47	23,23,12D	43	1	23	43	1	23	43	1	23	43	1	12
Class 48	23,23,23A	43	1	23	43	1	23	43	1	23	43	1	23
Class 49	23,24,12A	43	1	23	43	31	24	43	31	24	43	1	12
Class 50	23,24,12B	23	1	23	43	31	24	43	31	24	12	1	12
Class 51	23,24,12C	1	1	23	21	31	24	21	31	24	1	1	12
Class 52	23,24,12D	1	1	23	43	31	24	21	31	24	1	1	12
Class 53	23,24,23A	23	2	23	43	31	24	43	31	24	43	2	23

Table 14: (continued)

		CPL			POT			HAB			PROG		
		1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd	1 st	2 nd	3 rd
Class 54	23,24,24A	23	1	23	43	31	24	43	31	24	43	31	24
Class 55	24,24,14A	43	31	24	43	31	24	43	31	24	43	31	14
Class 56	3,2,2A	2	31	3	3	31	2	3	31	3	2	31	2
Class 57	3,3,2A	3	31	3	3	31	3	3	31	3	2	31	2
Class 58	3,3,2B	1	31	3	21	31	3	21	31	3	1	31	2
Class 59	3,3,2C	21	31	3	21	31	3	21	31	3	21	31	2
Class 60	3,3,2D	3	31	3	3	31	3	3	31	3	2	2	2
Class 61	3,3,2E	1	31	3	43	31	3	43	31	3	43	31	2
Class 62	3,3,2F	3	31	3	3	31	3	3	31	3	3	31	2
Class 63	3,3,3A	21	31	3	21	31	3	21	31	3	21	31	3
Class 64	3/,24,12A	31	1	31	43	31	24	43	31	24	12	1	12
Class 65	3/,24,12C	43	1	31	43	31	24	43	31	24	43	1	12
Class 66	3/,24,24A	1	1	31	21	31	24	21	31	24	21	31	24
Class 67	3/,24,3/A	43	1	31	43	31	24	43	31	24	43	1	31
Class 68	3/,24,3/B	1	1	31	21	31	24	21	31	24	1	1	31
Class 69	3/,24,3/C	21	1	31	21	31	24	21	31	24	21	1	31
Class 70	3/,24,3/D	31	1	31	43	31	24	43	31	24	31	1	31
Class 71	3/,24,3/E	43	1	31	43	31	24	43	31	24	31	1	31
Class 72	3/,3/,3/A	31	1	31	31	1	31	31	1	31	31	1	31
Class 73	3/,3/,3/B	43	1	31	43	1	31	43	1	31	43	1	31
Class 74	1,21,1A	21	31	1	3	31	21	21	31	21	21	31	1
Class 75	1,21,1B	3	31	1	21	31	21	21	31	21	2	31	1

References

- de Angulo, Jaime. 1933. The Chichimeco language (Central Mexico). *International Journal of American Linguistics* 7–3/4: 152–194.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.
- Baerman, Matthew and Enrique L. Palancar. 2014. Baerman, Matthew and Enrique L. Palancar. 2014. The organization of Chinantec tone paradigms. *Carnets de Grammaire* 22 [Proceedings of the Décembrettes, 8th International conference on morphology], 46–59.
- Bonami, Olivier and Ana R. Luís. 2014. Sur la morphologie implicative dans la conjugaison du portugais : une étude quantitative. *Mémoires de la Société de Linguistique de Paris*, Vol. 22: 111–151.
- Bonami, Olivier and Gilles Boyé. 2007. Remarques sur les bases de la conjugaison. In Elisabeth Delais-Roussarie and Laurence Labrune (eds.), *Des sons et des sens*, 77–90. Paris: Hermès.
- Campbell, Lyle, Terrence Kaufman, and Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62:530–70.
- Campbell, Lyle. 1997. *The historical linguistics of Native America*. [Oxford Studies in Anthropological Linguistics, 4]. Oxford: Oxford UP: Oxford University Press.
- Creisels, Denis, Anderson M. Chebanne, and Heather W. Nkhwa. 1997. *Tonal Morphology of the Setswana Verb*. München: Lincom Europa.
- Creisels, Denis. 2006. Tswana verb morphology and the lexical integrity principle. *Lingue e Linguaggio* 5:1, 1–18.
- DiCanio, Christian T., Jonathan Amith and Rey Castillo García. 2012. Phonetic alignment in Yoloxóchitl Mixtec. Paper read at the 2012 meeting of the Society for the Study of the Indigenous Languages of America, Portland.
- Finkel, Raphael and Gregory Stump. 2009. Principal parts and degrees of paradigmatic transparency. In James P. Blevins and Juliette Blevins (eds.), *Analogy in Grammar: Form and Acquisition*, 13–53. Oxford: Oxford University Press.
- Finkel, Raphael and Gregory T. Stump. 2007. Principal Parts and Morphological Typology. *Morphology* 17–1: 39–75.
- Foris, David Paul. 2000. *A grammar of Sochiapam Chinantec* [Studies in Chinantec Languages 6]. Dallas, TX: SIL International and the University of Texas at Arlington.
- Hernández Green, Néstor. 2015a. Acazolco Otomi. In Timothy Feist and Enrique L. Palancar. The Oto-Manguean inflectional class database. University of Surrey. <www.smg.surrey.ac.uk/oto-manguean>
- Hernández Green, Néstor. 2015b. Morfosintaxis verbal del otomí de Acazolco. Unpublished PhD thesis in Indoamerican Linguistics at the Centre for Research and High Studies in Social Anthropology (CIESAS), Mexico.
- Hernández-Green, Néstor. Forthcoming. The inflectional classes of Acazolco Otomi. In Enrique L. Palancar, Matthew Baerman, and Timothy Feist (eds.), *Inflectional complexity and verb classes: New facts from the Oto-Manguean languages of Mexico*.
- Levy, Paulette. 1987. ‘La irregularidad en otomangue’. Plenary talk given at *I Coloquio Internacional “Mauricio Swadesh”*. Mexico D.F.
- Merrifield, William R. 1968. *Palantla Chinantec grammar* [Papeles de la Chinantla 5, Serie Científica 9]. Mexico City: Museo Nacional de Antropología.

- Merrifield, William. R. and A. E. Anderson. 2007. *Diccionario Chinanteco de la diáspora del pueblo antiguo de San Pedro Tlatapuzco, Oaxaca*. [2nd Edition] [Serie de vocabularios y diccionarios indígenas “Mariano Silva y Aceves” 39]. Mexico DF: Summer Linguistic Institute.
- Montermini, Fabio and Olivier Bonami. 2013. Stem spaces and predictability in verbal inflection, *Lingue e linguaggio* 12: 171–190.
- Morin, Yves-Charles. 1987. Remarques sur l’organisation de la flexion en français, *International Review of Applied Linguistics* 77–78: 13–91.
- Pace, W. J. 1990. Comaltepec Chinantec verb inflection. In W. R. Merrifield and C. R. Rensch (eds.), *Syllables, tones and verb paradigms* [Studies in Chinantec Languages 4], 21–62. Arlington, Texas: The Summer Linguistic Institute and the University of Texas at Arlington.
- Palancar, Enrique L. 2014. Revisiting the complexity of the Chinantecan verb conjugation classes. In Jean Léo Léonard and Alain Kihm (eds.), *Patterns in Meso-American morphology*, 77–102. Paris: Michel Houdiard.
- Palancar, Enrique L. and Heriberto Avelino. Forthcoming. Inflectional complexity and verb classes in Chichimec. In Enrique L. Palancar, Matthew Baerman, and Timothy Feist (eds.), *Inflectional complexity and verb classes: New facts from the Oto-Manguean languages of Mexico*.
- Pike Kenneth. 1948. *Tone languages: A technique for determining the number and types of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor: University of Michigan Press.
- Rasch, Jeffrey. 2015. Yaitepec Chatino. In Timothy Feist and Enrique L. Palancar. The Oto-Manguean inflectional class database. University of Surrey. <www.smg.surrey.ac.uk/oto-manguean>
- Roberts, David. 2013. *Conjugaison des verbes en kabiyè (Togo): Tableaux types, règles d’emploi, et index kabiyè-français et français-kabiyè des verbes*. Paris: L’Harmattan.
- Suárez, Jorge A. 1983. *La Lengua Tlapaneca de Malinaltepec*. Mexico City: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.

1. Clases léxicas, posesión y estructuras de cláusula compleja en Mesoamérica: una introducción

*Enrique L. Palancar** y *Roberto Zavala Maldonado***

Este libro contiene 12 trabajos de investigación morfosintáctica en varias lenguas del área lingüística de Mesoamérica, que aparecen en los capítulos del 2 al 13. Como denominador común, estos estudios presentan un acercamiento morfosintáctico contemporáneo a varios fenómenos sobre estas lenguas que se desconocían hasta el momento y que aportan nuevas e interesantes luces para incrementar el conocimiento lingüístico del área. Los diferentes capítulos son estudios de lenguas específicas, a excepción de dos contribuciones que hacen una indagación del fenómeno estudiándolo en varias lenguas representativas de una misma familia lingüística. Las lenguas tratadas provienen de diferente genética y son las siguientes: *(i)* el p'orhépecha o tarasco, una lengua aislada del estado de Michoacán; *(ii)* dos lenguas yutoaztecas: el cora meseño del grupo corachol y el náhuatl de Amanalco, una lengua nahua de Texcoco; *(iii)* el otomí, que constituye una familia de lenguas de la rama otopame de la familia otomangue; *(iv)* de la familia mixe-zoque, se presenta un trabajo sobre una lengua del grupo mixeano oaxaqueño, el mixe de Ayutla, y un estudio sobre las lenguas del grupo zoqueano; y *(v)* sobre cinco lenguas mayas, el tojol'ab'al y el q'anjob'al de la rama q'anjob'aliana, el mam de la rama mamiana, y el chol y el tselal del grupo tselalano mayor.

Los fenómenos estudiados abarcan tres grandes temas: las clases léxicas, la posesión, y la integración sintáctica de predicados. En este respecto, tres estudios en el libro giran en torno a problemas analíticos para la categorización de clases léxicas: el de Capistrán-Garza, en el capítulo 2, plantea el problema sobre la distinción léxica entre adjetivos y verbos que codifican conceptos de propiedad en p'orhépecha; el de Vázquez Soto, en el capítulo 3, versa sobre la interfaz léxico-sintáctica entre verbos y sustantivos que designan relaciones de parentesco en cora; y el de Curiel, en el capítulo 6, estudia cómo un verbo en tojol'ab'al ha dado pie a la emergencia de una marca de caso gramatical de dativo. El capítulo de Vázquez

* Structure et Dynamique des Langues (UMR8202), CNRS. Surrey Morphology Group, Universidad de Surrey

** CIESAS-Sureste.

Soto sobre los verbos de parentesco en cora gira asimismo en torno a la dimensión semántica de la posesión, como lo hacen los trabajos de posesión externa de Peralta Ramírez, en el capítulo 4 sobre el nahúatl, y el de Zavala Maldonado, en el capítulo 5 sobre lenguas de la rama zoqueana. El fenómeno complejo de la integración sintáctica de predicados se contempla, por un lado, en las construcciones de verbos seriales de dos lenguas mayas q'anjob'alianas —el tojol'ab'al y el q'anjob'al—, a cargo de Curiel y de Francisco Pascual, respectivamente, en los capítulos 6 y 7. El resto de las contribuciones del libro, de los capítulos 8 al 13 estudian fenómenos de integración sintáctica en estructuras de complementación en lenguas distintas, tales como el mixe de Ayutla, tseltal, chol, mam, q'anjob'al y otomí.

1.1. Clases de palabras: un problema de categorialidad

En el capítulo 2 de este libro, Capistrán-Garza estudia la expresión de los conceptos de propiedad en p'orhépecha o tarasco, lengua aislada hablada en Michoacán. En esta contribución se aborda el problema de la categorización léxica entre adjetivos y verbos siguiendo la metodología sintáctica propuesta por Dixon (1982) y Beck (2002). En este capítulo, Capistrán-Garza demuestra que en el p'orhépecha realmente existen *sólo* dos adjetivos básicos; *sapí*, “pequeño”, y *t'arbé*, “grande”. El resto de las palabras de esta lengua, que en otros estudios sobre la lengua se han clasificado simplemente como adjetivos, en realidad son derivaciones verbales, bien sea como participios pasivos, con el sufijo *-kata*, como un grupo reducido que toma diversos sufijos, *v.gr.* *-pu* y *-sí*, o como palabras que presentan el sufijo *-ri* (o su alomorfo *-ti*). Las palabras con *-ri* o *-ti* cubren la referencia de la mayor parte del espacio semántico de los conceptos de propiedad en p'orhépecha. El capítulo se centra en el análisis de estas palabras. La autora propone que son formas estativas provenientes de verbos incoativos que se pueden emplear como modificadores nominales solamente cuando expresan conceptos de propiedad.

Estas formas estativas se distinguen de otras palabras que funcionan como modificadores nominales en cuanto a que pueden tomar flexión verbal y aparecer en construcciones sintácticas solamente accesibles a los verbos y, por lo tanto, restringidas a adjetivos básicos y otras palabras adjetivales deverbales. El análisis de Capistrán-Garza permite ahora situar el p'orhépecha dentro de una de las tres grandes tendencias tipológicas que observamos en Mesoamérica en relación con el trato léxico de los conceptos de propiedad. Por un lado, están las lenguas adjetivales que tienen una gran clase de adjetivos básicos, como las lenguas mayas, las mixeanas y las lenguas nahuas. Por otro, están las lenguas nominales, las cuales presentan

una clase adjetival con muy pocos adjetivos básicos la que la mayoría de los miembros son adjetivos denominales, como el totonaco (Levy, 2006b). Por último, están las lenguas predicativas. En las lenguas predicativas encontramos dos grandes tipos. En primer lugar, se tienen lenguas en las que la atribución de conceptos de propiedad se codifica en su mayoría o en su totalidad como predicados que son parte de oraciones de relativo, como por ejemplo en cora (Vázquez Soto, 2004), en las lenguas otopames (Palancar, 2006) y en el tlapaneco (Suárez, 1983) en las cuales los predicados que expresan conceptos de propiedad suelen ser verbos estativos (es muy probable que en otras ramas del otomangue suceda lo mismo), o en las zoqueanas, en las que pueden ser tanto verbos como predicados no verbales (*v.gr.* posicionales) que son parte de construcciones de relativo. Un ejemplo de este tipo de lengua se da en 1, en el otomí de San Ildefonso Tultepec, Querétaro.

- (1) OTOMÍ DE SAN ILDEFONSO TULTEPEC
 'ně ga='óť'-a=nö má=n'á
 y 1.PRES.IRR=trazar[3OBJ]-D=DEF.SG otro=uno
 [nó xi=n-tsánt'-i]
 REL PRF=EST-ser_redondo[3OBJ]-L
 'Y dibujo otra (pieza) redonda.' (Palancar, 2009: 16.)
 (Lit. 'Y dibujo otra que es redonda.')

Dentro del otro gran grupo de las lenguas predicativas, en el que se encuentra el p'orhépecha, los conceptos de propiedad con función atributiva se codifican de manera no marcada, a modo de adjetivos, pero son, en su inmensa mayoría, formas verbales estativas que provienen de verbos intransitivos incoativos. Esta estrategia se encuentra en varias formas de zapoteco, como el de Santiago Apóstol (Padilla, 2010) y en el huave (Kim, 2008).

- (2) ZAPOTECO DE SANTIAGO APÓSTOL
 g-uu=m nis n-ald
 POT-tomar=s3SG.RESPE TO agua EST-enfriarse
 'Él tomará agua fría.' (Padilla, 2010: 114.)
- (3) HUAVE
 n-a-lyuy piats s-a-lyup tiol a najngoy
 EST-TEM-ablandarse tortilla s1-TV-remojarse en DET caldo
 'Las tortillas blandas, las remojo en el caldo.' (Kim, 2008: 197.)

La investigación de Capistrán-Garza hace posible que se identifiquen de forma precisa las subclases de palabras empleadas por el p'orhépecha en el contexto de la atribución, de tal manera que se observe claramente el gran peso que los verbos incoativos tienen en la expresión de conceptos de propiedad.

1.2. Clases de palabras y posesión

El problema de clasificación léxica también es tema de estudio de Vázquez Soto, en su capítulo “Predicación de términos de parentesco en cora meseño”, donde se discute el comportamiento gramatical y semántico de los términos de parentesco en el cora meseño, lengua sureña de la familia yutoazteca. A diferencia de lo que pasa en la gran mayoría de las lenguas del mundo, en las que los términos de parentesco (“madre”, “hijo”, “tío”, etcétera.) se codifican únicamente como sustantivos, en cora, al igual que en otras lenguas de Mesoamérica, como el oluteco (mixezoqueano), el náhuatl de Oapan y el huichol (ambas yutoaztecas), estos elementos léxicos son parte de la categoría de verbos transitivos. Esto se debe a que estos términos son semánticamente biactanciales; es decir, establecen una relación entre dos individuos humanos, el poseedor de la relación (*referente*) y el elemento identificado con el poseído (*relatum*). Hasta ahora se han identificado una serie de regiones en Norteamérica, Mesoamérica y Australia donde se encuentran lenguas que de manera más o menos aislada (*v.gr.* no suele ser un comportamiento de familia) tratan estos términos igualmente como verbos. En el cora meseño, existe una alternancia para la codificación de los términos de parentesco entre la construcción verbal, como en 4a, y la construcción no verbal, donde a los términos de parentesco se los trata como nominales, como en 4b (véase el capítulo 3, ejemplos 37a y 37b).

- (4) CORA MESEÑO
- a. néenu mwa-ya'ube'e
 S1SG OP2SG-ser_cuñado_de
 ‘Yo soy tu cuñado.’
 (Lit. ‘Yo te “cuñadeo”.’)
- b. néenu há'a-ya'ube'e
 S1SG POS2-cuñado
 ‘Yo (soy) tu cuñado.’

El fenómeno permanece escasamente documentado y el trabajo de Vázquez Soto representa una importante contribución a su estudio. En primer lugar, es el primer trabajo exhaustivo en cora que muestra conclusiones muy diferentes a los que hicieron Amith y Smith-Stark (1994a; 1994b) para el corachol a partir de datos de fuentes secundarias inadecuadas. En segundo lugar, el estudio está basado en la revisión de los principales parámetros tipológicos identificados en la investigación de verbos de parentesco desde el punto de vista semántico y formal, y por ello el capítulo es un modelo metodológico para realizar un acercamiento riguroso al fenómeno en cualquiera de las lenguas de Mesoamérica con verbos de parentesco aún no documentados. En este capítulo, Vázquez Soto muestra que el cora se comporta de manera muy distinta a los casos reportados en otras lenguas dentro y fuera del área de Mesoamérica, y por ello esta investigación contribuye a enriquecer nuestro entendimiento tipológico del fenómeno. En este sentido, resulta relevante que las dos construcciones ejemplificadas en 4 presentan una distribución libre cuya alternancia no la restringe, como en otras lenguas, la jerarquía de persona o la direccionalidad descendente de los términos. Por ejemplo, mientras que en cahuila (Seiler, 1980; 1982a; 1982b), en el náhuatl de Oapan (Amith y Smith-Stark, 1994a), y en oluteco (Zavala Maldonado, 2006a) se requiere que la expresión “yo soy su hermano” se codifique obligatoriamente como verbo, porque el referente, la primera persona sujeto de la predicación, es más alto en la jerarquía de persona que el *relatum*; por otro lado, en expresiones como “él es mi hermano”, donde el referente, tercera persona sujeto de la predicación, es más bajo en la jerarquía de persona que el *relatum*, se tiende a codificar preferentemente con la construcción nominal. En cora no existe tal restricción, esta como lo demuestra la existencia de usos de la construcción verbal tanto en casos en los que el referente es tercera persona y el *relatum* es primera, como en 5 (véase el capítulo 3, ejemplo 9b), como en casos en los que el referente es primera persona y el *relatum* es tercera, como en 6 (véase el capítulo 3, ejemplo 23d).

- (5) CORA MESEÑO
 me-ne-yá ube'e-te-mwa'a=mu mwá'a=hu'u
 S3PL-OP1sg-ser_cuñado_de-INT-PL.I=S3PL S3PL.P=FUT
 'Ellos van a ser mis cuñados.'

- (6) CORA MESEÑO
 néenu ra-náanaa
 s1SG OP3SG-ser_madre_de
 'Yo soy su madre.'

Al contrario de lo que reporta Evans (2006) para las lenguas del norte australiano, donde sólo se admiten construcciones verbales con “direccionalidad descendente”, es decir, aquellas en las que el pariente mayor es codificado como sujeto y el pariente menor como objeto, el cora admite construcciones con “direccionalidad ascendente”, es decir, casos en lo que el pariente menor es codificado como sujeto, y el pariente mayor como objeto, tal como se ilustra en 7 (véase el capítulo 3, ejemplo 39b).

- (7) CORA MESEÑO
 mwéépe wa'a-yauh-ka'a
 S2SG OP3PL-ser_nieto_de-IMPF
 'Tú eras su nieto (de ellos).'

Otro aspecto en el cual los verbos de parentesco del cora se separan de lo reportado para otras lenguas mesoamericanas se refiere a la semántica. El cora, al igual que las lenguas australianas que presentan este fenómeno, tiene una semántica “Y es pariente de X”, donde el sujeto (Y) refiere al referente (REF) y el objeto (X) refiere al *relatum* (REL) (véase el capítulo 3, ejemplo 39c).

- (8) CORA MESEÑO
 ne-REF_aREL-yáuh
 S1SG-OP3SG-ser_hijo_de
 'Yo soy su hijo.'
 (Lit. 'Yo_{REF} le_{REL} "hijeo".')

En contraste, las construcciones posesivas con términos de parentesco en otras lenguas mesoamericanas en la que se ha reportado el fenómeno, como el náhuatl de Oapan, presentan una semántica “X tiene o trata a Y como pariente”, donde, en oposición al cora, el sujeto (X) esta vez refiere al *relatum* (REL) y el objeto (Y) refiere al referente (REF), tal como se ilustra en 9.

- (9) NÁHUATL DE OAPAN
 tí_{REL}-ne:ch_{REF}-kone:-w
 S2SG-OP1SG-niño-POSD
 'Yo soy tu hijo.' (Amith y Smith-Stark, 1994a: 349.)
 (Lit. 'Tú_{REL} me_{REF} (tienes como) hijo.')

Una tercera diferencia del cora respecto al náhuatl de Oapan se refiere a la categorialidad más verbal de los términos de parentesco. Mientras que en el náhuatl los términos de parentesco retienen rasgos nominales, como la capacidad de aparecer poseídos y la imposibilidad de admitir aspecto, como en 9, en el cora estos términos se tratan como verbos plenos y aceptan la marca de imperfectivo propia de los verbos estativos, tal como se muestra en 10 (véase el capítulo 3, ejemplo 7b).

- (10) CORA MESEÑO
 pe-ni-yáuh-ka'a
 s2SG-OP1SG-ser_hijo_de-IMPF
 'Tú eras mi hijo.' {Txt}

Tanto el capítulo de Capistrán-Garza como el de Vázquez Soto suponen dos acercamientos novedosos al problema de cómo categorizar palabras que se encuentran en frontera entre dos clases léxicas; en el caso del p'orhépecha, entre verbos y adjetivos, y en el la del cora, entre verbos y sustantivos. Los dos casos están circunscritos a ámbitos semánticos específicos, el primero a conceptos de propiedad (Dixon, 1982) y el segundo a la dimensión de la posesión, una amplia y compleja área que tiene otros y variados reflejos en la configuración de la gramática en las lenguas mesoamericanas.

1.3. Posesión externa

La codificación de poseedor y poseído dentro de la estructura argumental de los predicados juega también un papel igualmente importante en la gramática de las construcciones de posesión externa en el náhuatl de San Jerónimo Amanalco, fenómeno que investiga en su capítulo Peralta Ramírez.

Payne y Barshi (1999: 3) definen el caso prototípico de una construcción de posesión externa como “aquellas construcciones en que la relación semántica poseedor-poseído se expresa codificando al poseedor como un argumento central del verbo y en un constituyente separado del que codifica al poseído”. En toda construcción de este tipo, el trato gramatical privilegiado que recibe el poseedor se interpreta como afectación semántica. Tipológicamente, se han identificado cuatro grandes tipos de construcciones de posesión externa: de marcación oblicua, de dativo, de incorporación nominal y de aplicativo (Haspelmath, 1999). Mesoamérica es un área lingüística particularmente rica en este sentido, ya que se encuentran

todos los tipos y existen incluso lenguas que emplean más de un tipo, entre ellas, en este volumen, el náhuatl de Amanalco y las lenguas de la familia mixezoqueana.

El tipo en el que el poseedor se codifica como argumento central y el poseído como oblicuo se ha reportado en Mesoamérica en el zoque de Tecpatán. En 11, el participante que refiere al poseedor se expresa como sujeto absolutivo de la oración por medio de la marca se sujeto *ny-* y del enclítico pronominal *=mij* en correferencia con la segunda persona. Por otro lado, el poseído se expresa como un sintagma oblicuo marcado por la posposición *=pit*. Nótese que el participante que refiere al poseedor permanece registrado en el sintagma oblicuo, además de ser el sujeto absolutivo de la oración (véase el capítulo 5, ejemplo 5b).

(11) ZOQUE DE TECPATÁN

ngyokaʔu mij yōmopit

ny-ko-kaʔ-u=mij

n-yomoʔ=pit

2ABS-APL:MAL-MORIR-COM=PRO2ABS 2PSR-mujer=CON

‘Se te murió tu mujer.’ (Lit. ‘Te moriste con/por tu mujer.’)

El segundo tipo, la estrategia con dativo, en el que el poseedor se codifica como objeto indirecto, es una construcción de posesión externa característica de las lenguas de Europa y que Haspelmath (1999) trata como rasgo de área. Realmente, la construcción es muy rara fuera de Europa y se encuentra sólo esporádicamente en otras lenguas que tienen también dativos. En Mesoamérica, la construcción existe en lenguas otomangues con dativos, por ejemplo en las de la familia otomí (Palancar, 2008, 2009). En 12, del otomí de San Ildefonso Tultepec, se ofrece el contraste entre posesión interna exclusivamente (12a) y posesión externa (12b). En 12a, la frase nominal *má ts’óí*, “mi olla”, es el objeto directo del verbo *ñú ts’í*, “llenar”. Este sintagma nominal está gramaticalmente poseído, y el poseedor, que es la primera persona, no funciona como argumento del verbo. En 12b, en cambio, el poseedor aparece codificado argumentalmente en el complejo verbal con un sufijo de dativo *-k*. Desde un punto de vista semántico, la codificación argumental en 12b dispara una interpretación obligatoria sobre el evento del llenado de la olla (en 12a) en la que el hablante se ve afectado por el resultado de la acción, bien sea en su beneficio o en su perjuicio. Esta implicación no se extrae de 12a.

(12) OTOMÍ DE SAN ILDEFONSO TULTEPEC

a. bi=ñúʔs-a=má ts’óí

3.PSD=llenar.TA[3OBJ]-D=1POS olla

‘Llenó mi olla.’

- b. bi=ñü's-k-a=má ts'óí
 3.PSD=llenar.TA-1DAT-D=1POS olla
 'Me llenó mi olla.'

En la construcción de posesión externa en las lenguas otomías, como en muchas otras lenguas de Mesoamérica, el sustantivo poseído aparece siempre marcado por la marca de poseedor, es decir, el poseedor semántico se manifiesta en dos lugares, como dativo de primera persona (el sufijo *-k* en el verbo) (poseedor externo) y como modificador del sustantivo poseído (el morfema *má*).

El tercer tipo presente en Mesoamérica es la construcción de posesión externa por medio de la incorporación nominal que la tipología de Mithun (1984) denominó incorporación por "manipulación de caso". En esta construcción, el sustantivo que expresa lo poseído se incorpora en el verbo dejando vacante una posición argumental, la cual es ocupada por el poseedor sin provocar ningún cambio en la valencia verbal, es decir, se presenta un reacomodo en la estructura argumental sin alterar la valencia original. Este tipo se ha reportado en muchas lenguas de México con incorporación nominal, por ejemplo en oluteco (Zavala Maldonado, 1999), zoque de San Miguel Chimalapa (Johnson, 2000), soteapaneco (de Jong Boudreault, 2009), p'orhépecha, restringido a partes (Capistrán-Garza, 2006), y totonaco, a modo de merónimos (Levy, 1999: 325). El ejemplo 13 proviene del zoque de San Miguel Chimalapa y el 14 del soteapaneco (también conocido como popoluca de la Sierra).

- (13) ZOQUE DE SAN MIGUEL CHIMALAPA
 də=kopak-toy-wə
 1ABS=cabeza-doler-COM
 'Me duele la cabeza.' (Johnson, 2000: 277.)

- (14) SOTEAPANECO
 mi=puy-ku.wi?ks-w
 2ABS=pie-doblarse-COM
 'Te torciste el pie.' (de Jong Boudreault, 2009: 666.)

Esta construcción se encuentra también en náhuatl de Amanalco, ilustrada en 15b, la cual se puede contrastar con la construcción de posesión interna en 15a (véase el capítulo 4, ejemplo 19).

- (15) NÁHUATL DE AMANALCO
- a. ni-k-paʔ-pa:k mo-ma:n
 s1SG-OP3SG-RED-lavar.PRF POS2SG-mano
 ‘Lavé tu mano.’
- b. ni-mits-ma:paʔ-pa:k
 s1SG-OP2SG-mano-RED-lavar.PRF
 ‘Te lavé las manos.’
 (Lit. ‘Te-mano-lavé.’)

Además de los tres tipos ya ejemplificados, es relativamente frecuente en lenguas de Mesoamérica emplear un cuarto tipo conocido como posesión externa con aplicativo, el cual se ha reportado, por ejemplo, en lenguas de la rama tseltalana de la familia maya, como el tzotzil (Aissen, 1987), tseltal (Polian, 2006), chol (Vázquez Álvarez, 2002), chontal (Osorio May, 2005), o en el huasteco, también maya (Constable, 1989). También se ha documentado en lenguas mixezoques, como el oluteco (Zavala Maldonado, 1999), soteapaneco (Marlett, 1986; de Jong Boudreault, 2009), zoque de San Miguel Chimalapa (Johnson, 2000), así como en lenguas totonacas, como el totonaco de Papantla (Levy, 2002) y el tepehua de Huehuetla (Smythe Kung, 2004). La encontramos asimismo en p’orhépecha (Capistrán-Garza, 2006). En la construcción con aplicativo el poseedor de la frase nominal pasa a ocupar la posición de objeto primario en la cláusula. Peralta Ramírez muestra en su trabajo que el náhuatl también la posee, tal como se ve en 16, en el contraste del par 16a – 16b. (véase el capítulo 4, ejemplo. 28).

- (16) NÁHUATL DE AMANALCO
- a. ti-k-ihta-k no-kone:-w
 s2SG-OP3SG-ver-PRF POS1SG-niño-SG.PSDO
 ‘Viste a mi hijo.’
- b. ti-ne:ch-ihtli-li no-kone:-w
 s2SG-OP1SG-ver-APL.PRF POS1SG-niño-SG.PSDO
 ‘Me cuidaste a mi hijo.’
 (Lit. ‘Me viste a mi hijo.’)

En la construcción de posesión externa por aplicativo, las lenguas mayas, totonacas y mixezoques no solamente marcan al poseedor como objeto primario sino que lo preservan en el interior del sintagma nominal como dependiente del

núcleo (fenómeno al que apuntábamos antes, en los ejemplos 11 y 12b). En el siguiente ejemplo del oluteco, el poseedor semántico de segunda persona se expresa como proclítico antes del sustantivo *tuku*, “ropa”, y como marca de absolutivo de segunda persona en el verbo *ko:puj-aʔx*, “lavar-APL”, en función de objeto primario.

Este comportamiento lo muestra también el náhuatl, tal como se observa en 15b.

(17) OLUTECO

min=tuku pün mi=ko:puj-aʔx-anüpa=aʔ
 2PSR=ropa quién 2ABS=lavar-APL-IRR.INV=ANIM
 ‘¿Quién te va a lavar tu ropa?’ (Zavala Maldonado, 1999: 256.)

Una lengua de Mesoamérica, posiblemente la única, en la que se expresa el poseedor en la estructura de posesión externa, con aplicativo únicamente como objeto primario y no como dependiente del poseído, es el p’orhépecha. Esto se ilustra en el contraste entre la estructura de posesión interna en 18a, en la cual el poseedor aparece marcado como un dependiente en genitivo, con la estructura de posesión externa en 18b, en la que aparece con marca de objeto.

(18) P’ORHÉPECHA

- a. xuká-s-Ø-ti María-eri k’waníntikwa-ni
 ponerse-PRF-PRES-3IND María-GEN rebozo-OBJT
 ‘Se puso el rebozo de María.’ (Capistrán-Garza, 2006: 295.)
- b. xuká-ku-s-Ø-ti k’waníntikwa-ni María-ni
 ponerse-APL3-PRF-PRES-3IND rebozo-OBJT María-OBJT
 ‘Se puso el rebozo de María.’ (Capistrán-Garza, 2006: 294.)
 (El rebozo era de María y le afectó a ella la acción.)

Aparte de las construcciones de incorporación y de aplicativo, el náhuatl de Amanalco tiene una tercera construcción de posesión externa que surge de la combinación de ambas estrategias, esta vez con verbos resultado del proceso de incorporación por composición que baja la valencia verbal (Mithun, 1984) más el aplicativo benefactivo *-li(a)*, tal como se ilustra en el contraste del par en 19 (véase el capítulo 4, ejemplo. 5).

(19) NÁHUATL DE AMANALCO

- a. ni-k-im-pia mo-chka-wa:n
 S1SG-OP3-PL-cuidar POS2SG-borrego-PL.PSDO
 ‘Cuido tus borregos.’

- b. ni-mits-ichka-pia-li
 S1SG-OP2SG-borrego-cuidar-APL.PRF
 ‘Te cuidé tus borregos.’ (Fui tu pastor por mucho tiempo.)

La contribución de Peralta Ramírez identifica los rasgos estructurales y semánticos de los tres tipos de construcción de posesión externa presentes en el náhuatl e investiga las restricciones de concurrencia con los sustantivos que pueden ser posibles dentro de esos tres tipos mayores. La construcción de incorporación ocurre con sustantivos inalienables de partes del cuerpo e inherentes (*i-yol*, “su semilla”; *i-ma:n*, “su mano”; *i-yes-yo*, “su sangre”; *i-nelwa-yo*, “su tendón”); la de aplicativo con sustantivos términos de parentesco (*i-kniw*, “su hermano”; *i-we:litiw*, “su cuñada”); mientras que la que combina ambas estructuras aparece con sustantivos alienables (*chichi*, “perro”; *cha:ntli*, “casa”; etcétera).

Igualmente rica en construcciones de posesión externa es la lengua zoque de Tecpatán que se presenta en el capítulo de Zavala Maldonado. Esta lengua de la familia mixezoque tiene, al igual que el náhuatl de Amanalco, tres construcciones de posesión externa: dos son de aplicativo y una tercera que combina la estrategia de incorporación con la de aplicativo. Las dos construcciones de aplicativo presentan idiosincrasias significativas que las hacen comportarse de manera muy distinta a sus contrapartes en otras lenguas del área, e incluso de la misma rama zoqueana. En la primera construcción de aplicativo se emplea *ko-* y aunque promueve al poseedor como argumento tiene dos características muy relevantes: (i) el aplicativo *no* altera la valencia verbal y (ii) el sintagma nominal que codifica lo poseído se demueve siempre a una posición de oblicuo, codificado por una posposición =*pit* que indica causa o instrumento. Esto se ilustra en los ejemplos 20 y 21. En 20a, se muestra un verbo intransitivo que en 20b aparece con el aplicativo *ko-*. El aplicativo hace posible la promoción del poseedor a sujeto sin modificar la valencia. Lo mismo sucede en 21a, esta vez con un verbo transitivo, si se compara con 21b. En este caso, el poseedor se promueve a objeto primario (en forma de pronombre reflexivo).

- (20) ZOQUE DE TECPATÁN
 a. Ø-kaʔ-u n-yomoʔ
 3ABS-morir-COM 2POS-mujer
 ‘Se murió tu esposa.’

- b. ny-ko-kaʔ-u=mij n-yomoʔ=pit
 2ABS-APL-morir-COM=PRO2ABS 2POS-mujer=INSTR
 ‘Se te murió tu esposa.’
 (Lit. ‘Te moriste con/por tu esposa.’)

(21) ZOQUE DE TECPATÁN

- a. n-jak-pəʔ-u=ʔəʂ n-kəʔchus
 1ERG-cortar-TOTALMENTE-COM=PRO1ERG 1POS-uña
 ‘Corté mis uñas.’
- b. n-ko-jak-u=ʔəʂ n-win n-kəʔchus=pit
 1ERG-APL-cortar-COM=PRO1ERG 1POS-REF 1POS-uña=INSTR
 ‘Me corté las uñas.’
 (Lit. ‘Me corté yo mismo con/por mis uñas.’)

En la segunda construcción se tiene el aplicativo benefactivo *-jay*, que, empleado como estructura de posesión externa, tiene un comportamiento anómalo. En primer lugar, en la construcción, el sintagma nominal que codifica al poseedor mantiene su estatus de modificador genitivo del sintagma nominal que expresa lo poseído, por lo que no aparece *strictu sensu* como un constituyente independiente. Como consecuencia de ello, la construcción *no* promueve al poseedor como objeto primario —función que mantiene el sintagma nominal poseído; nótese que en este caso tampoco está demovido como en 20b y 21b— pero *sí* le da rango de objeto secundario, ya que el verbo puede recibir marcas de concordancia de plural a las que solamente son candidatos los argumentos centrales —sean sujeto u objeto—. Esto se puede ver en 22.

(22) ZOQUE DE TECPATÁN

- n-paʔt-jay-yaj-u=ʔəʂ ((teʔ pən=ʔis=taʔ) y-kəʔchus)
 1ERG-encontrar-APL-3PL-COM=PRO1ERG DET hombre=GEN=PL 3POS-uña
 ‘Les encontré la uña a los señores.’

Por último, la tercera construcción de posesión externa en el zoque de Tecpatán se emplea como en el náhuatl de Amanalco (véase el capítulo 4), un aplicativo (*ko-*) más la incorporación de un sustantivo poseído que refiere a partes inalienables de un todo, tal como se ilustra en el par del ejemplo 23.

- (23) ZOQUE DE TECPATÁN
- a. Ø-putz-u ((teʔ kuytəm=ʔis) y-puj)
 3ABS-pudrir-COM DET aguacate=GEN 3POS-semilla
 ‘Se pudrió la semilla del aguacate.’
- b. Ø-puj-ko-putz-u teʔ kuytəm
 3ABS-semilla-APL-pudrir-COM DET aguacate
 ‘Al aguacate se le pudrió su semilla.’

Con estos nuevos datos se confirma el hecho de que el área lingüística de Mesoamérica es particularmente rica en manifestaciones gramaticales para la codificación de PE. Los datos del zoque de Tecpatán indican adicionalmente, que dentro del ámbito de la posesión externa, una lengua puede emplear un aplicativo sin hacer que el argumento poseedor pase a ser objeto primario. Los nuevos datos sugieren también que cuando una lengua tiene más de una construcción de posesión externa como alternativa, la distribución de estas construcciones es muy probable que esté regulada por criterios léxicos de alienabilidad e inalienabilidad. Podemos igualmente proponer la siguiente generalización de alcance tipológico al menos para el área mesoamericana: “cuando una lengua tiene una construcción de incorporación en construcción de posesión externa, tiene también una de aplicativo, pero no viceversa”.

Las lenguas de Mesoamérica que hacen uso de la construcción de PE con aplicativo recurren también a esta construcción para codificar al recipiente como objeto primario de los verbos bitransitivos derivados por el aplicativo. Nótese el siguiente par de ejemplos del oluteco: 24a es una construcción bitransitiva en la cual el aplicativo introduce al recipiente, mientras que 24b es una construcción bitransitiva en la que el aplicativo hace posible que el poseedor del tema tome la posición de objeto primario

- (24) OLUTECO
- a. jeʔ tye:mpo tan=joʔn-küx-aʔx-i-y min=piyu
 aquel tiempo 1ABS=robar-3PL-APL-COM.DEP-INV.DEP 2PSR=pollo
 ‘En aquellos tiempos ellos me robaron tus pollos.’
- b. ta=tun-aʔx-ü-w=ak tan=kama-nak
 1ABS=hacer-APL-INV-COM=ANIM 1PSR=milpa-DIM
 ‘Él me preparó mi milpita.’

Además de la estrategia de aplicativo para la codificación del receptor o destinatario como objeto primario, las lenguas del área hacen uso de otras estructuras para introducir esta función gramatical como parte de la estructura argumental de la cláusula (López Nicolás, 2009). Entre las estrategias comunes en el área, están la del tratamiento del receptor como objeto simétrico, como poseedor proléptico, la de reacomodo de argumentos por incorporación, y la marcación del receptor como objeto asimétrico por medio de caso oblicuo o por medio de objeto seriales, estrategia de la que el tojol'ab'al reclutó la marca de caso dativo, objeto de estudio del capítulo de Curiel.

1.4. Verbos seriales

Hasta muy recientemente, Mesoamérica fue un área donde el estudio sobre verbos seriales había sido ignorado, sobre todo si lo comparamos con la investigación exhaustiva que se ha hecho de este fenómeno en otras áreas del mundo, principalmente en el sureste asiático; Papua, Nueva Guinea, y África occidental. Con la investigación del fenómeno en estas tres regiones se ha podido identificar a los verbos seriales como secuencias de dos o más verbos que actúan juntos como un solo predicado, que no tienen marcas expresas de coordinación, subordinación o de dependencia sintáctica o discursiva, y que comparten al menos un argumento (Aikhenvald, 2006a). El siguiente ejemplo del tojol'ab'al es un caso de una construcción serial debido a que: 1) entre las secuencias del verbo “decir” y “oír” no hay dependencia sintáctica; 2) ambos verbos comparten el objeto lógico, y 3) los dos verbos forman una unidad estructural que codifica un solo evento (véase el capítulo 6, ejemplo 9).

(25) TOJOL'AB'AL

y-al-a-Ø=ni k-ab' ke mi=ni jas=ni wa x-'ek'-Ø
 A3-decir-ET-B3=ENF A1-oír COMP NEG=ENF qué=ENF INC INC-pasar-B3
 s-k'ujol=a

A3-corazón=DIST

‘Claro que me dijo que nada le pasaba por la mente.’

(Lit. ‘Sí lo dijo y yo lo escuché que nada le pasaba por la mente.’)

En Mesoamérica sabemos que lenguas con “genio” polisintético hacen uso de estructuras seriales de tipo nuclear en las que las secuencias entre los verbos son

contiguas y llegan a formar palabras flanqueadas por las marcas de argumentos y TAM. Los ejemplos siguientes son del oluteco, (en 24) y zoque de San Miguel Chimalapa, (en 25), ambas lenguas mixezoqueanas.

(26) OLUTECO

maku ʔu:ra tuk=xü ʔi=xej-pük-i
 diez hora uno=CIT A3(ABS)=exhalar-agarrar-INCD
 ‘Uno descansa a las diez.’ (Zavala Maldonado, 2006: 273.)

(27) ZOQUE DE SAN MIGUEL CHIMALAPA

bi nuʔ ʔəy=waš-num-təʔ-pa bi cinuʔ ʔəy=tək
 DET perro 3ERG=morder-robar-querer-INC DET abeja 3PSR=casa
 ‘El perro quería husmear y morder el panal de abejas.’ (Johnson, 2000: 218.)

La investigación en diversas áreas del mundo y también en Mesoamérica ha dado cuenta de la manera en que lenguas que hacen uso de la construcción con verbos seriales han reclutado núcleos verbales que son parte de las estructuras seriales para volverlos morfemas gramaticales. El tipo de estructuras seriales donde se observan estos procesos de gramaticalización se conoce como “construcciones multiverbales asimétricas” (Aikhenvald, 2006a) debido a que uno de sus miembros proviene de una clase cerrada de raíces o bases verbales que da lugar a morfemas gramaticalizados que expresan marcas aspectuales, marcas de concordancia, marcas de caso, direccionales, marcadores de cambio de valencia, complementizadores y subordinadores. Varios estudios han demostrado que los miembros de las secuencias seriales que se reanalizan como morfemas gramaticales son las raíces verbales con alto grado de frecuencias (Durie, 1997; Givón, 1991; Foley y Olson, 1985; Lord, 1993; entre otros). Curiel, en el capítulo titulado “Construcciones gramaticalizadas de verbos seriales en tojol’ab’al”, muestra que construcciones como la de 28 es una estructura serial y que el segundo verbo de la secuencia v1+v2 se ha gramaticalizado como una marca de caso dativo (véase el capítulo 6, ejemplo 8).

(28) TOJOL’AB’AL

mok wa-b’at-Ø y-i s-k’ab’ ja=’alats=i
 NEG.IRR A2-pisar-B3 A3-tomar A3-mano DET=bebé=TOP
 ‘No le pises la mano a la bebé.’
 (Lit. ‘No la pises ella-tomar su mano de la bebé.’) (Ttxt)

El autor presenta evidencia fonológica, morfológica, sintáctica y semántica que demuestra el cambio de estatus gramatical de verbo a marca de caso. La evidencia crucial que demuestra este cambio de estatus del v2 se encuentra en su estructura argumental defectiva. En su función de marca de dativo, el verbo original *i'*, “traer”, no marca al argumento absoluto (en 29b), a pesar de que en su estructura argumental original la marca de paciente es obligatoria (véase el capítulo 6, ejemplo 31).

(29) TOJOL'AB'AL

a. mok wa-chon-on y-i'
 NEG.IRR B2-vender-B1 A3-DAT
 ‘No me vendas a él.’

b. * mok wa-chon-on y-i'-won
 NEG.IRR B2-vender-B1 A3-DAT-B1
 Lectura buscada: ‘No me vendas y que él me lleve.’

El proceso de gramaticalización de donde se recluta la marca de dativo es único entre las lenguas de la familia. Otros miembros de la familia han reanalizado construcciones asimétricas similares dando como resultado morfemas direccionales que son un paradigma cerrado de elementos gramaticales que codifican trayectoria y/o aspecto. Este proceso ha sido investigado en varias lenguas de la familia (Aissen, 2006; Zavala Maldonado, 1992; England, 1976; entre otros; véase también el capítulo 7).

Por otro lado, en el capítulo “La construcción resultativa de verbo serial en q'anjob'al”. Francisco Pascual estudia otra construcción de verbo serial donde el v1 funciona como predicado principal y el v2 funciona como predicado secundario con semántica resultativa. El ejemplo 30 ilustra esta construcción donde el segundo predicado *kay*, “desaparecer”, (véase el capítulo 7, ejemplo 9) copredica sobre el participante que tiene el papel de paciente del predicado principal. Debido a que la predicación expresa un resultado del evento causativo *b'eq*, “soltar”, se le conoce como predicado *resultativo*.

(30) Q'ANJOB'AL

P1°	P2°res			
<u>max-Ø-s-b'eq</u>	<i>k'ay</i>	naq	Wel	s-q'oxq'om
COM-B3-A3-soltar	desaparecer	CLF	Manuel	A3-sombrero

‘Manuel perdió su sombrero.’

A diferencia de la construcción de predicación secundaria depictiva, el la que el predicado secundario precede al predicado primario, en la construcción resultativa, los predicados toman un orden que es icónico con el orden en que suceden los eventos en el mundo; el verbo que codifica la causa antecede al verbo que codifica el resultado. Este tipo de construcción de predicación secundaria que recurre a la serialización es poco conocida en Mesoamérica (para el caso del oluteco, véase Zavala Maldonado, 2006), y no se había discutido respecto de ninguna otra lengua maya. En el capítulo, Francisco Pascual presenta los rasgos formales que definen a la construcción resultativa como serial: *a)* no se permiten marcas de coordinación ni de subordinación que separen a los dos verbos, *b)* no se permite FFNN que intercedan entre los dos verbos, *c)* hay una sola marca de aspecto y valor polar con alcance hacia los dos verbos, y *d)* los dos verbos comparten un argumento. Las estructuras resultativas del q'anjob'al constan de dos núcleos independientes que forman una estructura integrada formal y semánticamente y que desde el punto de vista translingüístico ilustra estructuras a medio camino entre las formas analíticas y las formas sintéticas que propias de las lenguas polisintéticas.

1.5. De las formas no finitas y la complementación

Los fenómenos morfosintácticos más complejos, en el proceso de integración de predicados en cláusulas complejas, se observan mejor en el ámbito de la complementación clausal; en éstas un verbo matriz (*v.gr.* “querer”, “acabar”, etcétera) presenta un argumento eventivo (*v.gr.* la acción que se quiere hacer, acabar, etcétera) que se codifica con un predicado normalmente verbal. Este predicado, argumento de la matriz, aparece a su vez en diferentes tipos de constructos sintácticos llamados “cláusulas de complemento”, cuya estructura oscila en grado de complejidad desde las más complejas, o cláusulas plenas (“quiero que tú vinieras a casa”), a las de sintaxis más reducida (“quiero venir”).

Para el estudio de fenómenos de este tipo en lenguas poco documentadas cuya descripción cubre en el mejor de los casos aspectos morfológicos fundamentalmente, es necesario explorar primero de forma metódica todas las posibilidades estructurales que presentan los tipos de cláusulas de complemento en la lengua en cuestión para poder estar en posición de entrar en más detalles en las características peculiares de cada tipo. El capítulo 8, a cargo de Romero Méndez es una muestra precisamente de este proceder. En este trabajo, Romero Méndez hace una indagación de todos y cada uno de los tipos de complementos clausales que encontramos en la lengua mixe hablada en Ayutla, Oaxaca.

Para ello, Romero Méndez emplea la estricta metodología de contraste sintáctico desarrollada por Judith Aissen, inspirada a partir de los lineamientos tipológicos Cristóforo (2003), y que la misma Aissen impartió en dos talleres de complementación en lenguas de Mesoamérica organizados respectivamente por el CIESAS-Sureste, en San Cristóbal de las Casas, Chiapas, y por la Asociación Oxlajuuj Keej Maya' Ajtz'iib' (OKMA), en la Antigua, Guatemala, en colaboración con el proyecto SEP-Conacyt-47475 de la Universidad Autónoma de Querétaro.

En estos talleres participaron de lingüistas que describieron y documentaron diferentes lenguas de Mesoamérica y que poseían en el momento una sólida formación en varios aspectos formales de las lenguas objeto de estudio; estos lingüistas contaban con una amplia base de datos primarios de texto para realizar apropiadamente los respectivos análisis. De los diferentes aspectos trabajados por los autores en esos talleres, surgen asimismo las contribuciones de los capítulos restantes de este libro, a cargo de Mateo Toledo, England, Vázquez Álvarez, Polian y Palancar.

En la lengua mixe de Ayutla, Romero Méndez propone la existencia de hasta cinco tipos distintos de complementos clausales, organizados bajo dos grandes rubros: en estructuras subordinadas, que agrupan los Tipos 1 y 2, en los que hay independencia de marcación morfosintáctica entre el predicado matriz y el complemento, y estructuras co-dependientes, con los Tipos 3-5, en los que se presenta una mayor integración entre los predicados implicados. El Tipo 1, ilustrado en 31, corresponde a un complemento clausal con estructura sintáctica plena: entre otras características, lo introduce un complementante; puede aparecer en polaridad negativa y puede tener constituyentes en posiciones informativas prominentes (véase el capítulo 8, ejemplo 16).

(31) MIXE DE AYUTLA

te'n=ējts n-mētoo-y [kuu jajp=ja'a tāt y-t<ää>n-y]
 así=1PRO 1A-OÍ-DEP COMP DEIC-DEM3 señor 3S-<COM>quedarse-DEP
 'Escuché que allá se quedó el señor.'

Un segundo tipo de subordinación son los complementos no finitos, como en 30, donde se puede observar que, en contraste con 32, el predicado no finito precede al predicado matriz (véase el capítulo 8, ejemplo 24).

(32) MIXE DE AYUTLA

mejts [uuk-p] m-jënmay-py
 2PRO.SG beber-NF 2A-pensar-INDEP.TR
 'Piensas tomar.'

Los predicados no finitos en mixe, como *uukp*, son exclusivamente intransitivos. Éste es un rasgo que se encuentra asimismo en lenguas mayas, tal como apuntaremos más adelante.

En el tipo de complemento que Romero Méndez trata como Tipo 3, el predicado dependiente aparece después del verbo, pero restringido a modo irrealis o desiderativo. Cuando es transitivo, como en 33, su FN objeto aparece en posición inicial de la cláusula matriz. Este desplazamiento sintáctico se interpreta como evidencia de la existencia de un cierto grado de integración clausal (véase el capítulo 8, ejemplo 29).

(33) MIXE DE AYUTLA

<i>akxäjk</i>	nojty	Karlos	y-tsok-py	t-jě'kx-t
chayote	IMP	Carlos	3A-querer-INDEP.TR	3A-comer-IRR.DEP

‘Carlos quería comer chayote.’

Por otro lado, en las lenguas mixeanas, el comportamiento de las marcas explícitas de dependencia en los verbos se emplea como criterio morfosintáctico crucial para distinguir tipos de complemento. Estas marcas suelen correlacionarse con marcas de aspecto y paradigmas de personas que en lenguas innovadoras han dado pie, además, a sistemas de alineamiento complejos (Santiago Martínez, 2008). Estas marcas de dependencia —el término no se debe entender como subordinación— reflejan la sensibilidad del verbo a la estructura sintáctica de ciertos constituyentes dentro de su ámbito clausal. Por ejemplo, el verbo aparece marcado como “independiente” cuando lo preceden constituyentes argumentales, como en 34a, pero como “dependiente” cuando aparece un adverbio en posición inicial, como en 34b (véase el capítulo 8, ejemplo 7).

(34) MIXE DE AYUTLA

- a. Karlos yäjks y-kay-py
 Carlos memelas 3A-comer-INDEP.TR
 ‘Carlos come memelas.’
- b. xem Karlos yäjks t-kay-y
 DEIC Carlos memelas 3A-comer-DEP
 ‘Carlos come memelas ahí.’

En este sentido, la estructura no finita del ejemplo 33 parece comportarse como el argumento objeto en 34a, ya que el verbo matriz se codifica como independiente.

En la estructura del ejemplo 33 el predicado del complemento es finito y aparece en dependiente (presumiblemente) porque se tiene el marcador de imperfecto *nojty*, un antiguo adverbio gramaticalizado, parecido al comportamiento del déictico en 34b. Este comportamiento sugiere que existe un cierto grado de integración entre los predicados matriz y complemento como una unidad predicativa, ya que el segundo está siendo sensible a los parámetros a los que tendría que atender por sí solo el predicado matriz.

Los dos últimos tipos de complemento se ejemplifican en 35 y 36. En estos tipos, el orden del dependiente es el inverso de 33 y recuerdan más al de 32. En el tipo en 35, se puede explicitar el objeto del complemento, pero no así en el de 36, donde es necesariamente un predicado intransitivo, al que Romero Méndez denomina “mínimo”, y que aparece en posición dislocada; contrástese al respecto el posicionamiento del adverbio de tiempo *ojts* en 36 con *nojty* en 33 (véase el capítulo 8, ejemplos 34 y 41).

(35) MIXE DE AYUTLA

ja'a kipy-u'unk y-mäts-py y-kaxë'ëk-y
 DEM3 palo-DIM 3A-agarrar-INDEP.TR 3S-parecer-DEP
 'Se ve que está agarrando el palo.' {Txt}

(36) MIXE DE AYUTLA

(**ojts*) Ø-tsoon-p *ojts* x-jënmay-y
 (PRET) 3S-irse-INDEP PRET 2A-pensar-DEP
 'Pensaste que se fue.'

El capítulo de Romero Méndez es una contribución metodológica importante que abunda en detalles sobre los diferentes contrastes formales que presenta el espectro de la complementación clausal en una lengua mixe. Cada uno de los tipos se acompaña asimismo de un pequeño listado de índole orientativo de los verbos matrices que los subcategorizan en su estructura léxica, pero la misma existencia de tantos tipos en una misma lengua plantea la pregunta de por qué se tiene esa riqueza y cuáles son el reducto de estructuras que históricamente pudieron haber sido más productivas. El fenómeno sugiere principios de integración regulados diacrónicamente por focalización y grados de finitud que dan pie a complejas estructuras de integración, tal como nos las encontramos hoy en la familia.

Aspectos del vasto tema de la finitud estructural en los complementos se tratan en detalle en el capítulo 9, de Mateo Toledo, sobre el q'anjob'al, así como en el acercamiento morfológico a la finitud, en el capítulo 10, de England, sobre el mam;

pero también de forma tangencial en el capítulo 11, de Vázquez Álvarez, sobre el chol, y en el capítulo 12, de Polian, sobre el tselal y sus dialectos. El interés en los fenómenos estudiados así como la constitución de los trabajos resultaron coincidentemente paralelos con el interés de Nikolaeva (2007a) de recoger y contrastar de manera ejemplar lo que en la lingüística, moderna se sabe hasta la fecha del fenómeno de la finitud. El concepto de finitud se ha manejado de manera diversa en lingüística dependiendo de la escuela en vigor en el momento; los acercamientos tradicionales favorecen el énfasis en lo morfológico, mientras que en los formalistas se prefiere el énfasis sintáctico. En realidad, y tal como sugiere Nikolaeva (2007b), los rasgos que se toman en cuenta para definir el grado de finitud a veces son tan específicos de una lengua que resulta difícil establecer una definición de la finitud que resulte válida interlingüísticamente, y parece sensato establecer que los criterios de definición, cualesquiera que sean, para lenguas particulares, han de presentar una correlación entre una reducción y/o restricción de marcación en los rasgos gramaticales en el dominio morfológico con una reducción y/o restricción en comportamiento de complejidad sintáctica.

Las lenguas mayas son particularmente interesantes, en este sentido, para un estudio teórico de la finitud; y si bien nos parece una lástima que en Nikolaeva (2007a) no se encuentre ningún estudio sobre estas lenguas, tenemos la satisfacción de haber podido compilar una serie de trabajos que esperamos que contribuyan, por su calidad e innovación, como valioso “grano para el molino de la teoría lingüística”, por citar las palabras de Nora England (1983: 4).

En el capítulo 9, Mateo Toledo presenta la distribución de dos tipos de estructuras clausales con finitud reducida en la lengua maya q'anjob'al. Tal como también sucede en el mam, como se muestra en el capítulo de England, estas estructuras no están restringidas a la función de complementos, ya que la morfología que las caracteriza aparece asimismo en construcciones adverbiales y en aquellas con predicación secundaria depictiva y/o resultativa, aunque a veces su comportamiento es diferente. A pesar de ello, el ámbito sintáctico de la complementación se mantiene con mucho como el área más fructífera en la cual se puede encontrarlas. Las dos estructuras se ejemplifican en 37 y 38, respectivamente, y se tratan morfológicamente como no finitas por el simple hecho de que no portan marcas de aspecto (véase el capítulo 9, ejemplos 1 y 2).

- (37) Q'ANJOB'AL
 x-Ø-y-il ix [ha-way-i]
 COM-B3SG-A3SG-ver CLF:mujer A2SG-dormir-VI
 'Ella te vio durmiendo.'

- (38) Q'ANJOB'AL
 x-ach y-*uqtej-toq* ix [man-*oj* jos]
 COM-B2SG A3SG-perseguir-DIR CLF:mujer comprar-INF huevo
 'Ella te mandó a comprar huevos.'

Si bien ninguna de las dos estructuras tiene marca de aspecto, la finitud de estas estructuras es también de índole sintáctica ya que, como ilustra Mateo Toledo, ninguna puede recibir ni negación ni elementos prepuestos al verbo en posiciones de prominencia informativa. El contraste morfológico más significativo entre el tipo que se ejemplifica en 37 y el que se muestra en 38, radica en que, en 37, se tienen marcas de concordancia de persona, mientras que el tipo, en 38, realmente equivale estructuralmente a los infinitivos de lenguas indoeuropeas, de ahí la glosa “infinitivo”, *v.gr.* sin flexión de aspecto ni de persona, muy parecida a la forma no finita, en mixe, del ejemplo 32. Si bien el predicado recibe marcación de persona, el tipo, en 37, presenta sin embargo un alineamiento nominativo-acusativo. Nótese al respecto, por ejemplo, que si el patrón fuera ergativo, se esperaría una marca de absolutivo o juego B en el verbo intransitivo *wayi*, “dormir”, en 37, en vez del ergativo o juego A, que es, sin embargo, la concordancia que se encuentra.

Otra diferencia entre ambas estructuras radica en que el tipo de complemento en 37 puede tener un complemento transitivo, como se muestra en 39 (véase el capítulo 9, ejemplo 27), mientras que en la construcción en 38 se tiene solamente un verbo intransitivo. El nominal *jos*, “huevos”, en 38, está incorporado,¹ y esto se observa en el hecho de que no puede recibir determinantes, tal como se ilustra en la agramaticalidad de 40 (véase el capítulo 9, ejemplo 58b). Es decir, el constituyente objeto no puede constituirse en una frase determinante (en inglés, *DP*) porque no puede recibir caso sintáctico de objeto del predicado no finito.

- (39) Q'ANJOB'AL
 max-Ø w-il [*hach* s-sah-on-toq ha-mam b'ay xal]
 COM-B3SG A1SG-ver B2SG A3SG-regalar-MD-DIR A2SG-padre a CLF
 'Yo vi que tu padre te regaló a la señora.'

- (40) Q'ANJOB'AL
 * q-Ø-j-*uqtej-toq* ix [jun waj-*oj* sakate ti]
 POT-B3SG-A1PL-mandar-DIR CLF:mujer INDEF juntar-INF zacate DEM
 Lectura buscada: 'La mandaremos a juntar zacate.'

¹ La incorporación es de un tipo especial que en Gerdtz (2001) se denomina *noun stripping* (“desvestimiento nominal”).

En el capítulo 11, Vázquez Álvarez muestra que el chol tiene dos tipos de cláusulas sin aspecto a las que se denomina “no finitas”. Un primer tipo, ilustrado en 41, es equivalente al tipo estructural de forma con persona ejemplificado en 37 o 39, en q’anjob’al (véase el capítulo 11, ejemplo 50).

(41) CHOL

tyi k-il-ä-Ø [a-wuts’-oñ-el]
 PRFV A1-ver-VTD-B3 A2-lavar.ropa-AP-NF
 ‘Vi que lavabas ropa.’
 (Lit. ‘Vi tu lavar ropa.’)

Del segundo tipo, que se diferencia del de 41 en que no muestra concordancia de persona, se tienen dos subtipos, ilustrados en 42, respectivamente (véase el capítulo 11, ejemplos 1 y 34a).

(42) CHOL

- a. a’bi tä=x=tyo k-ñop-o-Ø-loñ [tyäl-el]
 ayer PRFV=ya=todavía A1-intentar-VT-B3-PL venir-NF
 ‘Ayer todavía intentamos venir.’ {T_{xt}}
- b. ya’=bi kaje muk’-Ø-ob [tyi wäy-el]
 allí=REP empezar IMPF-B3-PL P/SUB dormir-NF
 ‘Dicen que allí estaban empezando a dormir.’ {T_{xt}}

Las dos construcciones en 42 son intransitivas y contrastan en que, en 42a, se emplea solamente la forma mínima del predicado, mientras que, en 42b, el complemento se introduce por el marcador *tyi*, que el autor trata como subordinador o complementante y que en otros contextos funciona como una preposición locativa genérica. Tanto las formas de finitud reducida con persona que aparecen en 41 como las que aparecen en 42 presentan el sufijo *-el*, razón que motiva a Vázquez Álvarez a tratar este sufijo como “no finito”, en vez de “infinitivo”, el cual se reservaría solamente para formas verbales como las del q’anjob’al ejemplificadas en 38.

El objetivo del estudio de Vázquez Álvarez es presentar argumentos que expliquen la distribución de los dos subtipos mostrados en 42. Para ello, el autor propone que la estructura en 42b, la de complementante, emerge como consecuencia de un proceso sintáctico subyacente que implica la ascensión del argumento sujeto del predicado subordinado al argumento matriz. El proceso se puede ver en su estadio inicial en 43a, donde la cláusula de complemento con persona del tipo ilustrado en

41 es el único argumento del verbo *mejl*, “poder” (Lit. “ser posible”). Como resultado de la ascensión, el argumento del verbo *wäyel*, “dormir”, aparece codificado como argumento de *mejl* en 43b (véase el capítulo 11, ejemplos 2a y 2b).

(43) CHOL

- a. *mejl-Ø* [k-wäy-el]
 poder-B3 A1-dormir-NF
 ‘Puedo dormir.’
 (Lit. ‘Es posible mi dormir.’)
- b. *mejl-oñ* [tyi ___ wäy-el]
 poder-B1SG P/SUB dormir-NF
 ‘Puedo dormir.’

Vázquez Álvarez argumenta que como consecuencia del proceso de ascensión se obtiene una saturación de argumentos en el verbo matriz, de forma que no se puede asignar caso sintáctico al complemento no finito. En este contexto, se requiere de la preposición *tyi* —reanalizada en este contexto como complementante— para dotar de caso sintáctico a la forma no finita, que sigue en superficie manteniendo el comportamiento de un nominal. El autor argumenta que la asignación de caso es un proceso independiente del proceso de ascensión y que también se puede evidenciar en estructuras como la del ejemplo 44, donde se explica que los verbos de movimiento toman como complemento una cláusula de propósito y que, al no poder asignarle caso ya que tienen solamente un argumento (*v.gr.* la persona que se mueve), requieren del complementante para tal efecto (véase el capítulo 11, ejemplo 39).

(44) CHOL

- ts-ajñ-Ø-ob* [tyi mǎñ-oñ-el]
 PRFV-llegar-B3-PL P/SUB comprar-AP-NF
 ‘Llegaron a comprar.’ {T_{xt}}

El proceso de ascensión sucede también con verbos causativos y con los de percepción directa, en los que el argumento ascendido funciona como objeto del predicado matriz, como en 45. El resultado petrificado del proceso se observa con los auxiliares aspectuales que en su día fueron verbos, como el marcador de imperfectivo ejemplificado con anterioridad en 42b (véase el capítulo 11, ejemplo 32).

(45) CHOL

tyi aw-il-ä-y-oñ [tyi — uk'-el]
 PRFV A2-ver-VTD-EPN-B1 P/SUB llorar-NF
 'Me viste llorar.'

Los predicados no finitos sin aspecto y sin persona en las lenguas mayas son típicamente intransitivos. El q'anjob'al muestra el caso más común. El sufijo *-oq* se asocia directamente a una raíz intransitiva para formar un infinitivo (*way-oq*, "dormir"). Cuando el verbo es transitivo de origen, se requiere primero modificar la valencia del verbo mediante sufijos de voz: antipasivo (por ejemplo, *q'an-waj-oq*, "pedir", o *kol-waj-oq* "ayudar") para dejar el argumento A como S, o pasivo (por ejemplo, *ante-lay-oq*, "ser curado") para que O sea S. Cuando el verbo transitivo se emplea en una construcción de incorporación, no se requiere el sufijo de antipasivo ya que la construcción en sí sirve de reductor de valencia (por ejemplo, *loh-oj china*, "comer naranja"). En algunos casos, la intransitividad permea incluso morfológicamente los predicados transitivos en la formas sin aspecto pero con persona. Esto sucede, por ejemplo, en q'anjob'al, donde estos predicados no sólo presentan un patrón de alineamiento no canónico como en otras lenguas mayas, sino que además aparecen marcados con el sufijo *-on* (46a) (nótese además el uso del sufijo de estatus de los verbos intransitivos en frontera prosódica, tal como se ilustra en 46b) (véase el capítulo 9, ejemplos 14 y 26a), y que aparece también en las construcciones de enfoque de agente que presentan igualmente transitividad defectiva (Aissen, 2006).

(46) Q'ANJOB'AL

- a. q-Ø-ey-ab' [he-way-i]
 POT-B3SG-A2PL-oír A2PL-dormir-VI
 'Se fijan como duermen.' {Txt}
- b. k'am ch-Ø-je' [Ø-ha-ten-on-i]
 NEG INC-B3SG-poder B3SG-A2SG-tocar-MD-VI
 'No lo puedes tocar/no es tocable.' {Txt}

Dos lenguas mayas en este libro divergen innovadoramente de este patrón de intransitividad, pero de manera distinta. Se trata, por un lado, del mam, y por el otro, del tselal. En mam, tal como se observa en el capítulo de England, el infinitivo de un verbo se codifica mediante el sufijo *-l* (por ejemplo, *b'eeta-l* "caminar"). Los verbos transitivos pueden aparecer intransitivizados en la construcción de incorpo-

ración nominal, como en *tx'eema-l sii*, “cortar leña”, pero a diferencia de todas las demás lenguas mayas pueden hacer igualmente explícito su argumento objeto de forma oblicua codificando la FN dentro de una frase de sustantivo relacional encabezada por *ee*, que en otras construcciones tiene varias funciones como marcador de dativo/benefactivo y posesivo, como en 47 (véase el capítulo 10, ejemplo 20),

- (47) MAM
 n=chi ku' teen xjaal [belaara-l t-e weech]
 INC=B3PL DIR estar gente velar-INF A3SG-SR:PAC gato_de_monte
 ‘La gente empezó a velar al gato de monte.’ {T_{xt}}

El sustantivo relacional aparece poseído en núcleo con una marca del juego A que hace referencia cruzada al sustantivo argumental del objeto y que se puede emplear por sí sola de forma pronominal y anafórica, como en 48 (véase el capítulo 10, ejemplo 12).

- (48) MAM
 n=ch=ul aaj=x=kya xjaal tzluu'
 INC=B3PL=DIR regresar=siempre=3PL.ENF gente aquí
 [b'inchaa-l t-ee]
 hacer-INF A3SG-SR:PAC
 ‘Viene la gente aquí a hacerlo.’

England apunta que este sustantivo relacional (junto con la marca pronominal) se suele repetir para indicar énfasis sobre el objeto (49) (véase el capítulo 10, ejemplo 34a).

- (49) MAM
 asta nn=Ø=ok teen meeb'a [woons-l t-e-t-ee
 hasta INC=B3SG=DIR estar huérfano tostar-INF A3SG-SR-A3SG-SR:PAC
 t-b'aq-al masaat]
 A3SG-hueso-ABST venado
 ‘[...] y se puso el huérfano a tostar los huesos del venado [...].’ {T_{xt}}

El *tzeltal* es otra lengua maya que ha innovado respecto a la restricción de intransitividad en las formas no finitas sin aspecto o persona con la creación de un infinitivo transitivo, tal como lo argumenta Polian en su capítulo. En esta lengua, al igual que en el *mam*, las formas no finitas están marcadas con el sufijo cognado

-*el*, que se emplea con los verbos intransitivos, como en *way-el*, “dormir”. Asociado a las raíces transitivas, el resultado es un predicado intransitivo que sorprendentemente, en el contexto de la familia maya, tiene orientación pasiva, es decir, se cancela el argumento A, y O se trata como S, tal como se ve en 50, donde hay una identificación de argumentos entre el agente de *mulan*, “gustar”, y el sujeto paciente de la raíz *jot'*, “rascar” (véase el capítulo 12, ejemplo 16).

(50) TSELTAL

ya j-mulan-Ø [jot'-el]
 INC A1-gustar-B3 rascar-NF.PAS
 ‘Me gusta ser rascado.’

Si se quiere expresar un objeto (subyacente) de tercera persona, se emplea una forma como la que se ejemplifica en 51 (véase el capítulo 12, ejemplo 19).

(51) TSELTAL

ya j-mulan-Ø [s-jot'-el j-jol]
 INC A1-gustar-B3 A3-rascar-NF.PAS A1-cabeza
 ‘Me gusta rascarme la cabeza.’
 (Lit. ‘Me gusta su ser-rascada de mi cabeza.’)

Esta estructura recuerda a la estructura común de posesión en chol del ejemplo 52, en el cual el nominal que funciona semánticamente como paciente del evento aparece poseído como sujeto de un predicado pasivo que, según Vázquez Álvarez, se puede interpretar como una nominalización con estructura posesiva semejante a la de otras FFNN, como en el ejemplo 53 (véase el capítulo 11, ejemplos 60b y 60a).

(52) CHOL

tyi k-il-ä-Ø [y-ixm-ään-ty-el ixim]
 PRFV A1-ver-VTD-B3 A3-desgranar-VTD-PAS-NF maíz
 ‘Vi desgranar el maíz.’
 (Lit. ‘Vi su ser desgranado del maíz.’)

(53) CHOL

tyi k-il-ä-Ø i-ts'i' aj-Wañ
 PRFV A1-ver-VTD-B3 A3-perro CLF-Juan
 ‘Vi el perro de Juan.’
 (Lit. ‘Vi su perro (de) Juan.’)

Aunque rebasa nuestro cometido el discutir si las formas mostradas en 52 se tratan en realidad de nominalizaciones poseídas o de formas verbales con el aspecto superficial de un nominal pero con la proyección de argumentos de un verbo, resulta sin embargo pertinente apuntar que el fenómeno es generalizado en otras lenguas mayas, como por ejemplo el k'iche' (Can Pinxabaj, 2009), y que estructuras no finitas de este tipo en estas lenguas muestran evidentemente su origen como nominales.

Sin embargo, en contraste con el chol, en la estructura tselal del ejemplo 51, Polian muestra que la FN que explicita el argumento paciente se puede elidir, tal como se ilustra en 54 (véase el capítulo 12, ejemplo 21a).

- (54) TSELTAL
 ya j-mulan-Ø [s-jot'-el]
 INC A1-gustar-B3 A3-rascar-NF.PAS
 'Me gusta rascarlo.'

Este comportamiento le hace a Polian postular que en tselal la marca *s-* del juego A en 54 *no* tiene función de poseedor de tercera persona, sino más bien funciona como un índice de concordancia de persona del argumento absoluto, solamente que con un juego de concordancia no esperado (*v.gr.* se tiene el juego A en vez del juego B). Esto resulta, por otro lado, morfosintácticamente coherente dentro de la familia, ya que las formas no finitas sin aspecto con persona presentan un alineamiento no canónico en otras lenguas. La prueba crucial que esgrime Polian de que *s-* no es un posesivo, radica en la imposibilidad de reemplazarla con una marca de poseedor de primera o de segunda persona que refiera al paciente del verbo transitivo. Esto se ve en la agramaticalidad de 55 (véase el capítulo 12, ejemplo 23), la cual, sin embargo *sí* puede aparecer en lenguas con estructuras de posesión como el k'iche', en 56 (Mondloch, 1981: 142).

- (55) TSELTAL
 *ya j-mulan-Ø [a-jot'-el]
 INC A1-gustar-B3 A2-rascar-NF.PAS
 Lectura buscada: 'Me gusta rascarte.'

- (56) K'ICHE'
 x-Ø-ki-maji-j [qa-yaaj-iik]
 COM-B3SG-A3PL-empezar-VT A1PL-regañar.PAS-NF
 'Empezaron a regañarnos.'
 (Lit. 'Empezaron nuestro ser-regañados.')

Para codificar un objeto de primera y segunda persona, el tselal ha innovado una forma sin aspecto transitiva marcada con el sufijo *-bel*, el cual recibe marcas de persona con alineamiento canónico, tal como se ve en 57 (véase el capítulo 12, ejemplo 23).

(57) TSELAL

ya j-mulan-Ø [j-jot'-bel-at]
 INC A1-gustar-B3 A1-rascar-NF.TR-B2
 'Me gusta rascarte.'

Por otro lado, la intransitividad de las formas no finitas no parece ser un rasgo exclusivo de las lenguas mayas. En muchas lenguas otomangués se tienen nominalizaciones verbales en contextos de complementación que funcionan en la inmensa mayoría de los casos como predicados igualmente intransitivos. El capítulo de Palancar, el último de este volumen, muestra en detalle la distribución sintáctica de formas nominalizadas verbales en la familia otomí (otopame); éstas se emplean como estrategia de complementación clausal y son también intransitivas. Un ejemplo de la estructura en cuestión se ofrece el ejemplo en 58, del otomí de la Sierra (Voigtlander y Echegoyen, 2007: 4).

(58) OTOMÍ DE LA SIERRA

i=pä^h=na mpefi yo ts'unt'u
 3.PRES.R=saber.D=SG.NOM trabajo PL muchacho
 'Los muchachos saben trabajar.'
 (Lit. 'Los muchachos saben el trabajo.')

El tipo de estructura que aparece en 58 no es una construcción favorita en estas lenguas dentro del ámbito la complementación y por ello tiene un ámbito funcional restringido a ciertos verbos matrices y a ciertos nominales. Sin embargo, la mayoría de sus instancias implican nominalizaciones intransitivas de verbos que denotan actividades humanas habituales.

A pesar de la predilección común por la intransitividad en los complementos clausales, las lenguas otomíes, tomándolas como representantes del espíritu que caracteriza a las otopames y a las otomangués en general, difieren extraordinariamente de las mayas en que por lo general no existe en éstas una operación morfológica sincrónica con la que los hablantes puedan crear formas no finitas de los verbos de forma totalmente productiva, y en consecuencia, la mayoría de las ins-

tancias son constructos gramaticales que funcionan como idiomatismos, mostrando una estructura petrificada hoy de lo que parece haber sido una sintaxis productiva en algún momento de la historia de estas lenguas.

1.6. Conclusión

Las lenguas de Mesoamérica son bien conocidas por sus características tipológicas compartidas, pero también por la gran diversidad de sus genios estructurales presentes en la región, desde lenguas altamente polisintéticas hasta otras que son puramente analíticas, lenguas con distintos alineamientos para marcar relaciones gramaticales, lenguas que recurren a distintos mecanismos de codificación de las relaciones gramaticales, lenguas con marcación en los núcleos o en los dependientes, lenguas en las que las categorías gramaticales son claramente distinguibles y otras en las que se presentan como híbridas, entre muchas características notables. En esta área lingüística se encuentran rasgos compartidos no solamente dentro de los límites de las familias, sino entre familias diversas, debido al contacto que tuvieron los hablantes de estas lenguas en épocas precolombinas.

Los siguientes capítulos, en este volumen abren vetas de investigación poco conocidas en la sintaxis de las lenguas de Mesoamérica y reflejan la gran diversidad de los dominios gramaticales y funcionales que son los ejes temáticos de la obra: clases léxicas, posesión e integración oracional. Las contribuciones contenidas en el libro son aportes que dialogan con la teoría sintáctica abocada al estudio de corte contrastivo, tipológico y con la teoría de la gramaticalización. De esta manera, los distintos estudios no solamente son contribuciones que incrementan el conocimiento de la morfosintaxis de las lenguas específicas sino que son aportes a las teorías interesadas en comprender la diversidad de los patrones estructurales, teniendo como base de estudio la diversidad de las lenguas habladas en los distintos continentes. Este volumen es una muestra de la riqueza de los diversos patrones a los que recurren algunas de las lenguas pertenecientes a cinco familias mesoamericanas.

A typology of split conjunction

ENRIQUE L. PALANCAR

Abstract

In this article, I study instances of noun phrase conjunction where the conjoined noun phrase is subject and the referents of the conjuncts are human, of the type 'John and Mary are having lunch'. More specifically, different splits are investigated that can occur in such structures, which involve the disruption of the phrasal continuity of the conjuncts, resulting in structures roughly equivalent to 'they are having lunch with Mary' and 'John are having lunch with Mary'. I claim that such splits are allowed by the inherent asymmetries of conjunctive structures, and I see them as functionally motivated by a play of prominence of the conjuncts involved. Three subtypes of such splits are distinguished: (i) splits by elision, where the conjuncts form a constituent but their phrasal contiguity is disrupted because the more salient conjunct is elided due to its topicality; (ii) splits by extraction, where the salient conjunct is given even more prominence by being extracted to a topical position thus breaking the continuity of the conjuncts; and (iii) splits by integration, where the role of the less salient conjunct is upgraded by means of dual or plural agreement on the verb. The constructions equally involve standard coordinative structures of the type 'Mary and I' as well as other non-coordinative types such as inclusory conjunction among others.

Keywords: agreement, associative, comitative, constituency, coordination, noun phrase, split conjunction, syntax, topicalization

1. Introduction

The study of coordinating constructions has been a subject of continuous interest for linguistic typology because of their great diversity in the world's languages, both language internally and crosslinguistically. It is for this reason that the literature is vast, and I will mention just a few works here. The various

studies in Haspelmath (ed.) (2004), together with the comprehensive surveys in Haspelmath (2004, 2007), are foundational in this respect. More specifically for the purposes of the present study, the typology of noun phrase (henceforth NP) conjunction in the various works by Stassen (2000, 2001, 2003) has been fundamental. Similarly, the studies in Corbett (1991, 2000, 2006) have become a landmark for the general agreement properties of conjoined NPs and their crosslinguistic peculiarities.

As novel linguistic data from lesser known languages expands our awareness of the existence of previously unknown possibilities in natural language, the enterprise of linguistic typology continuously faces the need to expand and accommodate this new knowledge. This article is intended as a contribution to the expansion of a typology of conjunctive constructions, more specifically in the area of NP conjunction, precisely in the light of aspects of conjoining constructions of a few languages which have not yet gained the attention of the typological literature and which pose interesting challenges for a taxonomy of conjunction.

In particular, I study here the different possibilities of a phenomenon I call “split conjunction”. Split conjunction comprises various types of conjunctive structures which share two properties: (i) the fact of having two discontinuous conjuncts in a clause and (ii) a clausal predicate that shows resolved agreement with both of the conjuncts. The discontinuity of the conjuncts may occur at a superficial level or at a deeper level in the syntax of clausal structure. In this connection, I briefly revisit the so-called “split inclusory pronoun constructions” in Lichtenberk (2000) – also called “asymmetrical conjunction” in Brill (2010, 2011, in press) – and propose to include them as a subtype of split conjunction. However, the article especially aims to present more severe cases of split conjunction that involve structures like the one shown in (1), from Santa Ana Otomi, an Oto-Manguean language of Mexico, where the two NP conjuncts appear in different positions in the clause, while the verb indexes both as subject via resolved agreement.

- (1) Santa Ana Otomi (Oto-Pamean, Oto-Manguean)
 [ma mbe] nuya bá=n-tsi-hme=wi
 1.POSS mother today 3.PST.TRANSLOC=ANTIP-ingest-tortilla=DU
 [a mbahä]
 DEF.SG priest
 ‘My mother ate today with the priest.’
 ‘My mother and the priest ate (together) today.’ (Selene Hernández,
 personal communication)

Conjoined NPs may appear in all levels of clausal structure, and may involve conjuncts ranking differently in the animacy hierarchy and bearing different

feature values (typically gender, number, and case). Because of this, for convenience, I have restricted the object of my study to instances of conjunctive strategies where the conjuncts rank highest in the animacy hierarchy (i.e., they are humans) and where the resulting conjoined NP functions as the subject of the clause (i.e., I exclude cases where the conjoined NP is object or possessor). Beyond convenience, the restriction of the object of study is also justified on the theoretical level, because the range of the conjunctive strategies involving the splits that I present in this article are often restricted only to subject conjoined NPs which involve humans acting together as a team in the completion of the event.¹

The article is structured as follows. First, in the next section I give an overview of what I believe are the principal parts in the state of the art regarding NP conjunction. This involves the relation between comitative and coordinative constructions, as well as other types of conjunctive strategies such as inclusory conjunction, summative conjunction, and associative conjunction. I intend this introduction to create a common background in a coherent and systematic way to be able to introduce split conjunction, which I do in Section 3. In Section 3, I define split conjunction and I propose that it comes in three disguises, as it were, namely in the form of splits by elision, splits by extraction, and splits by integration. I study these splits independently in subsequent sections, but advance areas of intersection in the end of the article in the conclusion section.

2. An overview of noun phrase conjunction

2.1. Defining conjunction: Coordination and comitative strategies

Although NP conjunction involves other types of conjunctive strategies (see Section 2.3), the phenomenon is commonly illustrated at a first instance by examples of NP coordination as in (2) from English, where the NP *John* and the NP *Mary* are conjoined into a larger NP structure *John and Mary*, by means of the relator or coordinator *and*.

(2) [*John and Mary*] were leaving for the station

The conjoining structure in English in (2) is further treated as syndetic because there is an overt relator (i.e., *and*) to link the two NPs. Example (3) illustrates an instance of NP conjunction by means of asyndesis, where the second conjunct is syntactically juxtaposed to the first one.²

1. In this article, I do not make a distinction between conjoined NPs involving full NP conjuncts, like 'the man and the woman', and others which involve pronominal conjuncts, like 'the man and I' or 'you and I'. Whether the subtleties of the distribution of conjunctive splits are due to the lexical nature of the conjuncts (i.e., nouns vs. pronouns) remains an open question.

- (3) Nyangumarta (Pama-Nyungan, Australian)
pala-nga ngatu jarri-nya-pinti-ngi, mima-nikinyi-yi
 that-LOC stationary INCH-NM-ASS-LOC wait.for-IMPV-3PL.SUB
puluku, [kujarra kangkuru-jirri waraja yalapara]
 3DU.DAT two kangaroo-DU one goanna
 ‘And there, on the finishing line, the two kangaroos and one goanna
 waited for those two.’ (Sadler & Nordlinger 2006: 340, citing Sharp
 2004: 315)

Both (2) and (3) are instances of NP coordination,³ which in this article, adopting Singer’s (2001a: 81) proposal, I define as a group of NPs in which elements of equal rank are conjoined together into a larger group, which is prototypically a group of two.⁴

On the other hand, the semantic relation between the linked NPs in (2) is one of “conjunction”, which is characterized in Stassen (2003: 763–764) as in (4):

- (4) A sentence contains a case of NP conjunction
- a. if it describes a single occurrence of an event (action, state, process, etc.)
 - b. and if this event is predicated simultaneously of two (and no more) participant referents, which are conceived of as separate individuals.

Notice, however, that the definition in (4) also accounts for the semantics of the typical comitative structure in (5), where only the participant John is subject, while Mary is encoded as an adjunct oblique.

2. An asyndetic conjoined NP is often framed within a special intonational phrase that cues the hearer to interpret a conjoined interpretation for the syntax. In example (i) of Jamsay, a Niger-Congo language of the Dogon family, the final syllabic nucleus of each conjunct is prolonged with a slowly falling pitch to which Heath (2008: 266) refers as “dying-quail intonation”, indicated by the sign “÷”.

(i) *mí÷ ñě-m÷*
 1SG woman-PL
 ‘me and the women’ (Heath 2008: 270)

3. According to Stassen (2001, 2003), asyndesis is much less frequent than syndesis as a coordinative structure, but it becomes quite common in other types of conjoining structures such as inclusory conjunction (see Section 2.3.1).

4. Different definitions of coordination abound. For example, Haspelmath (2007: 1) proposes that “the term COORDINATION refers to syntactic constructions in which two or more units of the same type are combined into a larger unit and still have the same semantic relations with other surrounding elements”, while Zwart (2005: 233) prefers a more formal definition as “a constituent *x* is an NP coordination iff *x* contains two or more NPs realizing a single argument or grammatical relation”. In principle all these definitions cover prototypical examples of the type in (2).

(5) [John] was leaving for the station [with Mary]

Both (2) and (5) are instances of (4) because for either of them there is an available interpretation that the act of leaving occurred only once (4a) and that this act was predicated simultaneously of both John and Mary (4b). However it is useful to bear in mind that while (2) and (5) are both seen as conjoining structures, only the structure in (2) is an instance of a coordination. Besides, the two structures are not perfectly synonymous because they have different semantic entailments if the state of affairs predicated is an action event, instead of a motion event, as for example in (6) and (7).

(6) [John and Mary] are having lunch

(7) [John] is having lunch [with Mary]

In the coordinative structure in (6), the content of (4a) is an implicature that can be cancelled. For example, while the default interpretation of both (6) and (7) is that John and Mary had lunch and that they had it simultaneously, in the same place and together, a possible alternative interpretation of (6) is that John and Mary indeed had the meal but the act of eating could have happened in different places (and of course not together). In contrast, in the comitative structure of (7) the content of (4a) is an entailment.

Following Stassen's (2003) proposal, I view the coordinative structure in (2) and (6) and the comitative structure in (5) and (7) as two opposing extremes in a grammatical continuum to encode conjunctive semantics. The defining features characterizing each strategy as a prototype are given in (8) (Stassen 2003: 780).

(8)	Coordinative strategy	Comitative strategy
	a. NPs have same structural rank	NPs differ in structural rank
	b. Unique coordinative particle	Unique comitative marker
	c. NPs form a constituent	NPs do not form a constituent
	d. Plural/dual agreement on the verb	Singular agreement on the verb

Of the four features in (8), feature (a) refers to the dimension of the symmetry/asymmetry of the relation between the conjuncts involved. In the prototype, the coordinative structure is symmetrical (i.e., it portrays the conjuncts as bearing a symmetrical relation to each other), while the comitative structure is asymmetrical because the two NPs differ in structural rank. In (5) and (7), for example, the referents of the NPs hold a semantic relation of accompaniment, but only one (*John*) is topical and subject (when established, it is often elided or encoded with an anaphoric pronoun), while the other participant (*Mary*) is less prominent and often carries new information, although it can also serve

as a secondary topic for the following discourse. The topical NP conjunct in the comitative structure is treated by Stolz (2001) as the “focal participant”, while the non-topical NP conjunct is called the “associate participant”.⁵ Besides, the asymmetry between the two conjuncts is one of control, as pointed out in Haspelmath (2004: 15–16). Here, of the two participants, the topical subject is the one in full control of the action and is the only one to be held accountable for it.

2.2. *The origin of comitative coordinators*

On the other hand, it is well known that comitative markers are often used as coordinators across languages, and the works by Stassen (2001, 2003, 2011) have shown that comitative coordination is typologically common. My understanding of Stassen’s proposal is that this situation is functionally motivated by an increase in the degree of symmetry between the conjuncts.⁶ For convenience, I illustrate this dynamism here from Hausa. In this light, the comitative construction in (9) would represent a first stage for the dynamics, a sort of a point of departure.

- (9) Hausa (Chadic, Afro-Asiatic)
 [Feemi] yaa tafi lkko [dà ùba-n-shì]
 Femi 3SG.M.PFV go Lagos with father-of-3SG.M
 ‘Femi went to Lagos with his father.’ (Abdoulaye 2004: 182)

In this comitative structure, the following two principles are implemented:

Principle 1:

The relation between two participants is more symmetrical if they are conceived of as a unit in the form of a group (typically of two members).

Principle 2:

A group is efficiently perceived as a unit if its members are close to each other, that is, if they are together.

The principles can be first realized by making the phrases encoding the two conjuncts appear contiguously in the clause. This is achieved by moving the comitative conjunct closer to the subject conjunct, as in (10).

5. Also, according to Stolz et al. (2006: 17), the NPs play different semantic roles: the focal participant is ACCOMPANEE while the associate participant is the COMPANION.

6. A further way to increase the symmetry between the conjuncts is to mark them equally; this was first proposed by Haiman (1985a: 84) for Hua (Papuan). Stassen (2003: 793) found some other instances in Nubian (East Sudanic, Nilo-Saharan), Cuzco Quechua (Andean), and Kobon (East Highlands, Papuan).

- (10) Hausa (Chadic, Afro-Asiatic)
 [Feemì] [dà ùba-n-shì] yaa tàfi Ìkko
 Femi with father-of-3SG.M 3SG.M.PFV go Lagos
 ‘Femi with his father went to Lagos.’

But the stage in (10) is highly unstable in both SVOX and VSOX languages because of the prominent informative position of the comitative phrase, and hence it is rarely attested and is likely to be produced only in elicitation.⁷ Instead, the emerging structure commonly leads to a sudden reanalysis of the syntactic boundaries of the phrases realizing the two conjuncts into a unified constituent. The semantic role of this newly emerged constituent is interpreted following another principle:

Principle 3:

A group is more symmetrical if its members have an equal degree of participation, that is, if they act together or have the same semantic role.

Such a principle is then realized by making the verb agree with the two conjuncts via semantic agreement involving number resolution (e.g., plural/dual), as in (11).⁸ The new structure gives rise to a comitative coordinator.⁹

- (11) Hausa (Chadic, Afro-Asiatic)
 [Feemì dà ùba-n-shì] sun tàfi Ìkko
 Femi **and**/with father-of-3SG.M 3PL.PFV go Lagos
 ‘Femi and his father went to Lagos.’ (Abdoulaye 2004: 183)

In this conception, which I take to be representative of the general view shared by many typologists, the coordinative strategy is viewed as a beacon of symmetry when the two conjunctive strategies in (2) and (5) or (6) and (7) are compared in the encoding of conjunctive semantics.¹⁰

7. For VSOX languages, this stage is not available because of the surface contiguity of S and the comitative phrase.

8. In resolved agreement or agreement via resolution, the appropriate feature values must be computed (see Corbett 2006: 168–171, 238–263).

9. In languages where verbs are not targets of subject number agreement, the constructions may remain ambiguous (see Amfo 2010 for a discussion of Akan).

10. The concept of symmetry used here is equivalent to “conceptual symmetry” as defined by Haiman (1985a: 73–74) where two elements A and B are said to be symmetrical with respect to a relationship *r* in at least one of the following two situations: (i) whenever {A *r* B} and {B *r* A} are both true (this is the prototypical case of actors involved in reciprocal relations like in [Max and Harry] hit each other); or (ii) whenever {A *r* C} and {B *r* C} are both true (this happens when the two actors are portrayed as performing the same action, as in [Max and Harry] hit each other).

However, most would also agree that the symmetry of coordination is a canonical ideal rather than the common situation, because in most instances of coordination one of the conjuncts is usually more prominent than the other. In my view, the underlying asymmetry of coordinative structures is inherited from the conceptual asymmetry characterizing the essence and core of conjunction itself.¹¹ In this regard, adopting Smith's (2003) terminology, I call "primary conjunct" the prominent conjunct of a conjunctive structure, whatever its subtype, whether coordinative or other. Likewise, the second next prominent conjunct will be referred to as the "secondary conjunct".¹²

In this respect, the asymmetry of NP coordination can be seen in various phenomena; the most relevant one being word order. Here the order of the conjuncts is far from being random, but responds to a number of hierarchies comprehensively characterized in Allan (1987), including degree of familiarity of the conjuncts (e.g. the primary conjunct comes first and is more familiar), topic vs. comment, universal sequencing conventions, etc. Asymmetry is also revealed when the order is fixed, because the lack of syntactic freedom points to an unequal ranking, as in well-known examples from English, such as those in (12); see Wälchli (2005) for an extended discussion.¹³

- (12) a. *mum and dad*/*dad and mum*
 b. *brothers and sisters*/**sisters and brothers*

-
11. As extensively discussed in Haiman (1985b), the asymmetry of coordination rests on the inherent linearity of the linguistic sign, because in the phonological string when saying *A and B*, *A* is necessarily perceived before *B*, and word order is invested here, as elsewhere in the grammar, with cognitive symbolism.
12. In formal theories of syntax such as Principles and Parameters and Minimalism, the common view of the structure of NP coordination since the works by Johannessen (1996, 1998), Kayne (1994), and Radford (1993) among others, is to understand coordination as bearing the typical asymmetrical structure of a single-headed X-bar phrase called ConjP, where the coordinator is the head, and the primary and the secondary conjuncts occupy the specifier and the complement positions, respectively. Coordination is viewed as inherently asymmetric because these theories take it that grammar is not capable of generating symmetric structures (Cormack & Smith 2005). Other theories do not assume this structure, for example Lexical Functional Grammar (Dalrymple & Kaplan 2000, Bresnan 2000), or Head-driven Phrase Structure Grammar (Pollard & Sag 1994). For reasons why the ConjP analysis is not accepted in these frameworks see Borsley (2005).
13. Similarly, the politeness rule that dictates that conjuncts should be aligned against the person hierarchy $1 < 2 < 3$, e.g., *my friend and I/me* instead of *me and my friend* or Spanish *Juan, Antonio y yo* ('John, Anthony, and I') instead of *yo, Juan y Antonio*, also responds to asymmetry. Whether this is a learned rule rather than an acquired one is irrelevant for our purposes, because what is important is that it is based on the perception that the order of the conjuncts is meaningful: placing the higher pronoun in the hierarchy in the last position is regarded as a sign of humility, and showing humility is a well-known strategy to encode politeness. In some cases, it grammaticalizes so that the option *I/me and my friend* is no longer as frequent as *my friend and I/me*, and for some even ungrammatical.

c. *ladies and gentlemen/*gentlemen and ladies*

Other phenomena revealing asymmetry between coordinated NPs is further seen when the verb agrees with only one of the conjuncts, commonly the nearest to the verb (Corbett 2006: 168–170, 180, etc.; Benmamoun et al. 2009); when only one of the conjuncts serves as locus for case marking (Johannessen 1998, Smith 2003, among others);¹⁴ when only the primary conjunct has binding properties (e.g., *John_i and his_i mother*, but *his_i mother and John_j*); or when the comitative coordinator still keeps government properties from being an apposition (for example in Russian, where the comitative coordinator *s* governs instrumental case, Feldman 2002: 41).

There are other types of conjunctive strategies identified in the literature which are more asymmetrical than coordination, and I will only revisit them briefly in the next section. These include conjunctive constructions by means of inclusory and summative pronouns as well as associate plurals. While coordination remains the least asymmetrical of conjunctive structures, the asymmetries present in all such constructions both feed and allow the natural occurrence of the conjunctive splits I study in more detail in Section 3. This is why it is important to review them here using the same approach.

2.3. *Other conjunctive strategies*

2.3.1. *Inclusory pronoun constructions.* In this section, I revisit briefly other types of conjunctive strategies found across languages. Among these the most common is the type that Haspelmath (2007) calls “inclusory conjunction”, also called “inclusory pronoun construction” in Lichtenberk (2000) and formerly known as the “plural pronoun construction” in Schwartz (1985, 1988a) and Aissen (1989). Variants of this construction occur widely across languages on all continents, while they are especially characteristic of many Austronesian (Bril 2004, 2010, 2011, in press), Australian (Singer 2001a, b), and Afro-Asiatic languages.

Inclusory conjunction is used to conjoin a pronoun, which is formally and semantically treated as the primary conjunct, plus the NP (or another pronoun ranking lower in the person hierarchy) of the secondary conjunct. The primary conjunct defines a superset in which the reference of the secondary conjunct is included as a subset, and for this reason it is called the “inclusory pronoun”. Two typical examples are given in (13). Here, as elsewhere in the article, I have underlined the secondary conjunct.

14. Outside coordination, a similar mismatch in case marking is also found in the inclusory conjunction of Tagalog. Here the inclusory pronoun, being the primary conjunct, may appear in an unmarked case, while the secondary conjunct (the subset included in the pronoun) remains case marked. For more details see Bril (2011: 250).

- (13) a. Nêlêmwa (New Caledonia, Eastern Oceanic)
 [yaman ma Polie]
 1DU.EXCL.FR CONJ Polie
 ‘Polie and I’ (Bril in press: 213)
- b. Tzotzil (Mayan)
 [vo ʔoxuk xchiʔuk i j-tzebe]
 you.PL and/with DEF 1.POSS-daughter
 ‘You (SG) and my daughter’ (Aissen 1989: 524)

For the purposes of this article, two aspects of inclusory conjunction are relevant. One is its inherent asymmetrical nature; the other verb agreement. I will discuss them succinctly here.

Inclusory conjunction is more asymmetrical than coordination. The asymmetry lies in the fact that the inclusory pronoun clearly functions as the head of the phrase, while the secondary conjunct is perhaps better treated as its specifier. As we will see in Section 3.2, the split that is common in inclusory conjunction lies at the heart of this asymmetry because the head inclusory pronoun is often elided by default. In the construction, the inclusory pronoun refers to a superset that includes the reference of the secondary conjunct by indexing what Singer (2001b) calls the “central member” of the conjunction. This central member corresponds to the focal participant in comitative semantics (i.e., the 1st person in *I with you*). The included subset is the associate participant (i.e., the 2nd person in *I with you*). In this respect, Bril (2004: 509–510) has shown that a situation involving the 1st and the 2nd person is naturally construed as [we.two.INCL & you], but also as [we.two.INCL & I] for very specific circumstances where the speaker lacks control of the action.¹⁵

The use of an inclusory construction to convey conjunctive semantics imposes its own characteristic semantic frame upon the situation. On the one hand, inclusory constructions appear to be equally translatable with a coordinative strategy (‘I and the man’) or with a comitative strategy (‘I with the man’). In reality, they represent a third semantic type, lying in the midst of these two strategies, which is not easily translatable in English or for that matter in other languages lacking the type.¹⁶ In this connection, Bril (2004, 2010, in press) has pointed out that the structures are commonly used to express a nuance of accompaniment that profiles a close relation between the participants involved. This is the case to such an extent, that in most languages the construction can only be used with humans, as for example, in Hungarian and in Polish

15. The example in question was said by a blind speaker who used to go fishing with the hearer on a regular basis and was thus guided by him. Here the speaker naturally portrayed himself as being less in control of the situation.

16. To compensate for this, authors often add a literal translation of the sort ‘we two including the man’.

(Schwartz 1988a: 242, citing Edith A. Moravcsik, personal communication), or only with animates, as in Southern Zapotec (Beam 2006) and in Nêlêmwa (Bril 2004).¹⁷

Another important feature that becomes relevant with split cases is to keep in mind that when the inclusory phrase is subject,¹⁸ the verb agrees with the inclusory pronoun. This is clearly shown in Hausa in (14).

- (14) Hausa (Chadic, Afro-Asiatic)
 [muu dà shii] mun jee kàasuwaa
 1PL and/with 3SG.M 1PL.PFV go market
 ‘He and me, we went to the market.’ (Abdoulaye 2004: 166)

Apart from inclusory conjunction, two other minor types remain: summative pronouns and associative plurals. I introduce each succinctly next.

2.3.2. *Summative pronoun constructions.* In some languages, the conjuncts can also be conjoined by means of another construction that involves a head pronoun that computes the number values of each of the conjuncts put together, as illustrated in (15). I call this construction “summative conjunction” inspired by Bril (2010), who calls it a “summation pronoun strategy”.¹⁹ The asymmetry of summative conjunction lies in the fact that both conjuncts are headed by the

17. As Bril (2010) points out, the existence of the phrasal inclusory construction correlates with the avoidance of conjunction of pronouns in many languages; not a definitional feature of the construction, but a natural consequence of its existence for expressive purposes. These are banned in Nêlêmwa, but are to be found in Taqabaqita (Bril 2010: 8), where the use of and-coordination is restricted to same level salience. At times they use different means, asyndesis for the inclusory construction and syndesis for coordination, like in Ajië (Oceanic) (Bril 2010: 6). Notice that the use of an explicit conjunction commonly separates the conceptual unity of the inclusory group.

18. In some languages, mostly Oceanic and Chadic, the inclusory conjunction can also be used in phrases that work as objects and as possessives.

19. Aikhenvald (2008) prefers the name “argument elaboration construction”, whereas Haspelmath (2007) treats it as a subtype of inclusory conjunction on the basis that the superset of the summative pronoun in a way “includes” the reference of the subsets of the individual conjuncts.

summative pronoun.²⁰ The examples in (15) further illustrate that the verb agrees with the summative pronoun.²¹

- (15) a. Manam (Austronesian, Oceanic)
 [nau toʔa-gu ʔéru] ʔi-réba-ru
 1SG.FR older.brother-1SG 1DU.EXCL.FR 1DU.EXCL-sail-DU
 ‘I and my older brother sailed.’ (Bril 2010: 362, citing Lichtenberk 1983: 430)
- b. Mapudungun (isolate)
 [iñché eymi iñchiu] i-y-u
 I you we.DU eat-IND-1DU
 ‘You (SG) and I ate.’ (Smeets 2008: 138)

Alternatively to summative pronouns, in the Southern Zapotec languages of Mexico, one finds summative conjunction by means of quantifier verbs (Beam 2006), as illustrated in (16). This type will become relevant for the discussion of splits in Section 3.3 further below.

- (16) Quiégolani Zapotec (Oto-Manguean)
 r-oo men nisgaal [y-rup men Biki]
 HAB-drink 3 soda POT-be.two 3 Virginia
 ‘She and Virginia drink soda pop.’ (lit. ‘They drink soda being two she and Virginia.’) (Black 2000: 260)

2.3.3. *Associative conjunction.* Finally, there is another type of conjunction that involves an NP encoding the superset by virtue of being marked with an associative plural. This NP includes the reference of the secondary conjunct as its subset. The construction is not widespread, being limited already to languages with associative plurals. Haspelmath (2007: 34) acknowledges having found only one example of it, which is given in (17a) from Margi.²² The ex-

20. The ranking of the conjuncts is as symmetrical as they can get in standard coordinative constructions where their order is motivated by hierarchical principles.

21. Summative conjunction can become an inclusory construction if the primary conjunct (the one that selects the person value of the head summative pronoun) is elided. This is shown in (i) from Mapudungun, if compared with (15b).

(i) [ʔey iñchiu] nüttram-ka-y-u
 he we.DU conversation-FACT-IND-1DU
 ‘He and I talked.’ (Smeets 2008: 139)

22. The example, with a slightly different orthography, appears in Vassilieva (2005: 2). As for published material, the example first appears in Moravcsik (2003: 494), but it has errors, including the glossing of -yàr as a plural marker, not as an associative plural. Moravcsik acknowledges in Footnote 26 that she was made aware of this structure by Martin Haspelmath

ample in (17b) is from Central Alaskan Yup'ik, whose relevance is discussed in Corbett (2000: 107–110).²³

- (17) a. Margi (Chadic, Afro-Asiatic)
 [Kàmbèràwázhá'-yàr àgá màlà géndà]
 Kamburawazha-ASS.PL and/with wife of.him
 'Kamburawazha and his wife.' (Haspelmath 2007: 34, citing Hoffmann 1963: 57)
- b. Central Alaskan Yup'ik (Eskimo Aleut)
 [Cuna-nku-k arnaq=llu] ayag-tu-k
 Chuna-ASS-DU woman=too go-IND-3.DU
 'Chuna and the woman left.' (Corbett & Mithun 1996: 12)

In this section, we have seen various encoding possibilities for conjunctive semantics. These include coordinative, inclusory, summative, associative, and comitative conjunction. The schematic syntax of these strategies is given for convenience in (18). In this notation, the conjunct (nominal or pronominal) that serves as the primary conjunct in each of the structures is indicated by the subscript "1", the secondary conjunct by the subscript "2". The ampersand represents a relator. In syndetic inclusory, summative, and associative strategies the relator used is commonly a comitative one, but I have placed it in parentheses to include asyndetic instances as well, which are more common in these types of conjoining structures than in coordination.

and Maria Vassilieva. Another mention of the same example is in Bhat (2004: 92). All go back to a manuscript by Haspelmath dated 2000, which later became Haspelmath (2007).

23. Associative conjunction is the mirror-image of a special inclusory conjunction construction found in a number of Australian languages where the secondary conjunct is explicitly marked as being a subset of the larger set indexed by the inclusory pronoun (Singer 2001b). Such a construction is illustrated from Yidiñ (Pama-Nyungan, Australian) in (i).

(i) [ŋaŋđi buŋa:-ba] galiŋ
 1.NONSG woman-ONE.OF.GROUP go.PRES
 'A woman and I (and some others) are going.' (Dixon 1977: 178)

- (18) Different conjunctive strategies
- | | | |
|------------------|--|-------------------------------|
| a. Coordinative: | | |
| | $[NP_1 \ \& \ NP_2]_{SUB}$ | V+DU/PL |
| b. Inclusive: | | |
| | $[PRO_{1+DU/PL} [(\&COMIT) \ NP_2]]_{SUB}$ | V+DU/PL |
| c. Summative: | | |
| | $[PRO/V+DU/PL [NP_1 \ (\&COMIT) \ NP_2]]_{SUB}$ | V+DU/PL |
| d. Associative: | | |
| | $[NP_1+ASS [(\&COMIT) \ NP_2]]_{SUB}$ | V+DU/PL |
| e. Comitative: | | |
| | $[NP_1]_{SUB}$ | V+SG [COMIT NP ₂] |

To close this section it is sensible to remind ourselves that languages often have more than one conjunctive strategy at hand. When this happens, the distribution and usage of the alternatives is often conditioned by the semantics,²⁴ but also by discourse functions.²⁵ All the strategies schematized in (18), except comitative conjunction, typically involve conjuncts in contiguous phrases, as seen in the examples so far, but such a contiguity can be broken, crucially resulting in a phenomenon I call “split conjunction”. In the following sections, I study this phenomenon in detail. In the next section, I start with a definition, and then I present its various types.

3. Split noun phrase conjunction

3.1. Defining split conjunction

I use the term “split noun phrase conjunction”, abbreviated simply as “split conjunction”, to refer to a conjunctive structure involving two NPs (nominal or pronominal) which has the following two properties: (i) the conjuncts are discontinuous in the clause, and (ii) the clausal predicate, commonly a verb, shows resolved agreement with both of the conjuncts treating them equally as subject. None of the strategies presented so far, summarized in (18) above, qualify as instances of split conjunction. The comitative structure fulfills (i) but

24. An exemplary case is illustrated by Tariana (Arawakan), which employs asyndetic coordination for inanimates but comitative coordination for humans (Aikhenvald 2003). The distribution may equally respond to the type of conjuncts involved. For example, among other constructions, Nêlêmwa (Oceanic, Austronesian) has a syndetic coordinative strategy to conjoin full NPs, but restricts the use of the inclusive conjunction in (13a) to animate conjuncts involving a pronoun (Bril 2004). Likewise, Manam (Oceanic, Austronesian) uses summative conjunction when the conjunction involves a pronoun, but comitative coordination when conjoining full NPs (Bril 2010: 326 citing Lichtenberk 1983: 430).

25. For example, according to Aikhenvald (2008), Manambu (Sepik, Papuan) has up to four different strategies to conjoin NPs involving human referents: asyndetic and comitative coordination, inclusive conjunction, and summative conjunction.

not (ii). The rest of the strategies (i.e., coordinative, inclusory, summative, and associative) fulfill (ii), but not (i).

Split conjunction is a phenomenon that responds to the inherent asymmetry existing between the two conjuncts as it is perceived by speakers. In this respect, speakers play with the prominence of the conjuncts in the way they choose to present them in the description of the event, while they still portray them equally as co-actors through the resolved agreement in the verb. As a result of this interplay of prominence, we obtain different types or structural manifestations of split conjunction. In other words, the different types of split conjunction reveal different possible ways of resolving the conflict of which one of the two co-actors gets the stage, the protagonist or the antagonist. I propose there are three main types of conjunctive splits: “splits by elision”, “splits by extraction”, and “splits by integration”.

One common way to play with the asymmetry of the conjuncts is by downgrading the role of the primary conjunct. This results in “split conjunction by elision”, which involves both inclusory and coordinative constructions, as I show in the next section. In splits by elision, the secondary conjunct is the only overt constituent in the clause. This split is introduced in Section 3.2.

Alternatively, another way to deal with the conflict of prominence between the conjuncts is to bluntly present them as asymmetrical. This is commonly achieved by making the two conjuncts appear in different positions in the clause. However, to resolve the conflict, speakers may decide to reduce the asymmetry by presenting them equally as co-actors in the event. This alternative leads to at least two other types of splits. In “split conjunction by extraction”, presented in Section 3.3, the relevance of the primary conjunct is enhanced by extracting its phrase to a prominent syntactic position in the clause. This split is found in inclusory and summative conjunction, although only rarely. In “split conjunction by integration”, on the other hand, the discontinuity of the conjuncts is there by default, but the role of the secondary conjunct is upgraded from its background by integrating it as a co-actor in the event through the resolved agreement on the verb. This type of split is introduced in Section 3.4. Finally, in some languages of the Americas, there are also instances of conjunctive constructions featuring a split by integration which work as independent constructs in the grammar of conjunction. I refer to such cases as “severe split conjunction” and they will be dealt with more extensively in Section 3.5.

In principle, split conjunction may involve all types of the conjunctive strategies in (18), but because of its crosslinguistic rarity, we lack evidence that languages with the associative type have developed a split. Consequently, I will not be saying anything else about this type here.

3.2. Split conjunction by elision

3.2.1. *Illustration.* The primary conjunct is likely to be the established topic from the previous discourse. If this happens, it may be entirely downgraded by not being overtly expressed. This is possible because the verb allows for some type of resolved agreement with all the conjuncts together.²⁶ I treat such cases as instances of “split conjunction by elision”, where the term “elision” is intended here in a noncommittal manner as a way to refer to the process by which the NP (or a pronoun) of one of the conjuncts is not expressed in surface structure.

The elision of the primary conjunct produces a superficial type of split which is common across languages with inclusory conjunction. The conjunctive constructions bearing this split are currently known in the literature as “split inclusory pronoun constructions” from Lichtenberk (2000), but they were initially called “verb-coded coordination” in Schwartz (1988b). As a consequence of the elision, the secondary conjunct is the only conjunct of the whole conjoined NP that surfaces in the clause. A number of illustrative examples are given in (19).

- (19) a. Toqabaqita (Oceanic, Austronesian)
 [*doqora-mu*] mere ngata
 brother-2SG 1DU.EXCL.NONFUT speak
 ‘Your brother and I spoke.’ (Lichtenberk 2000: 3)

26. In the computation of feature values in resolved agreement, the person value is calculated following the same principles that rule the selection of the inclusory pronoun in inclusory conjunction. Such principles are spelled out in Corbett (2006: 240) as: “If the conjuncts include a first person, agreement will be first person. Otherwise, if the conjuncts include a second person, agreement will be second person. The default situation is that agreement is third person.” In coordination, resolved agreement commonly attends to such principles, but there can be exceptions. For example in French, Grevisse & Goose (1993: 1329–1330) report on various (although rare) instances where agreement is resolved with the 3rd person, as in *des efforts que faisaient Damien et moi-même* ‘the efforts Damien and I used to make’ (Dhôte, *Je ne suis pas d’ici*, p. 40) instead of more canonical *des efforts que faisons Damien et moi-même*, or *mes enfants et moi vous disent mille tendresses* ‘my children and I send you all our love’ (Sand, *Correspondence*, 20 juillet 1845) instead of *mes enfants et moi vous disons mille tendresses*. For other cases, agreement with the 3rd person is with the nearest conjunct as in *lorsque toi et tes frères ouvraient à Noël leurs paquets de cadeaux somptueux* ‘when you and your brothers opened your/their packets of sumptuous presents in Christmas’ (Simenon, *Mémoires intimes*, p. 11). In reality, when French native speakers are confronted with made-up examples showing a similar phenomenon such as for example *Jean et moi sont arrivés à la gare* for *Jean et moi sommes arrivés à la gare* ‘John and I arrived at the station’, the consensus is to regard them as ungrammatical. Speakers react to the deviating examples above as slips of the pen. In this respect, Grevisse & Goose (1993: 1329) comment: “il s’agit dans plus d’un cas, il est vrai, de lettres ou d’autres écrits peu élaborés”.

- b. Nêlêmwa (Oceanic, Austronesian)
io ma tu haga [ma Polie]
 FUT 1DU.EXCL go.down fish CONJ Polie
 ‘I’ll go fishing with Polie’ or ‘Polie and I will go fishing.’ (Bril 2004: 508)
- c. Tzotzil (Mayan)
l-i-bat-otikotik ta ch’ivit [xchi?uk li]
 COMPL-1.ABS-go-1PL.EXCL to market and/with DEF
Xun=e]
 Xun=ENC
 ‘I went to the market with Xun.’ (Aissen 1989: 522)
- d. Hausa (Chadic, Afro-Asiatic)
mun jee kàasuwaa [dà Abdù]
 1PL.PFV go market and/with Abdu
 ‘We went to the market, Abdu and me.’ (Abdoulaye 2004: 185)

In these examples where the conjoined NP functions as subject, the secondary conjunct designating the subset is the only conjunct overtly expressed as an NP. The primary conjunct, the inclusory pronoun indexing the superset, is elided because it is topical, although its reference can be recovered from the context and from the agreement index on the verb. The split construction in these examples can be viewed as an instance of syntactic pro-drop, see Aissen (1989) and Corbett (2000: 233).²⁷ In this direction, Schwartz (1988b), who sees this phenomenon under the general rubric of “verb-coded coordination”, proposes that allowing pro-drop elsewhere in the grammar is indeed one of the fundamental prerequisites for a language to have this type of split.

The schematic syntax of this split is given in (20). For languages with syn-thesis, splits by elision are only attested when the relator linking the secondary conjunct to the inclusory pronoun already works as a comitative relator. Split inclusory conjunction is not a fully independent construction in the grammar, but an alternative manifestation of the construction where both conjuncts form a contiguous phrase.

27. In my opinion, the wording used by some authors to describe the agreement pattern in split conjunction by elision is not a fortunate one because it ascribes an inclusory function to the agreement morphology, when this is not entirely necessary. For example Lichtenberk (2008: 669): “In split inclusory constructions, the inclusory pronominal and the included NP do not form a phrase together. There is no independent inclusory pronoun; rather, the superset IS ENCODED BY MEANS OF A DEPENDENT PRONOMINAL” (emphasis mine). The equation is taken a step further in Haspelmath (2007: 34), where the phenomenon is described as “many languages also have a construction in which THE INCLUSORY PRONOMINAL ELEMENT IS A CLITIC PRONOUN OR A COREFERENCE MARKER ON THE VERB” (emphasis mine).

(20) Split by elision and inclusory conjunction

Contiguous:

[PRO₁+DU/PL [(&COMIT) NP₂]]_{SUB} V+DU/PL

Split:

[____I_{GIVEN.TOP} +DU/PL [(&COMIT) NP₂]]_{SUB} V+DU/PL

The elided inclusory pronoun can be restored, but when this happens the clause is pragmatically charged, as the overt occurrence of the topical primary conjunct is construed as being emphatic, as in (21). Also, the full phrase may be treated differently in the syntax for information structure purposes. This may be seen in (21c) from Tzotzil, a VOS language where the subject is extracted as external topic by means of the enclitic =*ʔun*, or in Hausa in (21d), a SVO language where it is placed in topical position.

- (21) a. Toqabaqita (Oceanic, Austronesian)
 [*kamareqa doqora-ku*] *meki lae ma-i*
 1DU.EXCL brother-1SG 1DU.EXCL.FUT go DIR-at
 ‘I and my brother will come tomorrow.’ (Lichtenberk 2000: 10)
- b. Nêlêmwa (Oceanic, Austronesian)
mo pîla wuung [yamon ma âlô-raamwa
 2DU stroll together 2DU.FR CONJ young-woman
eli]
 that.ANPH
 ‘You took a walk, you and that girl.’ (Bril 2011: 241)
- c. Tzotzil (Mayan)
 [*li vo ʔotikotik xchiʔuk li Maryan=ʔun=e*] *te*
 DEF we.EXCL and/with DEF Maryan=ENC=ENC there
l-i-kom-otikotik
 COMPL-1.ABS-stay-1PL.EXCL
 ‘Maryan and I stayed there.’ (Aissen 1989: 524)
- d. Hausa (Chadic, Afro-Asiatic)
 [*muu dà shiî*] *mun jee kàasuwaa*
 1PL and/with 3SG.M 1PL.PFV go market
 ‘He and me, we went to the market.’ (Abdoulaye 2004: 185)

3.2.2. *Split coordination by elision.* Although splits by elision are typical of inclusory conjunction, they can also be found in coordinative conjunction, although apparently more rarely. An example of split coordination by elision is (22) from Chilean Spanish.

- (22) Chilean Spanish
fu-imos a-l cine [con mi madre]
 go.PST-1PL to-the cinema and/with my mother
 ‘My mother and I went to the cinema.’ (Schwartz 1988b: 54)

On the surface, the structure in (22) looks like an instance of the split construction in (19), but in reality Chilean Spanish, like any other dialect of Spanish, lacks an inclusory construction. This is illustrated by the impossibility of the inclusory reading of (23) (cf. example 21d).

- (23) [nosotros con mi madre] fu-imos a-l cine
 we.M and/with my mother go.PST-1PL to-the.M cinema
 Exclusory reading: 'We and my mother went to the cinema.'
 *Inclusory reading: 'I and my mother went to the cinema.'

In this way, the real elided primary conjunct in the split in (22) is the pronoun for the 1st person singular *yo*, as illustrated in (24). As in all examples in (21) above, once the primary conjunct is restored, the construction is emphatic, and like in the Hausa example in (21d), the full phrase occurs preverbally as topical subject.²⁸

- (24) [yo con mi madre] fu-imos a-l cine
 I and/with my mother go.PST-1PL to-the cinema
 'I and my mother went to the cinema.' (Schwartz 1988b: 64)

Just as in inclusory constructions, the split by elision in coordinative structures is only allowed with comitative relators. For example, in Chilean Spanish, while the genuine coordinator *y* 'and' would be perfectly possible in the contiguous construction in (24), it cannot be used in the split of (22), as shown by the infelicity of (25), presumably because of the Coordinate Structure Constraint by Ross (1967).

- (25) Spanish (infelicitous in all dialects)
 *fu-imos a-l cine [y mi madre]
 go.PST-1PL to-the cinema and my mother
 Intended reading: 'My mother and I went to the cinema.'

In this way, the schematic syntax of split coordination by elision is given in (26).

- (26) Split by elision and coordination
 Contiguous:
 [NP₁ &COMIT NP₂]_{SUB} V+DU/PL
 Split:
 [___₁ GIVEN.TOP &COMIT NP₂]_{SUB} V+DU/PL

28. Comitative coordination in Spanish has been studied in Camacho (1996, 2000). In the latter work, Camacho shows that this conjunctive strategy is relegated to subject position only and that it must have a collective interpretation.

Another example of split coordination by elision is found in the Boumaa dialect of Fijian. Consider, for instance, the coordinative structure in example (27a), which is emphatic, or (27b), where both Mary and John represent new information.

- (27) Boumaa Fijian (Polynesian, Oceanic, Austronesian)
- a. *'eirau aa sota vata [o yau 'ei Jone]*
 1DU.EXCL PST meet together ART 1SG and/with John
mai Viidawa
 at Viidawa
 'John and I met at Viidawa.' (Dixon 1988: 157)
- b. *'erau aa sota vata [o Mere 'ei Jone]*
 3DU PST meet together ART Mary and/with John
mai Viidawa
 at Viidawa
 'Mary and John met at Viidawa.' (Dixon 1988: 157)

When the primary conjuncts in (27) are the established topics of the discourse, older speakers have available the split encoding in (28),²⁹ while younger speakers prefer the comitative construction in (29), where the verb no longer shows resolved agreement.

- (28) a. *'eirau aa sota vata ['ei Jone] mai*
 1DU.EXCL PST meet together and/with John at
Viidawa
 Viidawa
 'John and I met at Viidawa.' (Dixon 1988: 157)
- b. *'erau aa sota vata ['ei Jone] mai Viidawa*
 3DU PST meet together and/with John at Viidawa
 'She and John met at Viidawa.' (Dixon 1988: 157)

29. Notice that the predicate *vata* 'together' in these examples is not a summative pronoun like in Section 2.3.2 above. The same structure obtains when *vata* is not used, like in (i) (Dixon 1988: 158).

- (i) *o rau saa sota 'ina [ei+na gone.tuuraga a*
 ART 3DU ASP meet PREP+3SG and/with+ART high.chief ART
tuuraga.bete ni Faranisee, o Paatere Lorosio]
 priest ASSOCIATED France ART Father Lorenzo
 'They two met there, (Tui Ca'au) and a great chief, a French priest, Father Lorenzo.'
 (Dixon 1988: 157)

The adverbial predicate *vata* 'together' is used in (28) and (29), just like in many other instances, as a semantic indicator that the conjuncts act as a group. For younger speakers, the predicate has become reanalyzed as part of the complex preposition *vata 'ei* 'with'.

- (29) *au aa sota vata* [*'ei Jone*] *mai Viidawa*
 1SG PST meet together with John at Viidawa
 'I met up with John at Viidawa.' (Dixon 1988: 159)

3.2.3. *The secondary conjunct in splits by elision.* Leaving out the primary conjunct may equally serve as a mechanism to increase the narrative salience of the secondary conjunct. An example of this is given in (30a) from Manambu, where the participant *Lumawandəm*, a chief of a renowned village to the hearers, is presented as a focal protagonist of the scene, yet encoded as the secondary conjunct. Aikhenvald (2008) treats this example as an instance where the inclusory pronoun is left out. Another example with comitative coordination in the same language is given in (30b), where the secondary conjunct is presented as an anti-topic.

- (30) Manambu (Sepik, Papuan)
- a. *aw a* [*Lumawadəm*] *kwa-brə-l*
 then then Lumawandəm stayed-3DU.SUB-3SG.F.BAS
 'Then Lumawandəm (and the village) stayed.' (Aikhenvald 2008: 141)
- b. *gra-kə-dana-di* [*də-kə*]
 cry-FUT-3PL.SUB.VT-3PL.BAS.VT 3SG.M-OBL.3SG.F
 [*takwa-wa*]
 woman.LK-COMIT
 'They all cried, (them) including his wife.' (Aikhenvald 2008: 142)

The latent asymmetry allows for the extraction of the secondary conjunct to a position of external topic (31),³⁰ and it can be further questioned (32) or relativized (33).³¹

- (31) Boumaa Fijian (Polynesian, Oceanic, Austronesian)
 [*a tuuraga.bete mayaa*] *'eirau na vei-vasa'i vata*
 ART priest DEM 1DU.EXCL FUT REC-talk together
 'That priest, he and I will talk together.' (Dixon 1988: 160, my glossing)

30. The construction in (31) is the most common, but a pronominal copy of the secondary conjunct in the form of *'ei 'ea* or *kaya* 'and/with him' can be left behind in the clause after the predicate *vata*.

31. The relative clause is based on a structure like *porque cuando nos casamos con mi esposa ...* 'because when my wife and I got married (lit., because when we married my wife)' from http://www.elsolquimes.com.ar/imprimir.php?n_id=43097.

- (32) Nêlêmwa (Oceanic, Austronesian)
hli u muuvi [ma ti]?
 3DU PRF stay CONJ who
 ‘Who did (s)he live with?’ (Bril 2004: 508)
- (33) Chilean Spanish
Por esos días también conocí a quien
 by those days also met.1SG.PST ACC.DEF.ANIM who
sería mi señora, un-a hermos-a
 would.be.3SG my wife a-F beautiful-F
puertomontin-a [con la que nos
 person.from.Puerto.Montt-F with the.F REL 1PL.REC
casa-mos a fin-es de 2003]...
 marry-1PL(PST) at end-PL of 2003
 ‘In those days I met (the person) who would be my wife, a beautiful woman from Puerto Montt to whom I got married (lit., with whom we got married) towards the end of 2003 ...’ (from <http://www.elrepuertero.cl/admin/render/noticia/15967>)

But not all languages allow the same syntactic freedom to the phrase of the secondary conjunct. For example, if the comitative phrase in the Hausa example in (34a) is extracted, the resulting structure is no longer an instance of inclusory conjunction, as the plural agreement on the verb is no longer interpreted as targeted by the inclusory pronoun, but by a plural pronoun with a genuine plural reference. This is illustrated in (34b), where the comitative phrase is external topic in the topic fronting construction, and in (34c) where it occurs in a cleft focus structure.

- (34) Hausa (Chadic, Afro-Asiatic)
- a. *mun jee kàasuwaa [dà Bàlaa]*
 1PL.PFV go market with Bala
 Inclusive reading: ‘I went to the market with Bala.’
 Exclusive reading: ‘We went to the market with Bala.’
- b. *[dà Bàlaa] kàm, mun jee kàasuwaa*
 with Bala indeed 1PL.PFV go market
 ‘As for Bala, we did indeed go to the market, he and me.’
- c. *[dà Bàlaa] nèe mu-kà jee kàasuwaa*
 with Bala COP 1PL-PFV go market
 ‘It is with Bala that we went to the market.’ (Abdoulaye 2004: 186)

In this section, I have presented instances of split conjunction by elision, where the primary conjunct is elided because it is the established topic of the

discourse. Alternatively to elision, we also find conjunctive structures where the asymmetry between the conjuncts is maximized by either breaking their contiguity and thus producing a different type of split I call “split conjunction by extraction”. This is studied in the next section.

3.3. *Split conjunction by extraction*

Speakers may also want to foreground the role of the primary conjunct in the event, increasing further the inherent asymmetry of conjunction. One way to achieve this is by breaking the contiguity of the conjuncts by placing the primary conjunct in a prominent informative position in the clause. This produces a more severe disruption in the syntax of the clause than in the split cases by elision presented in the previous section, which I call “split conjunction by extraction”. Splits by extraction produce a double effect. The primary conjunct is clearly presented as the participant who is in control of the action, while the secondary conjunct becomes backgrounded. However, as the verb maintains resolved agreement with both conjuncts, both participants are still portrayed as performing the action together.

In inclusory conjunction, a split by extraction takes the form of fronting the inclusory pronoun to a topical position. This solution is illustrated in Nêlêmwa in (35), an example where the topicalization is justified as a way to assert the creditable source of information to find out about the identity of the secondary conjunct, which is questioned. The position from where the primary conjunct is extracted is indicated by the underscore.

- (35) Nêlêmwa (Oceanic, Austronesian)
 [yamon xe] mo tu yhalap [__ ma ti]?
 1DU.EXCL.FR TOP 1DU.EXCL go.down gather CONJ who
 ‘You, you went food-gathering with whom?’ (Bril 2004: 513)

The schematic syntax of the split applied to an inclusory construction is given in (36).

- (36) Split by extraction and the inclusory construction
 Contiguous: V+DU/PL [PRO₁+DU/PL [(&COMIT) NP₂]]_{SUB}
 Split:
 [PRO₁+DU/PL]_{TOP/SUB} V+DU/PL [____₁ [(&COMIT) NP₂]]_{SUB}

An example of a split by extraction involving summative conjunction is given in (37a) from the Mixtepec dialect of Miahuatec Zapotec, where again the primary conjunct is fronted to a topical position. Another example is (37b), where it is focalized. It should be born in mind at this point that Miahuatec Zapotec, like any other Southern Zapotecan language, uses a quantifier verb

instead of a summative pronoun. This was illustrated in example (16) above in Section 2.3.2.

- (37) Miahuatec Zapotec, Mixtepec dialect (Zapotecan, Oto-Manguenan)
- a. [m-bě't] ngw-à g-ǎw [rǒp —
 ANIM.CLF-child COMPL-go POT-eat COMPL.be.two
 x-nà'a]
 POSS-mother
 'The child went to eat with his/her mother.' (Beam 2010: 45)
- b. [lě'ě mě] ngw-à [rǒp — tǔě'l
 FOC 3.HUM.KNOWN COMPL-go COMPL.be.two spouse
 mě] ndá kanĩ'
 3.HUM.KNOWN HAB.go dance
 'She (was the one who) went with her spouse to dance.' (Beam 2006: 11)

The schematic syntax of this construction is given in (38).

- (38) Split by extraction and the summative construction
- Contiguous: V+DU/PL [PRO/V+DU/PL [NP₁ (&COMIT) NP₂]]_{SUB}
- Split:
 [NP₁]_{TOP/SUB} V+DU/PL [PRO/V+DU/PL [____₁ (&COMIT) NP₂]]_{SUB}

The split construction in the specific examples (37) is of special interest here because it serves a special purpose in the grammar of Miahuatec Zapotec, namely, it allows the primary conjunct to be overtly expressed in the clause. This may sound surprising, but Zapotecan languages have a syntactic constraint that bans the overt occurrence of a subject NP that is coreferential with the possessor of some other constituent in the clause. Because of this constraint, the primary conjunct in examples like (37) would not be allowed to surface in subject position. This is where the split construction becomes useful. I will elaborate briefly on this point.

As an illustration of the constraint, consider the examples in (39) from Quiégolani Zapotec. In (39a), there is an explicit NP subject that occurs postverbally. In (39b), the verb shows agreement with a 1st person subject. In both instances the subject is NOT coreferential with a possessor. The examples in (40) contrast with those in (39), where the subject is coreferential with a possessor. In such cases, the subject participant is not expressed in the subject slot (the absence of agreement for 1st person in (40b) shows the same phenomenon).

- (39) Quiégolani Zapotec (Zapotecan, Oto-Manguéan)
- a. *w-ey-t Benit mēl*
COMPL-take-NEG Benito fish
'Benito didn't take fish.' (Black 2000: 173)
- b. *čín-aʔ x-ičR-boʔ*
POT.comb-1SG POSSD-head-3.FORMAL
'I will comb his hair.' (Black 2000: 72)
- (40) a. *bito b-nežRw ___ bg^wex [če noʔol-ənʔ]*
not COMPL-give broom of woman-the
'The woman_i did not lend her_i broom (lit., (she)_i did not lend the woman_i's broom).' (Black 2000: 73)
- b. *čín___ x-ičR-aʔ*
POT.comb POSSD-head-1SG
'I will comb my hair (lit., (I)_i will comb my_i head).' (Black 2000: 72)

The split in (37) is a useful way to go around this constraint so that the primary conjunct can be encoded overtly.³² Another way could involve the fronting of the entire conjoined NP, as in (41a). These two strategies can be compared with the form a conjoined NP takes when it occurs *in situ*, shown in example (41b), that is, abiding by the constraint. Here the conjoined NP is subject, but internally, just like in (40) the pronoun *mē* cannot be used to index the primary conjunct, because it is coreferential with the possessor.

- (41) Miahuatec Zapotec, Mixtepec dialect (Zapotecan, Oto-Manguéan)
- a. [*lè'ë mē rōp bē'l mē*]
FOC 3.HUM.KNOWN COMPL.be.two sister 3.HUM.KNOWN
nhé yù'g nītyèʔ
AUX cook food
'It is she and her sister that are preparing a meal.' (Beam 2010: 44)

32. As verbs in Miahuatec Zapotec are not targets of plural agreement, examples in (37) could in principle be mistaken as instantiating a comitative structure. However, there is another comitative structure in the language which is not summative as it does not include the reference to the primary conjunct. This construction, shown in (i) from the Loxicha dialect of Miahuatec Zapotec, may use the comitative marker *kon*, borrowed from Spanish *con*. Notice that here the primary conjunct is repeated in the comitative structure as a way to maintain co-reference of the possessor with the subject.

(i) [*Pedr*] *ngw-à zi'n [kon rōp xi'n Pedr]*
Peter COMPL-go work COMIT COMPL.be.two child Peter
'Peter went to work with his two sons.' (Beam 2006: 12)

- b. *nhé yù'g [röp (*me) b'él mě]*
 AUX cook COMPL.be.two sister 3.HUM.KNOWN
nityèt
 food
 'She and her sister are preparing a meal.' (Rosemary Beam, personal communication)

On the other hand, in the non-configurational languages of Australia, where discontinuity of constituents is a characteristic feature of the syntax (Hale 1983, Austin & Bresnan 1996, Haviland 1979, etc.), splits in conjunctive structures naturally occur triggered by the requirements of information structure (Blake 2001). The splits happen alike in inclusory conjunction (42) and in coordinative conjunction (43). In the latter example, the primary conjunct (i.e., the one marked with the coordinator) also appears in a prominent topical position, notice for example the translation.

- (42) Inclusory conjunction
 Kuuk Thaayorre (Pama-Nyungan)
wey, [ngali] yancm [ngan waanharr]
 hey 1DU.EXCL(NOM) go.PL.IMPV relative brother(NOM)
iipal
 from.there
 'Hey, my brother and I have come here.' (Gaby 2006: 303)
- (43) Coordinative conjunction³³
 Kalkatungu (Pama-Nyungan)
 [*kinja-(ng)ku=yana*] *intji-mi-ngi-yu ntiya-(ng)ku*
 female-ERG=and pelt-FUT-me-they.DU stone-ERG
 [*tjipa-yi kurlayingu-thu*]
 this-ERG male-ERG
 'The girl and boy will both pelt me with stones.' (Bril 2010: 326, citing Blake 2001: 423)

In configurational languages, splits by extraction involving coordinative structures are not found, allegedly obeying the Coordinate Structure Constraint

33. An example of a discontinuous conjoined NP functioning as object is given in (i) from Kuuk Thaayorre (Pama-Nyungan, Australian).

- (i) *ngul ngay kirk kempthe kal-m thul=yuk*
 then 1SG.ERG spear.ACC apart carry-IMPV woomera.ACC
 'I used to carry spears and woomeras separately.' (Gaby 2006: 320)

by Ross (1967) (see the examples of Nakanai (Oceanic, Austronesian) in Section 4 for possible counterevidence). In the next section, I present other conjunctive constructions which involve yet another type of split I call “split by integration”. Such a split has not received much attention in the typological literature but poses, nonetheless, intriguing questions for a natural theory of conjunction.

3.4. Split conjunction by integration

In the previous section, I studied a few cases of splits by extraction where the role of the primary conjunct is made more prominent. In contrast to such splits, speakers may instead want to upgrade the role of the secondary conjunct as a way of reducing the inherent asymmetry between the conjuncts. However, they do it in such a way that the secondary conjunct is still not given equal discourse prominence as the primary conjunct, that is, without resorting to the coordinative structure. All this is achieved by yet another very different type of conjunctive split I call “split conjunction by integration”, whose schematic syntax is given in (44).

- (44) Split conjunction by integration
 [NP₁]_{TOP/SUB} V+DU/PL [COMIT NP₂]

The conjunctive structure in (44) involves a split where the NP encoding the primary conjunct is syntactically treated as a canonical topical subject while the secondary conjunct is encoded discontinuously in a comitative phrase. In this respect, the structure in (44) is in essence the comitative structure in (18e) above, repeated here as (45a). The fundamental difference between them lies in the fact that the verb in the construction in (44) shows resolved agreement with both conjuncts, making the structure closer to the coordinative structure in (18a), repeated here as (45b).

- (45) a. Comitative: [NP₁]_{SUB} V+SG [COMIT NP₂]
 b. Coordinative: [NP₁ & NP₂]_{SUB} V+DU/PL

An example of split coordination by integration is given in (46a) in Manambu, an SOV non-Austronesian Papuan language. This example may be contrasted with the comitative structure in (46b) and the standard coordination construction in (46c). Notice that the conjuncts in (46a) do not form a constituent despite their phonological contiguity because in the contiguous structure in (46c), the comitative coordinator *-wa* is associated with the primary conjunct.³⁴

34. Another example of split coordination is found in Tariana (Arawakan), as shown in (i). Here the secondary conjunct is a list of participants. This structure contrasts with (ii), which has a contiguous conjoined NP.

(46) Manambu (SOV) (Sepik, Papuan)

- a. [də-kə-də asa:y] [nəkə-də du-a-wa]
 he-OBL-M.SG father other-M.SG man-LK-COMIT
 asa-wa ata yi-di
 dog.LK-COMIT then go-3PL.BAS.VT
 ‘His father with another man, with a dog then went.’ (Aikhenvald 2008:) 162)
- b. [wun] [lə-kə-wa] warya-k
 I she-OBL-COMIT fight-PURP.SS
 yi-k-na-dəwun-ək
 go-FUT-ACTFOC-1SG.M.BAS.VT-CONF
 ‘I will go to fight with her.’ (Aikhenvald 2008: 162)
- c. [a takwa-wa lə-kə ñan] ata
 DEM.DIST.F.SG woman.LK-COMIT she-OBL child then
 kwa-bər
 stay-3DU.BAS.VT
 ‘The woman and the child then stayed there.’ (Aikhenvald 2008: 160)

Another language with a similar split is K’ichee’, a Mayan language from Guatemala. The split structure is illustrated in (47). For the split to be structurally obvious, the primary conjunct is required to appear preverbally in a topical position; otherwise, as the language has a canonical VOS word order, the common position for the subject is to occur postverbally and it could induce the interpretation of (47b), which is an instance of comitative coordination. The comitative construction is given in (47c) where the verb agrees only with the primary conjunct.

-
- (i) [pa:-dapana]-sina na-yā na-yā-nhi [diha]
 one-CLF.HAB-REM.PST.INFR 3PL-stay 3PL-stay-ANT he
 di-we-ne-ne diha ha-niri-nipe-ne
 3SG.NF-younger.sibling-PL-COMIT he parent-M-PL-COMIT
 di-īdua-ne di-daki-ne
 3SG.NF-daughter.in.law-COMIT 3SG.NF-grandchild.PL-COMIT
 ‘He [a Tariana] with his younger siblings, with his parents, with his daughter-in-law, with his grandchildren, stayed in one house.’ (Aikhenvald 2003: 151)
- (ii) na:ka neyù na: [di-we-ri]
 3PL.come 3PL+climb 3PL.go 3SG.NF-younger.sibling-M
 di-phe-ri-ne
 3SG.NF-elder.sibling-M-COMIT
 ‘They came up, the younger and the elder brother.’

- (47) K'ichee' (VOS) (Mayan)
- a. [*le achi*] *x-ee-chaku-n* [*r-uk'*]
 DET man COMPL-3PL.ABS-work-ANTIP 3.POSS-and/with
le nu-xb'aal
 DET 1.POSS-brother
 'The man worked with my brother.' / 'The man and my brother worked.'
- b. *x-ee-chaku-n* [*le achi r-uk'*]
 COMPL-3PL.ABS-work-ANTIP DET man 3.POSS-and/with
le nu-xb'aal
 DET 1.POSS-brother
 'The man and my brother worked.'
- c. *x-∅-chaku-n* [*le achi*] [*r-uk'*] *le*
 COMPL-3SG.ABS-work-ANTIP DET man 3.POSS-with DET
nu-xb'aal
 1.POSS-brother
 'The man worked with my brother.' (Telma Can, personal communication)

In contrast to K'ichee', in the split construction of Tzotzil, another Mayan language, there is no need for the topical extraction of the primary conjunct because the comitative phrase encoding the secondary conjunct can precede the NP encoding the primary conjunct, and the two phrases could not possibly be interpreted as forming a contiguous constituent.

- (48) Tzotzil (VOS) (Mayan)
?i-ve?-ik [*xchi?uk jkaxlan*] [*li Xun=e*]
 COMPL-eat[3.ABS]-PL with Ladino DEF Xun=ENC
 'Xun ate with the Ladino.' (Aissen 1989: 533)

The split structure in (44), illustrated in (46a), (47a), and (48), is a unique hybrid of the comitative and the coordinative structures which nicely combines the functional strengths of both constructions. On the one hand, the resolved agreement on the verb semantically portrays both conjuncts equally as co-actors in the event. This morphological mechanism proves successful at increasing the degree of symmetry between the conjuncts, and in this respect it reminds one of the function of coordination. On the other hand, the syntactic discontinuity of the conjuncts allows the introduction of the primary conjunct as a more topical participant than the secondary conjunct, and in this respect it reminds one of the asymmetries proper of the comitative construction.

In syntactic terms, the split in (44) is structurally analogous to the comitative structure in (45a) and it is thus sensible to account for it as having developed from the comitative structure by adding resolved agreement on the verb. I can

think of at least four arguments in favour of such an analysis. Firstly, the split in (44) always involves a comitative linker and never a canonical *and*-type of coordinator. However, the fact that a language makes use of a comitative linker as a coordinator does not imply that it will allow the split. This is for instance illustrated by the ungrammaticality of (49) in Chilean Spanish.³⁵

- (49) Chilean Spanish
 ***[yo]** *fu-imos a-l cine [con mi madre]*
 I go.PST-1PL to-DEF.SG cinema with 1.POSS mother
 Intended meaning: ‘My mother and I went to the cinema.’ (Schwartz 1988b : 64)

Secondly, the syntactic position of the secondary conjunct in the split structure is the same as a canonical comitative participant. Thirdly, the two conjuncts in either structure are also treated alike for the purposes of information structure. In other words, the primary conjunct is topical in either structure and thus may serve as an established topic for the following discourse. In contrast, the secondary conjunct in either structure works as an anti-topic; that is, it may become the topic for the following discourse.

A clear example of the topicality of the primary conjunct in splits of this kind is seen in colloquial Hausa in example (50). Here the primary conjunct appears preverbally, while the secondary conjunct follows the verb in the comment string of the utterance. According to Abdoulaye (2004), the split in (50a) responds to the pressures of information structure and is a common reply to the question ‘Where is Abdu?’. In (50b), it is further shown that the NP *Abdù* is the established topic in the following string of discourse as it cannot be repeated, but the secondary conjunct can.

- (50) Hausa (Chadic, Afro-Asiatic)
- a. [**Abdù**] *sun tàfi makaranta [dà Bàlki]*
 Abdu 3PL.PFV go school and/with Balki
 ‘Abdu went to school with Balki . . .’
- b. *an baa [*Abdù]/[Bàlki] àlloo*
 1MPS.PFV give Abdu/Balki board
 ‘. . . and a writing board was given to *Abdu/Balki.’ (Abdoulaye 2004: 187)

The primary conjunct of the split structure in (50a) is treated like the topical subject of the comitative strategy in (51a). This may again be seen in the behavior of (51b). The same applies to the secondary conjunct.

35. The same case is illustrated for Russian in McNally (1993: 353).

- (51) Hausa (Chadic, Afro-Asiatic)
- a. [**Abdù**] *yaa* *tàfi* *makařantaa* [*dà* **Bàlki**]
 Abdu 3SG.M.PFV go school with Balki
 ‘Abdu went to school with Balki, ...’
- b. *an* *bâa* [***Abdù**]/[**Bàlki**] *àlloo*
 IMPS.PFV give Abdu/Balki board
 ‘... and a writing board was given to *Abdu/Balki.’ (Abdoulaye 2004: 181)

In contrast, in the coordinative strategy in (52a) both participants are equally topical and either of them can be elided (if chosen as topic) or repeated (if chosen as anti-topic) in the following context. This is shown in (52b).

- (52) Hausa (Chadic, Afro-Asiatic)
- a. [**Abdù** *dà* **Bàlki**] *yaa* *tàfi* *makařantaa*
 Abdu and/with Balki 3PL.PFV go school
 ‘Abdu and Balki went to school, ...’
- b. *an* *bâa* [*Abdù*]/[**Bàlki**] *àlloo*
 IMPS.PFV give Abdu/Balki board
 ‘... and a writing board was given to Abdu/Balki.’ (Abdoulaye 2004: 181)

Finally, there are three languages in Stassen’s (2003) sample which do not appear to have a coordinative structure: Tolai (Oceanic, Austronesian), Acooli (Nilotic, Nilo-Saharan), and Tera (Chadic, Afro-Asiatic). Accordingly, the closest equivalent to the symmetrical semantics evoked by a coordinative strategy in other languages is expressed in these languages by means of a split structure, as shown by the contrast in the translation between (53a) and (53b).

- (53) Tolai (Oceanic, Austronesian)
- a. [*nam* *ra* *tutana*] *i* *ga* *rovoy* [*ma* *ra* *pap*]
 DEM ART man he TA hunt with ART dog
 ‘That man hunted with the dog.’
- b. [*Telengai*] *dir* *rovoy* [*ma* *ra* *pap*]
 Telengai they.DU hunt and/with ART dog
 ‘Telengai and the dog hunted.’ (Mosel 1984: 176)

In the light of this evidence, the split construction in (44) inherits the discontinuity of the conjuncts typical of a comitative structure. I call this split “split conjunction by integration”, using the term “integration” to indicate that the upgrading of the secondary conjunct as an equal co-actor in the event is achieved by its semantic “integration” into the core of the verb by means of resolved agreement. Nonetheless, while the comitative construction may well

be the source of this conjunctive structure, at least for the instances presented so far (see Section 3.5 for other possibilities), how the resolved agreement on the verb is achieved remains an intriguing question.

As a possible answer, one may want to view the agreement on the verb as having a semantic function rather than a morphosyntactic one. In this respect, in an influential article for a theory of agreement, Corbett (1988) introduces a number of problematic cases in the Slavonic languages where plural agreement does not respond to requirements of the syntax as is otherwise expected in more canonical cases.³⁶ An exemplary case for this deviation is the associative construction found in Talitsk Russian, as in (54), where a plural verb can be used with a singular NP subject “to indicate that the noun refers to a person or persons besides the one indicated directly” (Corbett 1988: 36). Since then, a similar construction has also been reported in other languages, for example in Maltese, as shown by the contrast between (55a) and (55b).

- (54) Talitsk dialect of Russian
M'it'ixa dra-l'-i-s'
 Mitixa fight-PST-PL-REFL
 ‘Mitixa (and his wife) had a fight.’ (Corbett 1988: 36, citing Bogdanov 1968)
- (55) Maltese (Semitic)
- a. *Brian ġie*
 Brian came.3SG
 ‘Brian came.’
- b. *Brian ġew*
 Brian came.3PL
 ‘Brian and his family/friends came.’ (Corbett 2000: 191, citing Fabri 1993: 276–278)

This associative construction is also found in the Mayan languages K'ichee' and Tzotzil. Example (56) illustrates it from Tzotzil. Similarly, the split examples in (47a) and (48) could also have the alternative interpretations ‘the man

36. For example, plural agreement with a singular NP subject is used to indicate respect, as in the old construction in Russian:

- (i) [mamen'ka] *plačut* – *šepnula ona vsled uxodivšej Elene* – *a*
 mother cry.PL whispered she after leaving Elena and
 [papen'ka] *gnevajutsja*
 father are.angry
 ‘“Your mother is crying”, she whispered after Elena, who was leaving, “and your father is angry”.’ (Ivan Turgenev, *Nakanune*, 1860, cited in Corbett 1983: 25)

and some other people worked with my brother' and 'Xun and some others ate with the Ladino', respectively.

- (56) Tzotzil (VOS) (Mayan)
ʔi-veʔ-ik [li Xun=e]
 COMPL-eat[3.ABS]-PL DEF Xun=ENC
 'Xun ate with someone.'
 'Xun and some others ate together.' (Aissen 1989: 533)

To account for (56), Aissen (1989) proposes that proper nouns in Tzotzil like *Xun* 'John' can have a group interpretation and thus can trigger plural agreement on the verb. In K'ichee' the construction also allows common nouns providing they have a specific referent. In this respect, Aissen advances an analysis for the split structure in (48) as being a natural extension of the associative construction. The extension happens when the companion participant encoded in the comitative phrase, the secondary conjunct, is identified as the associate member of the set evoked by the group interpretation of the noun indicating the primary conjunct. The reading involving only two people is possible because there is no *a priori* expectation that the set evoked should include more than two members.

In this view, split conjunction by integration results from a semantic extension of the associative construction. However, it should be born in mind that not all languages that have the associative construction allow the two-participant interpretation proper of the split construction. This is illustrated by Maltese, where the example in (57) can only have a group interpretation, but never one that includes the secondary conjunct.

- (57) Maltese (Semitic)
 [Brian] *gew* *fil-karozza* [*mal-president*]
 Brian came.3PL in.DEF-car with.DEF-president
 'Brian (and others) came with the president.'
 *'Brian came with the president.' (Maris Camilleri, personal communication)

In this section, I introduced some illustrative cases of split conjunction by integration. A construction bearing this type of conjunctive split has the two conjuncts encoded discontinuously in the clause in the same way as a comitative structure, but the verb shows resolved agreement to indicate that the secondary conjunct has been semantically integrated in the event as an equally participating co-actor. The split is functionally motivated by a desire to reduce the asymmetry between the two conjuncts by upgrading the role of the secondary conjunct. I have further proposed that a construction that shows split conjunction by integration is commonly based on a comitative structure and

emerges by using dual or plural resolved agreement on the verb. One of the possible ways verbal agreement is implemented is via the exploitation of the possibilities provided by an associative construction of the same type found in Talitsk Russian or Maltese. However, it still remains an open question whether all the languages that display split conjunction by integration have such an associative construction. As I show in the next section, there are some other more extreme cases of split conjunction by integration that are not based on an alternative comitative structure.

3.5. *Exploring the limits of splits by integration*

3.5.1. *Severe split conjunction.* There are other conjunctive constructions showing a type of split reminiscent of the splits by integration in the previous section, but which are not based on a readily available comitative structure. In other words, such constructs are not structurally related to a comitative structure of the standard type. Because of this, they function as fully independent constructions for the encoding of conjunctive semantics. Nevertheless, I still treat them as a subtype of split conjunction by integration because of two reasons: (i) they respond to the very same functional motivation as the cases presented in the previous section and (ii) they share a considerable degree of structure with them. However, it is important to distinguish them clearly as a different ontology, and to do so, I treat these new cases as instances of “severe split conjunction by integration”, abbreviated as “severe split conjunction”, with the schematic syntax in (58). In the light of these severe cases, the instances presented in the previous section could be revisited as instances of “mild split conjunction by integration”, whose schematic syntax is repeated in (59) for convenience.

(58) Severe split conjunction by integration
 [NP₁]_{TOP/SUB} V+DU/PL [NP₂]

(59) Mild split conjunction by integration
 [NP₁]_{TOP/SUB} V+DU/PL [COMIT NP₂]

In severe split conjunction, the two conjuncts occur in different positions in the clause just like in the mild cases, but the secondary conjunct is promoted by the construction as an argument of the verb and it is thus encoded as an unmarked anti-topic NP.

Constructions showing severe split conjunction are found in three unrelated language groups of the Americas. In North America, it occurs in the Eastern Algonquian languages Passamaquoddy and Mi'kmaq, while in Mesoamerica it is found in the languages of the Otomi-Mazahua subgroup of Oto-Manguean and in Misantra Tononac of the Totonac-Tepehua family. The data further suggest that the constructions in Algonquian and the Otomi-Mazahua instantiate a

similar type of construction, whereas the one of Misantla Totonac instantiates its own type. Because of this, it is sensible to present them separately. The extent to which severe split conjunction occurs in the languages of the world goes beyond my knowledge, but one can hypothesize that the phenomenon will be found in linguistic areas or linguistic families where the asymmetrical encoding of conjunctive semantics is favored over the symmetrical one.

Severe split conjunction has been reported for Passamaquoddy in Bruening (2003, 2004, 2005) who treats it as a case of “split coordination”, also found in closely related Mi’kmaq in a manuscript by Carlos Quicoli mentioned in Bruening (2005). The construction in question is illustrated in (60). The primary conjunct appears in the position dedicated to a topical subject, the secondary conjunct appears postverbally, and the verb shows resolved agreement in the plural.

- (60) Passamaquoddy (Algonquian, Algic)
 [Piyel] *ali-wiciyew-t-uwok* [Mali-wol]
 Piyel around-go.with-REC-3PL Mary-OBV
 ‘Piyel and Mary are going around with each other.’ (Bruening 2005: 5)

A similar construction is also found in the Otomi-Mazahua languages of Mexico. The case is illustrated from Toluca Otomi in (61) and in Highlands Otomi in (62), which make use of verb agreement in the dual. Notice that these languages instantiate two different word orders, but in both cases the primary conjunct occurs in the position of subject.³⁷

- (61) Toluca Otomi (SVO) (Otomi-Mazahua, Oto-Manguean)
 [_{RA} Xúwa] *bo=’ñ>ó=wí* [_{RA} Pédro]
 SG John 3.PST=<SS>walk=DU SG Peter
 ‘John came with Peter.’ (modified from Lastra 1989: 121)
- (62) Highlands Otomi (VOS) (Oto-Manguean)
 ‘*buh=mi* [_{rá} *ndombe*] [_{nu} *ra t’uhni*]
 [3.PRES]live=DU 3SG.POSS grandmother DEF.SG SG child
 ‘That child lives with his grandmother.’ (Voigtlander & Echegoyen 1979: 307)

On the surface, these constructions show structural resemblance to mild instances of split construction by integration. One fundamental difference lies in the fact that the NP of the secondary conjunct is not introduced by a coordinator. In other words, the secondary conjunct is encoded in an unmarked NP.

37. The primary conjunct in a VOS language like Highlands Otomi can also be moved to preverbal position. This happens when it is highly topical as in (68) further below.

Evidence that there is no underlying zero coordinator comes from the fact that there is no asyndetic coordination in these languages. This is shown by the impossibility of the (a) examples in (63) and (64). Alternatively, as shown in the (b) examples, the languages also have a syndetic construction that is only used extremely rarely and under very specific circumstances. For example, the coordinative structure of Toluca Otomi in (64b), an instance of comitative coordination (with the borrowed marker *ko* ‘with’ from Spanish), is only used to emphasize that the participants act together.³⁸

- (63) Passamaquoddy (Algonquian, Algic)
- a. **[Piyel Mali(-wol)] ali-wiciyew-t-uwok*
 Piyel Mary-OBV around-go.with-REC-3PL
 ‘Piyel and Mary are going around with each other.’
- b. *[Piyel naka Mali] ali-wiciyew-t-uwok*
 Piyel and Mary around-go.with-REC-3PL
 ‘Piyel and Mary are going around with each other.’ (Bruening 2005: 5)
- (64) Toluca Otomi (Otomi-Mazahua, Oto-Manguean)
- a. **[r_A Xúwa r_A Pédro] bo=’ñ>ó=wi*
 SG John SG Peter 3.PST=<SS>walk=DU
 Intended reading: ‘John came with Peter.’
- b. *[r_A Xúwa ko r_A Pédro] r_{AK}’ú*
 SG John and/with SG Peter together
i=pěh=mi
 3.PRES.REAL=work.AS=DU
 ‘John and Peter work together.’ (Lastra 1989: 121)

Furthermore, this split structure is not structurally based on a comitative structure, but works independently from it if such a comitative construction exists. For example, while the split structure is the old and native Otomi-Mazahua means to encode conjunctive semantics, some Otomi languages have also borrowed a comitative construction from Spanish, but this is used only rarely, most significantly in the situations where the native structure has distributional restrictions (see Section 3.4.2) or as a stylistic alternative to it. The later point is illustrated in the textual extract in (65) from San Ildefonso Otomi. Compare for this purpose the split structure in (65c) with the comitative construction in (65b) and (65e).

38. Notice also the co-occurrence of the predicative *r_{AK}’ú* ‘together’.

- (65) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguean)
- a. *nú'mu gé ... jö di='rá-hk-a=r* 'befi
if DUB god 3.SUBJ=give.to.1/2-1.DAT-BS=SG work
'If god were to give me a job.'
 - b. *ga=m-pef-i* [*ko ngé='á*]
1.PRES.IRR=ANTIP-work-FS with FOC=3SG
'I'd work with that one.'
 - c. *nú'mu[=r zithü] ga=n-thé='bé*
if=SG devil 1.PRES.IRR=MIDDL-meet=DU.EXCL
'If I'd meet the devil.'
 - d. *di='rá-hk-a=r* 'befi
3.SUBJ=give.to.1/2-1.DAT-BS=SG work
'(And if) he were to give me a job.'
 - e. [*ko=r zithü*] *ga=m-pef-i* 'něhe
with=SG devil 1.PRES.IRR=ANTIP-work-FS as.well
'With the devil I'd work as well.'

In severe split conjunction, the conjuncts are treated as different constituents, and both can be fronted independently for questioning, relativization, or focus. Examples in (66) show questioning in Passamaquoddy (the literal translation is Bruening's; the idiomatic translation is mine). The examples in (67) are again from San Ildefonso Otomi. In (67a), the secondary conjunct is external topic, recovered in the clause by the resumptive pronoun *nó*. In (67b), it is internal focus, just like in (65c). In the Otomi examples, the primary conjunct is elided because it is the established topic. This elision further illustrates that one can have splits by elision within splits by integration (see Section 4 for discussion).

- (66) Passamaquoddy (Algonquian, Algic)
- a. [*Wen*] *ali-wiciyew-ti-htit* [*Mali-wol*]?
who around-go.with-REC-3PL.CONJTV Mary-OBV
'Who is going around with Mary? (lit., Who are ___ and Mary going around with each other?)'
 - b. [*Wen-il*] [*Piyel*]
who-OBV Piyel
ali-wiciyew-ti-htic-il ___?
around-go.with-REC-3PL.CONJTV-PART.OBV
'Who is Piyel going around with? (lit., Who are Piyel and ___ going around with each other?)' (Bruening 2005: 6)

(67) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguean)

- a. [nǒ=’ǒ=r zónē], [nó] *hín=tó*
 DEF.SG=DEM.III.SG=SG wizard 3SG NEG=someone
da=ñ-húx=kwí _____
 3.PRES.IRR=MIDDL=sit.on.AS=DU
 ‘That wizard, he has no equal (lit., nobody sits at the same level
 with him).’
- b. *ya* [hónt’ä nǒ=r k’éñö] *mí=’óh=wí*
 PARTCL NARRFOC DEF.SG=SG serpent 3.IMPF=sleep.AS=DU
 _____ *hyáxtho=’yá*
 daily=PARTCL
 ‘And she used to sleep daily with the serpent.’ (Palancar 2009:
 565–566)

While the role of the primary conjunct as topical subject is relatively straightforward because of its position, it is less clear what the syntactic role of the secondary conjunct can be in these constructions. I address this question briefly in the following section.

3.5.2. *The secondary conjunct in severe split conjunction.* The secondary conjunct in severe split conjunction is encoded as an unmarked NP. There is evidence that points in the direction that this NP is not treated as an adjunct by the syntax, but as an argument of the verb. In this sense, the construction can be seen as promoting the secondary conjunct to argument status, much in the vein of comitative applicative constructions. Nevertheless, the syntactic function of this new argument encoding the secondary conjunct remains difficult to characterize. In this section, I conclude that perhaps the best analysis is to treat it as a second subject.

The status of the secondary conjunct as an argument is clearer in a number of Otomi languages whose verbs display a number of interesting asymmetries for the syntax of agreement. In such languages (most notably in all Otomi languages except Toluca Otomi and South Western Otomi), verbs must show resolved agreement in the split conjunction construction, but outside it no verb can agree with a 3rd person plural or dual subject. This is illustrated in the textual extract in (68) from Highlands Otomi. Example (68a) is an instance of severe split conjunction, and dual resolved agreement on the verb is obligatory.³⁹ In contrast, the verb cannot agree in number in the clause in (68b) with the 3rd person dual subject, only in person.

39. The dual marker =*wi* also encodes dual for the 1st and the 2nd person. The language has a dedicated morpheme to encode the dual of the 1st person exclusive. Because of this, when =*wi* is associated to the 1st person, it realizes an inclusive value.

- (68) Highlands Otomi (Otomi-Mazahua, Oto-Manguenan)
- a. [nu=na zi xitsu] i='b_ym-manho*(=wi)
 DEF=DEM.SG DIM woman 3.PRES.REAL=live.AS-good=DU
 [a rá ndo]
 DEF.SG 3SG.POSS man
 'The lady lives well with her husband.'
- b. hinjäm'bu di-n-tsu-il
 never [3.PRES.REAL]INFL-MIDDL-fight-FS/
 *di-n-tsu_h=mi
 [3.PRES.REAL]INFL-MIDDL-fight.AS=DU
 'They never fight.' (Echegoyen & Voigtlander 2007: 49)

As a consequence of these distributional asymmetries, the use of resolved agreement in the verb is a clear indication that the structure in question is an instance of split conjunction. In this respect, consider first example (69). This is the most natural rendering in Santa Ana Otomi for the semantics of the comitative structure in English, but it could equally be said to translate the semantics of the coordinative structure.

- (69) Santa Ana Otomi (Otomi-Mazahua, Oto-Manguenan)
- [a hä'i] bi=n-tsi-hme=wi[=nu]
 DEF.SG man 3.PST=ANTIP-ingest-tortilla=DU=DEM.SG
 'The man ate together with him'/'The man and him ate together' (Selene Hernández, personal communication)

Notice that the enclitic pronoun *nu* serves as anaphora to the secondary conjunct, hence the translation. The use of this anaphora is the most natural choice in the language, but a marker encoding 3rd person object anaphora is not strictly necessary. Consequently, example (70), although less natural, remains grammatical.

- (70) Santa Ana Otomi (Otomi-Mazahua, Oto-Manguenan)
- [a hä'i] bi=n-tsi-hme=wi _____
 DEF.SG man 3.PST=ANTIP-ingest-tortilla=DU
 'The man ate together with him'/'The man and him ate together'
 *'The man and some other (unidentified) person ate together.' (Selene Hernández, personal communication)

What is crucial for our purposes is that example (70) is not an instance of an associative construction of the same type we observe in Tzotzil or K'ichee' where the presence of the associative participant is not explicit.⁴⁰ The structure

40. On the other hand, an associative construction with a verb in the plural is found in Plains

in (70) is always interpreted with an elided secondary conjunct (this is indicated by the underscore). When the secondary conjunct is restored in an NP, it must occur in a different constituent than the primary conjunct, like in (71a). If the secondary conjunct was restored by means of the coordinative strategy in (71b), the verb does not show number agreement. If number agreement were to be used in the verb, the construction would require plural as in (71c). This reveals that the structure has an inherent split requiring the secondary conjunct as an argument, so that the conjuncts involve three participants now: the man and the priest, and another person.

- (71) Santa Ana Otomi (Otomi-Mazahua, Oto-Manguean)
- a. [*a hä'i*] *bi=n-tsi-hme=wi* [*a*
 DEF.SG man 3.PST=ANTIP-ingest-tortilla=DU DEF.SG
mbahä
 priest
 'The man ate together with the priest.'/'The man and the priest
 ate together.'
- b. [*a hä'i 'ne a mbahä*]
 DEF.SG man and DEF.SG priest
bi=n-tsi-hme(=wi)*
 3.PST=ANTIP-ingest-tortilla(=DU)
 'The man and the priest ate together.'
- c. [*a hä'i 'ne a mbahä*]
 DEF.SG man and DEF.SG priest
*bi=n-tsi-hme=hul*wi* _____
 3.PST=ANTIP-ingest-tortilla=PL/DU
 'The man and the priest ate with him.' (Selene Hernández, per-
 sonal communication)

The split between the conjuncts is so severe that the verb can even receive two distinct anaphoric pronouns for each conjunct, as in (72) from San Idefonso Otomi.

Cree, another Algonquian language like Passamaquoddy. This is illustrated in (i). Whether a similar construction exists in Passamaquoddy, I do not know, but if it does, it could account for the existence of the split construction in the same way as for the Mayan case presented above.

- (i) *ēkot[a] ānima Macōhōsis k̄ā-k̄ī-ayā-cik*
 there indeed Macohasis COMPL-PRF-live-3PL
 '... that, indeed, Macohasis and others used to live there' (Ahenakew & Wolfart
 (eds.) 1992: 373, Daniel & Moravcsik 2005, citing H. C. Wolfart, personal commu-
 nication)

- (72) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguean)
ha n-ja-hp-i
 how [3.PRES.REAL]MIDDL-make-3.DAT-FS
ga=ñöm= 'bé[=gá][=nú]
 1.PRES.IRR=speak.AS=DU.EXCL=1SG.EMPH=DEM.SG
 'How is it done so that I can talk with her?' (Palancar 2009: 587)

To account for these facts, I propose that severe split conjunction is a mechanism to promote the secondary conjunct to argument position. The question still remains of what syntactic role this new argument has.

In principle, it could be argued that it functions as an object. Such an account would have the advantage of making severe split conjunction typologically coherent with the peculiar abundance of comitative applicatives in the languages of North America which promote comitative adjuncts to object position (Stassen 2003). To North America we can add the languages of Mesoamerica as well, as comitative applicatives are widely found in the heavily synthetic families Mixe-Zoque and Totonac-Tepehua,⁴¹ for example in (73) from Upper Necaxa Totonac. But most notably, they are also found in a number of Oto-Manguean languages where applicatives are not that common. For the latter cases, the applicative is often a transparent emergent structure from the incorporation of the comitative coordinator, although not always. The case is illustrated in (74) from Zochina Zapotec, a VSO language.

- (73) Upper Necaxa Totonac (Totonac-Tepehua)
tu: ik-lakaskín na-[kin]-ta:-pína
 NON.HUM.REL 1SG.SUB-want FUT-1.OBJ-COMIT-go.2.SUB.IMPFV
 'What I want is for you to go with me.' (Beck 2004: 62)
- (74) Zochina Zapotec (Zapotecan, Oto-Manguean)
 a. *nhá b-yêgh [Pónch-é?] [lhénh Sàbàz-á?]*
 and COMPL-go.SG Alfonso-DET with Sabás-DET
 'And Alfonso went with Sabás.'
 b. *nhaá b-yêgh-lhénh [Pónch-é?] [Sàbàz-é?]*
 and COMPL-go.SG-COMIT Alfonso-DET Sabás-DET
 'And Alfonso went with Sabás.' (López 2009: 115)

In this respect, the secondary conjunct of severe split conjunction appears to share a number of properties with objects, although only indirectly.⁴² For

41. Recently, Brown et al. (2011) have proposed that Mixe-Zoque and Totonac-Tepehua are genetically related in a group they call "Totozoquean".

42. In the Algonquian languages, the secondary conjunct is always encoded as obviative. In these languages, when a given clause involves two 3rd person participants, they are ranked attend-

example, it occurs in the same syntactic position as an object. As an illustration, consider (75), where I show that in San Ildefonso Otomi no adjunct can intervene between the verb and the conjunct, itself a characteristic behavior of objects.

- (75) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguen)
- a. *hábu gí=pe=wí [ma nõnõ] M'ónda*
 where 2.PRES.IRR=go.DU=DU 1.POSS mother Mexico
 '... where you went with my mother to Mexico.'
- b. **hábu gí=pe=wí M'ónda [ma nõnõ]*
 where 2.PRES.IRR=go.DU=DU Mexico 1.POSS mother
 Intended reading: '... where you went to Mexico with my mother.'
 (Palancar 2009: 564)

Similarly, in both Algonquian and Otomi-Mazahua, the construction is only possible with intransitive verbs, as illustrated by the ungrammaticality of (76) and (77).⁴³ This restriction could be interpreted as resulting from the existence of a covert comitative applicative that turns the intransitive verb into a transitive one and treats the secondary conjunct as an object, banning verbs which are inherently bi-actantial.

- (76) Passamaquoddy (Algonquian, Algic)
- **[Mali] nomiy-a-wa-l [Piyel-ol] Susehp-ol Kehlis-k*
 Mary see-DIR-3PL-OBV Piyel-OBV Susehp-OBV Calais-LOC
 Intended reading: 'Mary and Piyel saw Susehp in Calais.' (Bruening 2005: 8)
- (77) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguen)
- **dá=tam=é [nó ma 'indõ]*
 1.PST-buy.AS[3.OBJ]=DU.EXCL DEF.SG 1.POSS brother.of.woman
 Intended reading: 'I bought it with my brother.' (Palancar 2009: 561)

ing to information structure principles so that the one ranking lower is marked as obviative, a grammatical person which is often treated as a 4th person. Objects often are obviative participants, as for example in (i) from Passamaquoddy. However, obviation is not a reliable diagnostic for objecthood, and thus cannot help to discern the role of the secondary conjunct in severe split conjunction. (I want to thank an anonymous referee for this observation.)

- (i) *koluskap 't-oli asitem-a-l nicalku-l*
 Koluskap 3-thus answer-DIR-OBV 3.uncle-OBV
 'Koluskap answers his uncle' (Bruening 2003: 5, citing Mitchell 1976: 7)

43. The only exception are a group of defective transitive verbs in Passamaquoddy which do not allow the direct/inverse indexing proper of more canonical transitive verbs, and which have a word order such as NP1 V NP2 OBJ.

In this connection for example, in the Mixean languages (Mixe-Zoquean), a Mesoamerican language family, comitative applicatives are only restricted to intransitive verbs (Zavala Maldonado 2012). This is illustrated in Olutec in (78). Compare (78a) with the impossibility of (78b), for which the only encoding alternative is (78c).⁴⁴

- (78) Olutec (Mixean, Mixe-Zoquean)
- a. *porke tan=mi:-kay-am=ak* [kumpa:ne]
 because 1.A.INDP=COMIT-eat-IRR.INDP=ANIM buddy
 ‘Because I’m going to eat with my buddy.’ (Zavala Maldonado 2000: 815)
- b. **?e:m+e=x?k ?i=mi:-kay-pe* *mixtun*
 skin=REP=ANIM 3.A.INDP=COMIT-eat-INCOM.INDP.TR cat
 [xu?ni]
 dog
 Intended reading: ‘The cat’s eating skin with the dog.’ (Zavala Maldonado 2012)
- c. *?e:m+e=xi=k ?i=kay-pe* *mixtun [mi:t*
 skin=REP=ANIM 3.A.INDP=eat-INCOM.INDP.TR cat with
xu?ni
 dog
 ‘The cat’s eating skin with the dog.’ (Zavala Maldonado 2000: 815)

However, an analysis of severe split conjunction based on a covert applicative would not account for the presence and relevance of verb agreement, which is so entrenched that it goes beyond the morphosyntax of affixal morphology. For example, a suppletive form of the intransitive verb ‘go’ must be used in the split construction in both these families when the subject is dual.⁴⁵ The selection of a lexical root can only be accounted for as motivated as a resolved agreement triggered by the subject.

- (79) Passamaquoddy (Algonquian, Algic)
 [*Mali*] *al-kawtuw-ok* [*Piyel-ol*]
 Mary around-walk.DU-3PL Piyel-OBV
 ‘Mary and Piyel are walking around.’ (Bruening 2005: 6)

44. The restriction does not exist in the Zoquean languages. In Mixean languages, a combination of the comitative and the instrumental applicative is used for transitive verbs. For more details, see Zavala Maldonado (2012).

45. The case at hand is root suppletion, not verbal number.

- (80) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguéan)
bi=me=wi[=r Zü'we]
 3.PST=SS/go.DU=DU=SG Devil
 'He went with the Devil.' (Palancar 2009: 567)

An alternative solution to account for the role of the secondary conjunct in severe split conjunction is to analyze it as a second subject. Such an analysis was first advanced in Bruening (2005), who treats it as a case of what he calls "split coordination" framed in the Principles and Parameters theory. The construction would involve generating the second subject in a second specifier of little vP by a type-shifting rule for multiple specifiers. In this account, the primary conjunct is generated in the first specifier of vP and is then moved to the topical position, whereas the NP generated in the second specifier (i.e., the secondary conjunct) is left *in situ*. The verb shows resolved agreement with both, because both are subject. The structure described is given in Figure 1.

The analysis of the secondary conjunct as a promoted second subject has two main advantages. It takes into account the argument status of the secondary conjunct and the different distributional properties associated with such status, and it accounts for the use of resolved verbal agreement.

In this connection, a conjunctive construction featuring severe split conjunction with a second subject is also found in Misantra Totonac, another American Indian language of Mexico. This language is unique within the Totonac family because it does not have a comitative applicative that promotes the secondary conjunct to object position, as is common in the rest of the family (cf. (73)). In the split construction of Misantra Totonac, the NPs realizing the

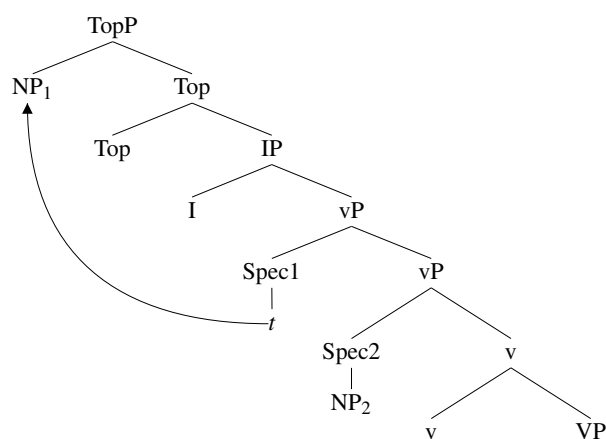


Figure 1. Generation of second subject for severe split conjunction

conjuncts also occur in different positions in the clause, the primary conjunct appearing in preverbal position. This is shown in (81).

- (81) Misantla Totonac (Totonac-Tepehua)
 [Pedro] *laa-laqaŋ-yaa-na* [Šwáan]
 Peter [3.SUB]REC-see[3.OBJ]-IMPFV-COMIT John
 ‘Peter saw it/(him/her/them) with John.’ (MacKay 1999: 300)

The main difference between the Algonquian/Otomi-Mazahua case and the construction in (81) lies in the fact that in Misantla Totonac the verb is marked with the comitative suffix *-na*, which anchors the conjunction relation between the two NPs,⁴⁶ but also the construction requires the reciprocal prefix *laa-* (MacKay & Trechsel 2003). MacKay (1999) claims that the two conjuncts are both treated as subject.⁴⁷ In most cases, the morphotactics of the verbal complex does not allow the co-occurrence of two subject pronominal indexes, and consequently only one subject affix is used. This only affix indexes the person ranking higher in the person hierarchy.⁴⁸ This is shown in (82) where the verb has the subject prefix for the 1st person. Notice, however, that the 1st person is treated by the syntax as the secondary conjunct (the idiomatic translation is mine; while the literal one is MacKay’s, which tries to reflect the complex semantics of the construction with the double subject).⁴⁹

- (82) Misantla Totonac (Totonac-Tepehua)
 [wiš] *ik-laa-qaala-yaa-na* [kit]
 you(SG) 1.SUB-COMIT-run-IMPFV-COMIT I
 ‘You run with me (lit., you run with me, I run with you).’ (MacKay 1999: 292)

There are also cases where the verb can reveal indexing of both subjects. This happens when one of the conjuncts is a 3rd person plural, as in (83a) and (83b). Here the verb indexes this participant by the prefix *ta-*, which indicates ‘plural subject’, as shown in (83c). Example (83a) further shows that transitive verbs can also be used in this construction, unlike the cases presented above in (74) and (75).

46. The suffix *-na* has morphotactic restrictions as it cannot co-occur with certain suffixes.

47. A similar phenomenon is found in Jacaltec (Mayan), where the preposition *-et* used to encode goals can be incorporated into the verb as a comitative applicative. In this construction, the secondary conjunct is not treated as a canonical object, but appears in topical position, see Craig (1977: 52–55) for more details.

48. With local participants, either 1st or 2nd person is marked.

49. In (80), the marking of a 3rd person subject is not obvious because it is encoded by means of a zero prefix or by the bare stem.

(83) Misantra Totonac (Totonac-Tepehua)

- a. [wiš_i] *ta_j-laa-č_{as}-t_i* [utun_j] *hun mesa*
 you(SG) SUB.PL-REC-lift[3.OBJ]-2.PFV they DET table
 ‘You lifted the table with them (lit., you lifted the table with them, they lifted the table with you).’ (MacKay 1999: 294)
- b. [kit_i] *ik_i-ta_j-laa-qawa-yaa-na* [utun_j]
 I 1.SUB-SUB.PL-REC-talk-IMPV-COMIT they
 ‘I talked with them (lit., I talk with them, they talk with me).’ (MacKay 1999: 288)
- c. *ta-laa-laqa_n-kan-yaa-tat*
 SUB.PL-REC-see-REF-IMPV-2PL.SUB
 ‘You (PL) see each other.’ (MacKay & Trechsel 2003: 283)

In this section, I have presented severe cases of split conjunction by integration where the conjunctive construction is not the extension of a counterpart comitative structure, and where the secondary conjunct works as an argument with the function of second subject. In functional terms, the secondary conjunct is a secondary topic, which can readily become an established topic for the following discourse. This is illustrated in (84). In (84a), the two participants are introduced by means of the split construction, while in (84b) they are both subject. However, the clause in (84c) has a singular 3rd person subject, whose reference is interpreted to be the secondary conjunct.

(84) Highlands Otomi (Otomi-Mazahua, Oto-Manguean)

- a. *di-m-bent*’=wi [rá mbare]⁵⁰
 [3.PRES.REAL]INFL-MIDDL-grasp=DU 3SG.POSS fellow.parent
 [a ra n’yohu]
 DEF.SG SG man
 ‘That man is fighting with his child’s godfather.’
- b. *ngetho i*=’iti
 because 3.PRES.REAL=get.drunk
 ‘Because they are drunk.’
- c. ’emme ra nexmate=’ä
 INT [3.PRES.REAL]SG provoker=3SG
 ‘He (the child’s godfather) is quite a provoker.’ (Echegoyen & Voigtlander 2007: 15)

Before concluding the article, a few observations about the semantics of split conjunction by integration are in order to compare the distribution of the construction in different languages.

50. The noun *mbare* is a kinship term that names the role the godfather of a child has with that child’s parents.

3.6. *The semantics of split by integration*

In general conjunctive constructions involving splits commonly predicate a state of affairs involving human beings only. For example, the use of a split construction is not natural if one of the conjuncts is a human but the other is an animal. For such situations a comitative strategy is widely preferred.⁵¹ The same is true about all types of splits by integration, mild and severe.

Nevertheless, exceptions to this generalization may occur, especially when the animal is part of a listing including other human beings, as for example in (85) from Highlands Otomi. Notice that to produce the listing, the secondary conjunct is discontinuous; that is, the other members in the list appear after the primary conjunct.

- (85) Highlands Otomi (VOS) (Oto-Manguéan)
gatho bi=mbahá [yó mbe]=[’u yə t’uhni]
 all 3.PST=SS/GO.PL 3PL.POSS mother=DEF.PL PL child
[’ne yó ta], [’ne=’u yó fo’yo buya]
 and 3PL.POSS father and=DEF.PL 3PL.POSS dog too
 ‘All the children went with their mothers, fathers, and also their dogs.’
 (Voigtlander & Echegoyen 2007: 240)

Only rarely does severe split conjunction involve an inanimate conjunct, and when it does, it almost always conveys a spatial ground, as shown in (86a). In other cases, the clause is used to describe the spatial disposition of an inanimate primary conjunct (86b). This does not happen in mild splits.

- (86) a. San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguéan)
[nǒ=r ndǒ] bi=n-tǒx=kwi[=r jǒdo]
 DEF.SG man 3.PST=MIDDL-bump=DU=SG wall
 ‘The man bumped against the wall.’ (Palancar 2009: 567)
- b. Misantra Totonac (Totonac-Tepehua)
[hun šaaluh] čunku† laa-wila-na [hun kaswela]
 DET pot next.to COMIT-seat.PFV-COMIT DET casserole
 ‘The pot is next to the casserole.’ (MacKay 1999: 299)

On the other hand, split conjunction by integration occupies an area in the semantic map of conjunctive semantics that lies in between the semantics conveyed by a coordinative structure at one end and the comitative structure at the other. Because of this, an instance of a construction bearing this type of split conveys semantic nuances that are in principle translatable by using

51. As expected, in fable stories or other special circumstances with human-like animal participants, the use of the split construction is more natural.

either structure in meta-languages like English. The choice of one construction over the other in the translation is at times random, but at times it also appears to respond to the feeling that the construction at hand is indeed used in a given language as semantically more equivalent to the nuances conveyed by the choice in the meta-language. If the comitative structure is used in the translation, the asymmetry in the topicality of the conjuncts is given prominence, while choosing a coordinative structure portrays the conjuncts with equal control. Constructions having splits by integration convey both these nuances at the same time.

While constructions with severe split conjunction may serve to convey typical comitative notions, such as actions or motion events that involve human participants acting together, some languages may take it further to the coordinative end of the spectrum and use it to convey stative situations which are only possibly encoded by means of typical coordination. This is illustrated for Passamaquoddy in (87), where the state of affairs described is predicated of both participants independently. In other words, the state of affairs does not happen only once, and certainly the participants do not have to be together for the situation to be true.⁵²

- (87) Passamaquoddy (Algonquian, Algic)
- a. *toqite* [Mali] *kisi-cipqahsin-uk* [Piyel-ol]
 both Mary PRF-have.nightmares-3PL Piyel-OBV
 ‘Both Mary and Piyel had nightmares.’
- b. [Susehp] *apolahsatpih-ik* [Piyel-ol]
 Susehp be.bald-3PL Piyel-OBV
 ‘Susehp and Piyel are bald.’ (Bruening 2005: 17)

In contrast to Algonquian, the Otomi-Mazahua languages represent the opposite case. Severe split conjunction in such languages cannot be used apart from prototypical instances of accompaniment – in other words, when both participants act as actors, not as undergoers, like in (88). For such situations, a standard coordinative structure would instead be used.

- (88) San Ildefonso Otomi (Otomi-Mazahua, Oto-Manguéan)
- a. **(nǒ=’ó)* *bi=dü=wí[=r* *’indö]*
 DEF.SG=3SG 3.PST=SS/die=DU=SG brother.of.woman
 Intended reading: ‘She and her brother died.’
- b. **(nǒ=’ó)* *bi=dang=wí[=r* *Pédro]*
 DEF.SG=3SG 3.PST=SS/fall.from.height.AS=DU=SG Peter
 Intended reading: ‘He and Pedro fell down.’ (Palancar 2009: 562)

52. Notice the use of the emphatic coordinator *toqite* ‘both’ in (87a), as appropriate in coordinative structures.

4. The whole picture

In this article, I first revisited well-established knowledge about the common strategies used crosslinguistically to conjoin NPs, restricting myself to cases where the conjuncts rank high in the animacy hierarchy and the resulting conjoined NP functions as subject. Given the inherent asymmetry of conjunction, I distinguished the conjuncts as primary and secondary according to their prominence both for the purpose of syntax and information structure.

In Section 3, I presented an overview of constructions involving NP conjunction which bear some type of a split in the phrasal continuity of the conjuncts while the verb shows resolved agreement with both conjuncts as a way to portray them with an equal rank as actors in the event. I have proposed that the splits are functionally motivated by a play of prominence of the conjuncts involved. In this regard, I have suggested that there are three main types of splits, i.e., splits by elision, by extraction, and by integration.

In splits by elision, the conjuncts form a constituent, but their phrasal continuity is disrupted because the primary conjunct is elided because of its topicality. As a consequence of this, the secondary conjunct is the only remaining element of the conjoined phrase on surface structure. In splits by extraction, the primary conjunct is given more prominence by being extracted or moved to a topical position, resulting in a surface structure where the two conjuncts are treated as two different constituents. Finally, in splits by integration, there is an upgrading of the secondary conjunct obtained by encoding it as a core participant of the action of the verb through resolved agreement. The conjuncts in splits by integration also appear in different positions in the clause. There are two subtypes of splits by integration. Those split constructions which are clearly based on a comitative structure existing in the language are treated as mild instances, but there are other, even rarer apparently, which function as fully independent constructs, where the NP encoding the secondary conjunct is treated as an argument of the verb.

Split conjunction involves all types of conjunctive structures, but to different degrees. Inclusive and summative conjunction commonly undergo splits by elision and by extraction. Splits by elision are also found in coordinative structures, but only rarely, as presented in Section 3.1.1, while splits by extraction are not found in coordination except in non-configurational languages. In contrast, splits by integration involve comitative structures at least for the mild subtype, and thus the relator involved is always a comitative one. Only in severe cases of split conjunction by integration, as found in a number of American Indian languages, such as Eastern Algonquian, Otomi-Mazahua, and Misantla Totonac, the second conjunct is an NP not linked by any coordinator. In such cases, following Bruening's (2005) account, the second conjunct could be treated as a second subject.

Constructions featuring splits by integration may also involve splits by elision responding to pressures of information structure. This interaction is mainly allowed because a split by elision is more superficial in the syntax, as it involves the surface effacing of the primary conjunct, more rarely the secondary one. In this regard, in languages having these splits it is relatively common to find surface manifestations of the kind in (89).

- (89) a. V+DU/PL [(COMIT) NP₂]
 b. V+DU/PL [___ (&COMIT) NP₂]
 c. ___ V+DU/PL [(COMIT) NP₂]

A structure like (89a) illustrates a split by elision of the primary conjunct, but the underlying structure is only revealed when the primary conjunct is restored, resulting in contiguous conjunction (inclusory or coordinative) as in (89b) or of a structure featuring split by integration as in (89c).⁵³ The latter case is illustrated from K'ichee' in (90) where example (90b) instantiates the situation in (89a) where the primary conjunct has been elided for being the established topic. The restoration of the primary conjunct is given in (90b') and follows the form of (89c).

- (90) K'ichee' (VOS) (Mayan)
- a. *x-Ø-'e* [*le achi*] *pa k'ayb'al*
 COMPL-3.ABS-go DET man LOCP market
 'The man went to the market.'
- b. *choqe' x-ee-chaku-n* [*r-uk'*] *le*
 and COMPL-3PL.ABS-work-ANTIP 3.POSS-and/with DET
nu-xb'aal
 1.POSS-brother
 'And worked with my brother.'
- b'. *choqe' [le achi] x-ee-chaku-n*
 and DET man COMPL-3PL.ABS-work-ANTIP
 [*r-uk'*] *le nu-xb'aal*
 3.POSS-and/with DET 1.POSS-brother
 'And the man worked with my brother.' (Telma Can, personal communication)

There are two further ways in which a conjunctive construction can undergo a split, and they must not go unmentioned. These are illustrated in Nakanai, an Oceanic language in (91), and may involve the extraction of the secondary conjunct to the right periphery of the clause as in (91a) or the extraction of both conjuncts independently to the external position of discourse

53. But notice the Nêlêmwa example in (35) above.

topic as shown in (91b), where a resumptive plural pronoun indexes both conjuncts as subject. The structure in (91b) can be further contrasted with the one in (91c), which illustrates the more canonical extraction of the whole (asyndetically) conjoined NP.

- (91) Nakanai (Oceanic, Austronesian)
- a. [*e hatavile*] *o_io po-pou*, [*la valalua*]
 NM women at.there RED-sit NM men
 ‘The women are there waiting, and the men.’ (Bril in press: 244
 citing Johnston 1980: 243)
- b. [*e hatavile*], [*la valalua*], *egite o_io po-pou*
 NM women NM men 3PL at.there RED-sit
 ‘The women and the men are there waiting.’ (Bril in press: 244,
 citing Johnston 1980: 243)
- c. [*e Marisa e latu-la*], *egira goata e Malalia*
 NM Marisa NM child-3.POSS 2DU go.up NM Malalia
 ‘Marisa and her child, they’ve gone up to Malalia.’ (Bril in press:
 214, citing Johnston 1980: 185)

Splits in NP conjunction are intriguing deviations from the canonical ideal of the symmetry of coordination structures, especially those involving coordinative structures, and deserve our attention. They are not common in Europe, but across the world’s languages, particularly in some language groups, they are resilient structures. As the achievements of the field of linguistic typology go hand in hand with the slow but extremely fruitful accomplishments in descriptive linguistics, the present study humbly hopes to serve as a map and inspire linguists to keep digging for more treasures in the deep sands of conjunctive strategies. Surprises are bound to await us.

Received: 21 July 2011

University of Surrey

Revised: 25 April 2012

Centre National de la Recherche Scientifique

Correspondence address: Surrey Morphology Group, Faculty of Arts and Human Sciences, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom; SeDyL-CELIA (UMR8202), Centre National de la Recherche Scientifique (CNRS), 7 rue Guy Moquet, Bat. D, 94801 Villejuif Cedex, France; e-mail: epalancar@hotmail.com

Acknowledgements: This article was written under the auspices of the ESRC/AHRC project RES-062-23-3126 “Endangered complexity: Inflectional classes in Oto-Manguean languages”. I heartily thank Greville G. Corbett and Isabelle Bril for their invaluable comments on an earlier draft of this article. I presented a version of the article as a talk at the Linguistics Department of the Karl-Franzens University of Graz on 9 June 2011, and I am grateful for all the comments received, but I want to especially thank Bernhard Hurch for his cordial invitation to Graz, which came as a blessing to me in very difficult times. I also want to thank Judith Aissen, Selene Hernández, Rosemary Beam, and Maris Camilleri for their wonderful help with Tzotzil, Santa Ana Otomi, Southern

Zapotec, and Maltese, respectively. I am indebted to Telma Can for her patience, generosity, and support with the K'iche' data. Jean-Léo Léonard and Gilles Polian confirmed my humble intuitions about resolved agreement in French for Footnote 26; Claudine Chamoreau provided me with just the right references. My thanks as well to Roberto Zavala, with whom I share the passion for this constructional labyrinth in the Mesoamerican linguistic landscape. Similarly, I am most grateful to the three anonymous referees of *Linguistic Typology* for their many comments and suggestions, especially to the first referee, whose rather insightful review inspired me to make a number of very important changes in both the presentation and the analysis which contributed substantially to the improvement of the article. Finally, I also want to thank Penny Everson for proof-reading the English in the text. All errors and deficiencies remain my responsibility.

Abbreviations: 1/2/3 1st/2nd/3rd person; A class A of pronominal indexes; ABS absolutive; ACC accusative; ACTFOC actor-focus; ANIM animate; ANPH anaphoric; ANT anterior; ANTIP antipassive; ART article; AS adjusted stem; ASS associative; AUX auxiliar; BAS basic cross-referencing; BS bound shape; CLF classifier; COMIT comitative; COMPL completive; CONF confirmation marker; CONJ conjunction; CONJTV conjunctive inflection; COP copula; DAT dative; DEF definite; DEM demonstrative; DET determiner; DIR directional/direct; DIM diminutive; DIST distal; DU dual; DUB dubitative; EMPH emphatic; ENC phrase final enclitic; ERG ergative; EXCL exclusive; F feminine; FACT factitive; FOC focus; FR free pronoun; FS free shape; FUT future; HAB habitual; HUM human; IMPF imperfect; IMPFV imperfective; IMPS impersonal; INCH inchoative; INCL inclusive; INCOM incompletive; IND indicative; INDP independent; INFL inflectional marker; INFR inferred; INT intensive; IRR irrealis; LK linker; LOC locative; LOCP locative preposition; M masculine; MIDDL middle; NARRFOC narrative focus; NEG negative; NF non-feminine; NM nominal marker; NOM nominative; NONFUT nonfuture; NONSG nonsingular; OBJ object; OBL oblique; OBV obviative; PART participle; PARTCL particle; PFV perfective; PL plural; POSS possessive; POSSD possessed; POT potential; PREP preposition; PRES present; PRF perfect; PROG progressive; PST past; PURP purposive; REAL realis; REC reciprocal; RED reduplication; REFL reflexive; REL relative; REM remote; REP reportative; SG singular; SS secondary stem; SUB subject; SUBJ subjunctive; TA tense/aspect; TOP topic; TR transitive; TRANSLOC translocative; VT versatile tense.

References

- Abdoulaye, Mahamane L. 2004. Comitative, coordinating, and inclusory constructions in Hausa. In Haspelmath (ed.) 2004, 165–197.
- Ahenakew, Freda & H. Christoph Wolfart (eds.). 1992. *Kōhkominawak otācimowiniwāwa – Our grandmothers' lives as told in their own words*. Saskatoon: Fifth House.
- Aikhenvald, Alexandra Y. 2003. *A grammar of Tariana*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2008. *The Manambu language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- Aissen, Judith L. 1989. Agreement controllers and Tzotzil comitatives. *Language* 65. 518–536.
- Allan, Keith. 1987. Hierarchies and the choice of left conjuncts (with particular attention to English). *Journal of Linguistics* 23. 51–77.
- Amfo, Nana Aba Appiah. 2010. Noun phrase conjunction in Akan: The grammaticalization path. *Pragmatics* 20. 27–41.
- Austin, Peter & Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language & Linguistic Theory* 14. 215–268.
- Barlow, Michael & Charles A. Ferguson (eds.). 1988. *Agreement in natural languages*. Stanford, CA: CSLI Publications.
- Beam, Rosemary. 2006. A Zapotec contribution towards the typology of inclusory constructions. Paper presented at the Annual Meeting of the Society for the Study of the Indigenous Languages of the Americas. Albuquerque, NM, 5–8 January.

- Beam, Rosemary. 2010. A Zapotec contribution towards the typology of inclusory constructions. Revised manuscript.
- Beck, David. 2004. *A grammatical sketch of Upper Necaxa Totonac*. München: Lincom Europa.
- Benmamoun, Elabbas, Archana Bhatia & Maria Polinsky. 2009. Closest conjunct agreement in head final languages. *Linguistic Variation Yearbook* 9. 67–88.
- Bhat, D. N. S. 2004. Conjunction and personal pronouns. In Haspelmath (ed.) 2004, 89–108.
- Black, Cheryl A. 2000. *Quiéjolani Zapotec syntax: A Principles and Parameters account*. Arlington, Texas: SIL International and the University of Texas at Arlington.
- Blake, Barry J. 2001. The noun phrase in Australian languages. In Jane Simpson, David Nash, Mary Laughren, Peter Austin & Barry Alpher (eds.), *Forty years on: Ken Hale and Australian languages* (Pacific Linguistics 512), 415–425. Canberra: Australian National University.
- Bogdanov, V. N. 1968. Osobyj slučaj dialektnogo soglasovanija skazuemogo s podležaščim po smyslu i kategorija predstavitel'nosti. *Naučnye doklady vyššej školy: Filologičeskie nauki* 4. 68–75.
- Borsley, Robert D. 2005. Against ConjP. *Lingua* 115. 461–482.
- Bresnan, Joan W. 2000. *Lexical-functional syntax*. Oxford: Blackwell.
- Bril, Isabelle. 2004. Coordination strategies and inclusory constructions in New Caledonian and other Oceanic languages. In Haspelmath (ed.) 2004, 499–533.
- Bril, Isabelle. 2010. Coordination inclusive et comitative dans les langues océaniques. In Franck Floricic (ed.), *Essais de typologie et de linguistique générale: Mélanges offerts à Denis Creissels*, 361–382. Lyon: ENS.
- Bril, Isabelle. 2011. 'And' and 'with' conjunctive strategies in some Austronesian languages: Syntax, semantics, pragmatics. *Language and Linguistics* 12(1). 239–272. Taipei: Academia Sinica.
- Bril, Isabelle (in press). Noun-phrase conjunction in Austronesian languages: Additive, inclusory and comitative strategies. In Claire Moyse-Faurie & Joachim Sabel (eds.), *Topics in Oceanic morphosyntax*, 211–263. Berlin: De Gruyter Mouton.
- Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters & Søren Wichmann. 2011. Totozoquean. *International Journal of American Linguistics* 77. 323–372.
- Bruening, Benjamin. 2003. Split coordination and successive-cyclic movement in Passamaquoddy. Paper presented at Western Conference on Linguistics of the Department of Linguistics at California State University. Handout at <http://udel.edu/~bruening/Downloads/WECOLHO4.pdf>
- Bruening, Benjamin. 2004. Verbal reciprocals and the interpretation of reciprocals. Invited paper at the Linguistics Department of the University of Chicago.
- Bruening, Benjamin. 2005. Split coordination in Passamaquoddy. Manuscript. <http://udel.edu/~bruening/Downloads/SplitCoord1.pdf>
- Camacho, José. 1996. Comitative coordination in Spanish. In Claudia Parodi, Carlos Quicoli, Mario Saltarelli & María Luisa Zubizarreta (eds.), *Aspects of Romance linguistics: Selected papers from the Linguistics Symposium on Romance Linguistics XXIV*, 107–122. Washington, DC: Georgetown University Press.
- Camacho, José. 2000. Structural restrictions on comitative coordination. *Linguistic Inquiry* 31. 366–375.
- Corbett, Greville C. 1983. *Hierarchies, targets and controllers: Agreement patterns in Slavic*. London: Croom Helm.
- Corbett, Greville C. 1988. Agreement: A partial specification based on Slavonic data. In Barlow & Ferguson (eds.) 1988, 23–54.
- Corbett, Greville C. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville C. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville C. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, Greville G. & Marianne Mithun. 1996. Associative forms in a typology of number systems: Evidence from Yup'ik. *Journal of Linguistics* 32. 1–17.

- Cormack, Annabel & Neil Smith. 2005. What is coordination? *Lingua* 115. 395–418.
- Craig, Colette Grinevald. 1977. *The structure of Jacalteco*. Austin, TX: University of Texas Press.
- Dalrymple, Mary & Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 4. 759–798.
- Daniel, Michael & Edith Moravcsik. 2005. The associative plural. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 150–153. Oxford: Oxford University Press.
- Dixon, R. M. W. 1977. *A grammar of Yidj*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: Chicago University Press.
- Echegoyen, Artemisa & Katherine Voigtlander. 2007. *Diccionario yuhú: Otomí de la Sierra Madre Oriental (Estados de Hidalgo, Puebla y Veracruz, México)*. Preliminary version. Ciudad de México: Instituto Lingüístico de Verano.
- Fabri, Ray. 1993. *Kongruenz und die Grammatik des Maltesischen*. Tübingen: Niemeyer.
- Feldman, Anna. 2001. Comitative and plural pronoun constructions. In Yehuda N. Falk (ed.), *Proceedings of Israel Association for Theoretical Linguistics 17*, The Hebrew University of Jerusalem. <http://linguistics.huji.ac.il/IATL/17/Feldman.pdf>
- Gaby, Alice. 2006. *A grammar of Kuuk Thaayorre*. Melbourne: University of Melbourne doctoral dissertation.
- Grevisse, Maurice & André Goose. 1993. *Le bon usage*. 13th edn. Liège: Duculot.
- Haiman, John. 1985a. Symmetry. In John Haiman (ed.), *Iconicity in syntax*, 73–95. Amsterdam: Benjamins.
- Haiman, John. 1985b. *Natural syntax*. Cambridge: Cambridge University Press.
- Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1. 5–47.
- Haspelmath, Martin. 2004. Coordinating constructions: An overview. In Haspelmath (ed.) 2004, 3–40.
- Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language typology and syntactic description* (2nd edn.), Vol. 2: *Complex constructions*, 1–51. Cambridge: Cambridge University Press.
- Haspelmath, Martin (ed.). 2004. *Coordinating constructions*. Amsterdam: Benjamins.
- Haviland, John B. 1979. Guugu Yimidhirr sketch grammar. In R. M. W. Dixon & Barry Blake (eds.), *Handbook of Australian languages*, Vol. 1, 26–180. Amsterdam: Benjamins.
- Heath, Jeffrey. 2008. *A grammar of Jamsay*. Berlin: Mouton de Gruyter.
- Hoffmann, Carl. 1963. *A grammar of the Margi language*. London: Oxford University Press.
- Johannessen, Janne Bondi. 1996. Partial agreement and coordination. *Linguistic Inquiry* 27. 661–676.
- Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.
- Johnston, Raymond L. 1980. *Nakanai of New Britain: The grammar of an Oceanic language* (Pacific Linguistics B-70). Canberra: Australian National University.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Lastra, Yolanda. 1989. *Otomí de San Andrés Cuexcontitlán, Estado de México* (Archivo de Lenguas Indígenas 13). Ciudad de México: El Colegio de México.
- Lichtenberk, Frantisek. 1983. *A grammar of Manam* (Oceanic Linguistics Special Publication 18). Honolulu: University of Hawaii Press.
- Lichtenberk, Frantisek. 2000. Inclusory pronominals. *Oceanic Linguistics* 39. 1–32.
- Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita, Volume II*. Berlin: Mouton de Gruyter.
- López Nicolás, Óscar. 2009. Construcciones de doble objeto en el zapoteco de Zoochina. Unpublished MA dissertation, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico.
- MacKay, Carolyn J. 1999. *A grammar of Misantra Totonac*. Salt Lake City: The University of Utah Press.

- MacKay, Carolyn J. & Frank R. Trechsel. 2003. Reciprocal /laa-/ in Totonacan. *International Journal of American Linguistics* 69. 275–306.
- McNally, Louise. 1993. Comitative coordination: A case study in group formation. *Natural Language & Linguistic Theory* 11. 347–379.
- Mitchell, Lewis. 1976. *Mikcic* ('Turtle'). Indian Township, ME: Wabnaki Bilingual Education Program. Edited and updated version of text in J. D. Prince (ed.), *Passamaquoddy texts* (Publications of the American Ethnological Society 10). New York: Stechert, 1921.
- Moravcsik, Edith A. 2003. A semantic analysis of associative plurals. *Studies in Language* 27. 469–503.
- Mosel, Ulrike. 1984. *Tolai syntax and its historical development* (Pacific Linguistics B-92). Canberra: Australian National University.
- Palancar, Enrique L. 2009. *Gramática y textos del hñöñhö: Otomí de San Ildefonso Tultepec*, Vol. 1: *Gramática*. Ciudad de México: Plaza y Valdés.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Quicoli, A. Carlos (undated). Conjunction in Micmac. Manuscript.
- Radford, Andrew. 1993. Head-hunting: On the trail of the nominal janus. In Greville Corbett, Norman Fraser & Scott McGlashan (eds.), *Heads in grammatical theory*, 73–113. Cambridge: Cambridge University Press.
- Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation. Published as *Infinite syntax!* Norwood, NJ: Ablex, 1986.
- Sadler, Louisa & Rachel Nordlinger. 2006. Apposition as coordination: Evidence from Australian languages. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG06 conference*. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/11/lfg06sadlernordlinger.pdf>
- Schwartz, Linda. 1985. Plural pronouns, coordination, and inclusion. In Nancy Stenson (ed.), *Papers from the Tenth Minnesota Regional Conference on Language and Linguistics*, 152–184. Minneapolis: Department of Linguistics, University of Minnesota.
- Schwartz, Linda. 1988a. Asymmetric feature distribution in pronominal ' coordinations '. In Barlow & Ferguson (eds.) 1988, 237–250.
- Schwartz, Linda. 1988b. Conditions for verb-coded coordination. In Michael Hammond, Edith A. Moravcsik & Jessica R. Wirth (eds.), *Studies in syntactic typology*, 53–73. Amsterdam: Benjamins.
- Sharp, Janet. 2004. *Nyangumarta: A language of the Pilbara Region of Western Australia* (Pacific Linguistics 556). Canberra: Australian National University.
- Singer, Ruth J. 2001a. The inclusory construction in Australian languages. *Melbourne Papers in Linguistics and Applied Linguistics* 1(2). 81–96.
- Singer, Ruth J. 2001b. The inclusory construction in Australian languages. Melbourne: University of Melbourne Honours thesis.
- Smeets, Ineke. 2008. *A grammar of Mapuche*. Berlin: Mouton de Gruyter.
- Smith, Marcus. 2003. A common structure for cross-linguistic coordination patterns. Paper at the 29th Meeting of the Berkeley Linguistics Society 29. <http://smithma.bol.ucla.edu/bls.pdf>
- Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4. 1–54.
- Stassen, Leon. 2001. Noun phrase coordination. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, Vol. 2, 1105–1111. Berlin: Mouton de Gruyter.
- Stassen, Leon. 2003. Noun phrase conjunction. In Frans Plank (ed.), *Noun phrase structure in the languages of Europe*, 761–817. Berlin: Mouton de Gruyter.
- Stassen, Leon. 2011. Noun phrase conjunction. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. München: Max Planck Digital Library, feature 63A. <http://wals.info/feature/63A>

- Stolz, Thomas. 2001. To be with X is to have X: Comitatives, instrumentals, locative and predicative possession. *Linguistics* 39. 321–350.
- Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2006. *On comitatives and related categories: A typological study with special focus on the languages of Europe*. Berlin: Mouton de Gruyter.
- Vassilieva, Maria Borisovna. 2005. *Associative and pronominal plurality*. Stony Brook, NY: State University of New York at Stony Brook doctoral dissertation.
- Voigtlander, Katherine & Artemisa Echegoyen. 1979. *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. Ciudad de México: Instituto Lingüístico de Verano.
- Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.
- Zavala Maldonado, Roberto. 2000. *Inversion and other topics in the grammar of Olutec (Mixean)*. Eugene, OR: University of Oregon at Eugene doctoral dissertation.
- Zavala Maldonado, Roberto. 2012. Instrumentos y comitativos en la familia mixezoque: Una exploración desde la diacronía y la tipología. Paper presented at the V Coloquio sobre Lenguas Otomangués y Vecinas (COLOV5), Oaxaca, 20–22 April.
- Zwart, Jan-Wouter. 2005. Some notes on coordination in head-final languages. In Jenny Doetjes & Jeroen van de Weijer (eds.), *Linguistics in the Netherlands*, 232–241. Amsterdam: Benjamins.

CHAPTER 37

VARIETIES OF
ERGATIVE

ENRIQUE L. PALANCAR

37.1 ERGATIVE CASE¹

ERGATIVE case – also ergative marker or simply ergative – is the term given to the grammatical morpheme associated with the noun phrase (NP) functioning as subject of a transitive clause (i.e. the A syntactic relation in Dixon 1994).² In semantic terms, ergatives mark NPs that typically play the role of agents in the transitive event rendered by such clauses. Two examples of this core function of ergative markers are given in (1):

- (1) a. Basque (LI) (Manandise 1988: 8; glosses adapted)
(Miren=e-k) liburu bat irakur-r-i d-u-Ø
 Mary=L-ERG book one.ABS read-L-PRF 3.ABS-UKAN('have')3SG.ERG
 'Mary has read a book'
- b. Tauya (P-TNG), (MacDonald 1990: 321)
(ʔe fanu-ni) fenaʔa-ra Ø-yau-a-ʔa
 DEM man-ERG woman-TOP 3SG-see-3SG-IND
 'The man saw the woman'

¹ This Chapter was written under the auspices of CONACyT Project No 47475. I want to thank Yuri Lander for all his comments on an earlier version of this chapter. To Even Hovdhaugen, for his support and inspiration while I worked on this topic in Oslo. I can be contacted at Facultad de Lenguas y Letras, Universidad Autónoma de Querétaro, CU, Querétaro 76010, Mexico, or at epalancar@hotmail.com.

² The term 'subject' is used here in a loose way.

The morphemes *=ek* in (1a) and *-ni* in (1b) are instances of ergative case by virtue of being morphemes associated with the NPs working as subjects in the transitive clauses in (1), which also express the agents of the events described.

37.1.1 Ergatives that are not ergative case

Ergative case is one of two possible morphological mechanisms to encode an ergative–absolute alignment. In such an alignment, A (transitive subject) is treated differently from S and O (intransitive subject and transitive object), which are commonly treated alike. As illustrated in (1), ergative case is used on the NP functioning as A; this is called a ‘dependent-marking’ strategy. The other way to encode such an alignment is using a so-called ‘head-marking’ strategy; that is, using a special pronominal morphology in the verb to cross-reference A, which is different from the one encoding S or O. Traditionally, this pronominal morphology has also been called ‘ergative’, but it should not be confused with ‘ergative case’. There are ergative languages that only use a head-marking strategy to encode their ergative–absolute alignment, for example Mayan languages, Mixean languages, and the Abkhaz-Abaza subgroup of North West Caucasian. An example of such a strategy appears in (2) in Mayan Awakatek (Mateo-Toledo et al., in press):

- (2) *ja* *Ø-s-tz'am* *(Wa'n)_A (tzee')_O*
 COMPL 3SG.ABS-3SG.ERG-GRASP John tree
 ‘John grasped the tree’

Notice that there is no ergative case on the subject NP *Wa'n* ‘John’. Reference to the A relation is only encoded in the verb with *s-*. Because of this, constructions of the type in (2) will not be regarded here as instances of ‘ergative case’.

37.1.2 Where is ergative case?

I base the observations in this chapter on the study of the behaviour of the ergative markers from about 140 languages. Much of the information comes from the analysis and understanding of the linguists who have described such languages. Ergative case is found in almost all ergative languages in Eurasia (N.B. the number of languages in parentheses represent the languages for which I had information about ergative case when compiling this chapter): in all Caucasian languages, except the Abkhaz-Abaza subgroup of North West Caucasian (18 languages); in the ergative languages of the Indo-Iranian subgroup of Indo-European (5); in all Tibeto-Burman languages (12); in a number of language isolates, such as Basque (1) and Sumerian (1); in a number of ergative Uralo-Yukaghir languages (1); and in Chukotko-Kamchatkan Chuckchee (1). In Oceania – including Papuan New Guinea, Australia, and the Pacific islands – it is found in many Papuan

languages (3); in all Australian languages that are ergative (16 Non-Pama Nyungan and 70 Pama-Nyungan; the sample comes mostly from Blake 1977); and in Austronesian (4 Oceanic). In the Americas, it is found in Eskimo (3), in a number of Zoquean languages from the Mixe-Zoque macrophylum (1), and recently, as descriptive work progresses in the area, it is being found in many South-American languages from various subgroups (Carib (2); Chibchan (1); Yanomam (1); and Panoan (1)). All in all, it appears that the number of ergative languages which employ ergative case as a morphological exponent of ergative-absolutive alignment far surpasses the number of languages which use head-marking patterns exclusively.

37.1.3 Formal varieties of ergative case

In phonological terms, ergative cases are in general highly grammaticalized morphemes: bisyllabic ergatives are scarce (e.g. both the allomorph *-njdja* of Djingili (AT-NPN) (Chadwick 1976) and Ngangikurrunggur *-ningi* (Hoddinott and Kofod 1976) should be considered rare typologically). Much more common are ergatives in /CV/ or /VC/, and even only /C/ (e.g. Kabardian (CA-NW) *-m*, Colarusso 1992); Basque (LI) = *k* (= *e-k*, after consonants); etc.). Interestingly, though perhaps irrelevant, ergatives in V also appear rarely (e.g. Khinalug (CA-E) *-i*, Kibrik 1994b).

Morphologically speaking, affixal ergative cases are more frequent than other morphological types, but this may just reflect the morphological typology of the language in question. When affixal, it is more common to find ergative cases of the agglutinative type than of the inflectional Indo-European-like type. In many languages of the Pacific, case is marked by adpositions, and thus ergative adpositions are found when the language has ergative-absolutive alignment (e.g. Tokelau (AU-Pol) *e te malo* [ERG SPC government] ‘the government’, Hovdhaugen et al. 1989: 49). Interestingly, nominal case in ergative languages is often encoded by means of phrasal clitics, and as clitics are syntactically-bound morphemes, this function fits the prototypical role of an ergative case. Consequently many ergative cases are clitics, and in this respect, ergative adpositions of the Polynesian type are syntactically equated with ergative clitics.

On the other hand, affixal or clitic-like, ergative allomorphy abounds, although it is commonly conditioned on morphophonological grounds. Nevertheless, lexical restrictions also occur. If the language allows ergative case with various types of nouns (animate vs. inanimate, etc.), the allomorphy is normally not conditioned by the semantic type of the noun; but in the Bzhedug and Shapsug dialects of Circassian Adyghe, proper nouns receive *-ə*, while common nouns receive *-m* (*Ahmed-ə* [Ahmed-ERG.PN] vs. *ps’as’e-m* [girl-ERG]; Zekokh 1969 in Kumachov et al. 1996). Similarly, the ergative case used with nominals is frequently the one

37.2 FUNCTIONAL VARIETIES OF ERGATIVE CASE

Besides encoding the subject of a transitive clause, ergative case also serves as grammatical exponency of an agent. The semantic role of agent may be defined as the participant that instigates and bears control upon a given action. In the prototypical case it is construed as a human being, and is thought to have acted intentionally and volitionally (for a full discussion see Palancar 2002). Expressing an agent is perhaps the foremost characteristic of ergative case, and as such it is exploited in sundry ways across ergative languages.

In a number of Australian and Papuan languages, ergative case is optional in transitive constructions (Blake 1977, Foley 2000). Some of these languages use it for pragmatic purposes to disambiguate the agent from the patient in a number of circumstances. In Dani (P-W), the use of ergative case on an agent noun is associated with events contrary to expectation (Foley 2000: 375). See (4a) where the ergative is not used on the noun *ap* 'man' because normally men eat pythons and not the other way around, as it is the case in (4b). The use of ergative case in (4a) would emphasize the agentive role played by the man:

- (4) a. (*ap*) *palu na-sikh-e*
 man python eat-RM.PAST-3SG.SBJV
 'The man ate the python'
 b. *ap (palu-nen) na-sikh-e*
 man python-ERG eat-RM.PAST-3SG.SBJV
 'The python ate the man'

In other languages where ergative case is grammatically obligatory in transitive constructions – besides possible aspectual splits that condition its application – the case may alternate with other encoding possibilities. The use of the ergative in such situations reinforces a reading of control, volition, intention, and/or responsibility. Citing various authors, Foley (2000) points out that the use of ergative case in Papuan languages often contributes to some subtle meaning where actors are depicted as wilful agents. Foley (2000: 375) mentions that 'in Fologa (Anderson and Wade 1988), certain verbs construed as agentive and willful, like "kill", always co-occur with subjects in ergative case, whereas non-agentive verbs, like "like", never do so. In between are verbs that can be construed either way, e.g. "do/say", "get"'. The same is true of Indo-Aryan languages (IE) like Urdu/Hindi with infinitival constructions (Butt and King 1991, Butt 2005: 15), as in (5):

- (5) (*nadya=ne/=ko*) *zu ja-na he*
 Nadya.F.SG=ERG/=DAT zoo.M.SG go-INF.M.SG be.3SG.PRES
 'Nadya wants to go to the zoo' (Erg=Control / Dat=Experiencer)

Along these lines, it is very common that languages use ergative case on the NPs functioning as subject of intransitive (unergative) verbs to render the actors as volitional or having control. This encoding possibility gives rise to the emergence of active/patient systems (Mithun 1991), and it is found across a number of Tibetan, Indo-Aryan (Kachru 1987), and Eastern Caucasian languages.

37.2.1 Ergative case polysemy patterns

Ergative case is often used across languages to express other semantic roles. Nevertheless, in many languages ergative case expresses nothing else but the agent in a transitive construction. I call the latter cases ‘asyncretic’ and the former ‘syncretic’. In principle, there is nothing typologically odd about asyncretic ergatives, as they are found in all families. Figures may give a useful indication: twenty-one out of a total of eighty-eight ergative markers in Australian languages are asyncretic (24 per cent). The same is true of Tibeto-Burman (four asyncretic vs. eight syncretic) and in Caucasian where the proportion is larger (seven asyncretic vs. twelve syncretic). At times we find the vestiges of old syncretisms that have long vanished. Basque asyncretic ergative =*k* used to encode cause with intransitive predicates. This old causal use still survives in sporadic, formal, idiomatic expressions (e.g. *hotz=a=k n-en-go-en* [cold=DEF=ERG 1SG-PAST-EGON ‘be.located’-PST] ‘I was cold’) and in the lexicalized adjective *hotzakil* ‘very cold’ (lit. ‘dead from cold’ *hotz-a=k (h)il* [cold=DEF=ERG die.PRF]).

37.2.1.1 Common syncretic patterns

Among syncretic instances, instrument is by far the most common semantic category found in ergative syncretisms. Figures speak for themselves: all syncretic ergative markers in Australian also express instrument (save the Djingili (AT-NPN) allomorph *-njdja*, which is better seen as an oblique). The pattern is found in the vast majority of East Caucasian languages (Lak being an exception; N.B. North West Caucasian languages have a separate Instrumental marker, and South Caucasian ergatives are asyncretic). It is also found in Indo-Aryan Brokskat, Konkani, and Marathi, but not in Urdu/Hindi, which has its own Instrumental. It is widespread across Tibeto-Burman and in Papuan languages, and a number of South-American languages have it too, for example, Shipibo-Konibo and Sanumá (Borgman 1990).

Occasionally, when an ergative expresses instrument, it may also serve to express cause in intransitive clauses (e.g. ‘they died from hunger’). The pattern is rather scattered, but appears in a number of East Caucasian, Tibeto-Burman, Papuan, and Indo-Aryan languages, as well as in isolated cases. It is illustrated in (6) from Sanumá (MJ) (Borgman 1990):

- (6) a. Agent (*kamisamakö-nö*) *hama sam_atöpö se kite*
 we.EXCL-ERG/AGEN visitor 1PL.EXCL hit FUT
 ‘We will hit the visitors’ (p. 29)
- b. Instrument (*kusiali a-nö*) *sa ia pia kule*
 spoon 3SG-ERG/INS 1SG eat intend PRS
 ‘I am about to eat with a spoon’ (p. 123)
- c. Cause (*kamali te wasu-nö*) *ipa ulu a*
 high.fever 3SG deadly-ERG/CAUS my son 3SG
noma-so-ma
 die-FOC-COMPL
 ‘My son died from a deadly high fever’ (p. 123)

Possessor is another semantic category that ergative markers also express cross-linguistically, but in overall frequency, the pattern is much less common than the syncretism with instrument. Ergative markers expressing a possessive relation are typical of Eskimo and Tibeto-Burman languages. It is illustrated in (7) in Ladakhi (TB) (Koshal 1979). The pattern is sporadically found elsewhere, as in Caucasian Lak (Van den Berg 2005) or in Mixe-Zoquean Chiapas Zoque (Faarlund p.c.). (N.B. A similar head-marked pattern is found in Mixean and Mayan languages.)

- (7) a. Agent (*thug-gu-yi*) *pəl-ldən-ni kə-ne şpe-čhə*
 boy-ABS-ERG/AGEN Paldan-GEN from-ABL book-ABS
khyers
 take.SIMPL-PRF
 ‘The boy took the book from Paldan’ (p. 75)
- b. Possessor (*khyi-yi*) *şŋə-mə-rin-mo duk*
 dog-ERG/GEN tail-ABS-long-ABS be.PRS
 ‘The dog’s tail is long’ (p. 74)

However, it is typologically rare to find an ergative case that expresses both instrument and possessor at the same time, only in some Tibeto-Burman languages is this pattern possible (e.g. Athpare *-ŋa* (Ebert 1997b), Limbu *-le* (Van Driem 1987), and Tibetan *-s* have it (Denwood 1999)). (N.B. Panoan Shipibo-Konibo has it too; Valenzuela 1997.)

37.2.1.2 Uncommon ergative syncretisms

Other syncretic patterns are less common or reveal syncretic phenomena that are yet to be uncovered in undescribed languages. In North West Caucasian languages with nominal case, the ergative case has been treated as an ‘oblique’ marker because of its divergent functions. Apart from their locative functions, which other syncretic ergative markers also have (see next section), in these languages ergatives are the case governed by adpositions; the case that marks the possessor in head-marked

possession relations; they cover adverbial functions with verbs (see 3); and they may even express a dative participant, as well as other demoted patients. The last two options are illustrated in (8) in Kabardian (CA-NW):

- (8) a. (*l'ə-m*) *tχətə-r* (*fəzə-m*)
 man-ERG/AGEN book-ABS woman-ERG/DAT
Ø-jə-ri-t-a-s'
 3SG.ABS-3SG.DAT-3SG.ERG-give-PRF-ASSRT
 'The man gave the book to the woman' (Kumachov et al. 1996: 100)
- b. *ɕ'a:lɔ-r* (*tχətə-m*) *Ø-j-aw-dʒa-Ø*
 boy-ABS book-ERG/DEMOTED.PAT he-it-DYN-read-INTR.PRES
 'The boy is reading the book' (Hewitt 2005: 124)

A somewhat similar phenomenon is found in Carib Tiriyo (Meira 1999), where the ergative marker is used to encode a dative participant as well as the causee in causative constructions.

37.2.1.3 *Local syncretisms*

Syncretic ergative markers often encode spatial categories. Ergatives in Australian languages may encode locative as well as instrument (seventeen out of a total of sixty-seven syncretic instances in sixty-five languages), but in this linguistic area, no marker appears to encode agent and locative alone, without instrument. The same pattern is found in Shipibo-Konibo (Pan), in Vach Khanty (UY), and in Manipuri Meithei (TB). All ergative (oblique) markers in North West Caucasian languages also express a locative relation. After locative, ergatives encoding spatial source (ablative) also occur, but more infrequently; they are especially typical of Tibeto-Burman and Papuan languages. As with the locative, all such markers encode instrument and cause as well (e.g. Dani (P-W) *-(n)en* (Bromley 1981); Tauya (P-TNG) *-ni* (MacDonald 1990); Athpare (TB) *-ŋa* (Ebert 1997b); Thakali (TB) *-se* (Georg 1996)). Other spatial cases are rare. Perlative ('through/by the park') is found in Indo-Iranian Konkani (Maffei 1986) and Marathi (Kashli Wali p.c.), but outside Indo-Iranian this syncretism is not found, as with the allative, which appears circumscribed to South-American languages.

37.2.1.4 *Non-existent syncretic patterns*

At this stage of the question and with scanty data, it is unwise to make any conclusive serious statement about the typological absence of certain syncretisms. However, one may advance at least one: even though the syncretism comitative-instrument is common cross-linguistically for instrumental cases, the comitative is

not found in ergative syncretisms that encode instrument or otherwise (Stolz 1996b, Palancar 2002).³

37.2.2 The origin of ergatives

What do these syncretic patterns reveal about ergative case? Ergativity has long been seen in linguistics as an emergent structure that came about through the reanalysis of a nominative–accusative alignment. In this view, passive has been believed to be the most common path for this reanalysis, especially for the Indo-Iranian, Polynesian, and Eskimo ergative languages (for a full discussion on the matter see Palancar 2002). However, the evidence provided is not fully convincing, and there are many who remain sceptical. Recently, Butt and King (1991, 2003, 2004) have challenged such a view with convincing arguments, for example the old Sanskrit inflectional instrumental with multiple allomorphy used with demoted agents in passive-like structures could not possibly have given rise to current ergative markers, as these are of recent development, and behave like clitics. Butt and King’s solution is that ergative case =*ne* in Urdu/Hindi stems from a grammaticalized adposition *janiyē* ‘for the sake of, because of’, which was used in the past tense to emphasize semantic agentivity. All in all, ergative syncretisms are often taken to shed light on the possible emergence of the alignment they manifest. Palancar (2002) treats ergative syncretisms involving instrument (and cause) as reflecting a more abstract energetic category he calls ‘energiser’, much along the lines of the ‘effector’ category in Van Valin and Wilkins (1997). Such a view proves more convenient than explaining which was first, instrument or agent. The oblique ergatives of the Caucasian type reflect very old markers whose history is difficult to trace. As for the spatial cases in the syncretisms, the spatial categories involved (mainly source and location) served as a model to metaphorically construe other more abstract, semantic categories such as agent/instrument/cause in order to foreground the energetic role played by such participants in the clausal event.

ABBREVIATIONS (LANGUAGES, LANGUAGE GROUPS, AND REGIONS)

AT	Australian
AU	Austronesian
CA	Caucasian

³ The apparent exception to this rule is found in Circassian languages. As the ergatives in these languages are really oblique markers, an NP functioning as comitative can also receive this case providing the relevant cross-referencing morphology appears in the verb. I thank Yuri Lander for this observation.

E	East
IE	Indo-European
LI	Language Isolate
MJ	Macro-Jê
NPN	Non-Pama-Nyungan
NW	North West
P	Papuan
Pan	Panoan
Pol	Polynesian
TB	Tibeto-Burman
TNG	Trans-New-Guinea
UY	Uralo-Yukaghir
W	West

5. Travaux sur la morphologie de langues otomangues

- TRAVAIL 1. Palancar, Enrique L. et Heriberto Avelino. A paraître. 547
 Inflectional complexity and verb classes in Chichimec. Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguéan languages*. (Morphological Investigations Series 1). Berlin: Language Science Press
- TRAVAIL 2. Palancar, Enrique L. et Leonardo Carranza. A paraître. 581
 Inflectional class interaction in Matlatzinca. Enrique L. Palancar, Timothy Feist et Matthew Baerman (eds.), *Inflectional class complexity in the Oto-Manguéan languages*. (Morphological Investigations Series 1). Berlin: Language Science Press
- TRAVAIL 3. Timothy Feist et Enrique L. Palancar. 2016. Tracing the 607
 emergence of inflectional tone in Cuicatec. Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection: New data and new perspectives*, pp. 267 – 294. Berlin: De Gruyter Mouton
- TRAVAIL 4. Palancar, Enrique L., Jonathan Amith et Rey Castillo 635
 García. 2016. Verbal inflection in Yoloxóchitl Mixtec. Enrique L. Palancar et Jean Léo Léonard (eds.), *Tone and inflection: New data under new perspectives*, pp. 295 – 336. Berlin: De Gruyter Mouton
- TRAVAIL 5. Palancar, Enrique L. 2015. A mixed system of agreement in 677
 the suffix classes of Lealao Chinantec. *Morphology* 25, 29 – 62.
- TRAVAIL 6. Palancar, Enrique L. et Timothy Feist. 2015. Agreeing with 711
 subjects in number: The rare split of Amuzgo inflection *Linguistic Typology* 19:3, 337 – 383.
- TRAVAIL 7. Baerman, Matthew et Enrique L. Palancar. 2014. The 759
 organization of Chinantec tone paradigms. *Carnets de Grammaire* 22: 46 – 59.
- TRAVAIL 8. Palancar, Enrique L. 2014. Revisiting the complexity of the 773
 Chinantecan verb conjugation classes. Jean Léo Léonard et Alain Kihm (eds.), *Patterns in Mesoamerican morphology*, pp. 77 – 102. Paris : Michel Houdiard.
- TRAVAIL 9. Palancar, Enrique L. 2012. The conjugations classes of 787
 Tilapa Otomi: An approach from canonical typology. *Linguistics* 50 – 4: 783 – 832.
- TRAVAIL 10. Palancar, Enrique L. 2011. The conjugations of Colonial 837
 Otomi. *Transactions of the Philosophical Society* 109: 1 – 19.

11. Inflectional complexity and verb classes in Chichimec

Enrique L. Palancar

SeDyL (UMR8202), CNRS

Heriberto Avelino

Abstract. *In this chapter, we propose that in the inflectional morphology of Chichimec, verbs can be classified attending to two different subsystems. One attending to the prefix set they select for the realization of notions such as tense/aspect/mood/polarity and person of the subject, and another attending to the type of stem alternation pattern they display. As a result of the interaction of these two inflectional subsystems, one obtains a very complex morphological system which is endemic in the Oto-Pamean branch of Oto-Manguean.*

17. Introduction

Chichimec (also known as Chichimeco-Jonaz or as ?Uzà? by its speakers) is an Oto-Manguean language of Mexico spoken in the community of Misión de Chichimecas in the town of San Luis de la Paz, in the state of Guanajuato (see Figure 1). In this paper, we show that the verbal inflection of Chichimec displays an interesting degree of complexity. Such complexity is created by the intersection of three different inflectional sub-systems. One such system is the one formed by the inflectional classes of prefixes encoding person of subject and tense/aspect/mood. A second and partially independent sub-system involves a rich array of stem alternation patterns while a third one involves the inflectional use of tone, which is, for the most part, lexically specified.



Figure 1: Location of Chichimec in Mexico.

We base our analysis on a revision of a seminal paper published in 1933 by Jaime de Angulo by studying patterns in the data in Feist and Palancar (2015). In his paper, Angulo presented a grammatical sketch of the Chichimec language which still remains the finest description of the language to date. He proposed the existence of five inflectional classes for verbs based on the selection of different sets of prefixes, but the descriptive material also includes information about other aspects of Chichimec verbal inflection such as stem alternations and tone. While our analysis is based on Angulo (1933), we also use the materials in Lastra (1984), and our own observations from the field, but only for the study of the prefix classes as the information involving other aspects of verbal inflection is scanty.

While the data in Angulo are of excellent quality, both prefix classes and stem patterns are presented in a scattered manner throughout his paper. For example verbs exhibiting different stem patterns may be presented in the same list, while others having the same pattern appear in different lists. This prevents the reader from gaining a comprehensive understanding of the inflectional system of Chichimec verbs, of the actual number of patterns we find and of their internal structure. The goal of this paper is precisely to fill this gap in the literature and to make the morphological structure of the verbal inflection of Chichimec more readily accessible. Furthermore, Angulo's paper does not address other issues which may be of interest to the modern reader, such as, for example, what the connection is between prefix classes and stem alternation patterns. In this paper, we also go beyond Angulo's (1933) analysis in providing cues as to the phonetic/phonological principles which are used to build the different stems we observe as well as to the internal structure of the stem zones in the paradigm where those stems are employed.

Chichimec and Pame form the so-called Pamean branch of Oto-Pamean, itself the northernmost branch of Oto-Manguéan. According to the most recent census by the National Geographical Institute of Mexico (INEGI), there were 1,362 speakers of the language in 2005. However, this figure is an oversimplification as it does not distinguish between fluent speakers and those with passive competence. From fieldwork observation, we estimate that the present day number of Chichimec speakers is no more than 1,000.

The structure of the paper is straightforward. In the next section, we introduce our analysis of the inflectional classes we can see in Chichimec verbs when one attends to prefix set selection. Where possible we also advance an indication as to the principles behind class membership assignment. In section 3, we study the system of stem alternation patterns and we put them in connection with the inflectional sub-system involving prefix class. In section 4, we introduce the different tone patterns we have been able to discern in Angulo's data which run partially independently from the other two sub-systems. In section 5, we finish with a discussion of our findings and make suggestions for further research.

18. Inflectional classes according to prefix set

Verbs in Chichimec inflect for tense/aspect/mood (TAM) and the person of the subject by means of cumulative prefixes. The examples in (1) illustrate the so-called ‘anterior past’ for the persons of the singular.¹¹⁵

- (1) a. *tu-nú*
 1.ANT.PST-see[3OBJ]
 ‘I saw it.’
- b. *ki-nú*
 2.ANT.PST-see[3OBJ]
 ‘You saw it.’
- c. *u-nú*
 3.ANT.PST-see[3OBJ]
 ‘S/he saw it.’

Dual and plural of subject is conveyed by number suffixes. This is indicated in Table 1 where the singular forms in (1) are also given for comparison. For the inflection of most verbs, however, there is a dedicated plural stem for a 3rd person plural subject, e.g. the verb in Table 1 has *-nú* for all cells except for the 3rd person plural where it requires *-nhú* without further affixation.

	SG		DU		PL
1 st	<i>tu-nú</i>	EXCL	<i>tu-nú-m^p</i>	EXCL	<i>tu-nú-hú</i>
		INCL	<i>tu-nú-s</i>	INCL	<i>tu-nú-n</i>
2 nd	<i>ki-nú</i>		<i>ki-nú-s</i>		<i>ki-nú-n</i>
3 rd	<i>u-nú</i>		<i>u-nú-s</i>		<i>u-nhú</i>

Table 1: Anterior past tense of the verb *-nú* ‘see’.

As the forms for the dual and plural are transparently derived from those of the singular, at least for most verbs,¹¹⁶ the forms for the singular together with the form for the 3rd person plural could be taken as sufficient to construct the entire paradigm of a verb. The basic cells of the paradigm of the verb *-nú* ‘see’ are given in Table 2, slightly modified from the original in Angulo (1933: 165). There are eight TAM categories: a present or incomplete, three past tenses (anterior past, recent past and immediate past), two irrealis values (future and potential) a mood called ‘contemporaneous’ that we prefer to call ‘sequential’, and a negative form used regardless of tense or aspect. In all the tables, we introduce the forms for the present, the anterior past and the future together. While we are aware that this grouping may not make sense from a semantic point of view, it is nonetheless a convenient way to

¹¹⁵ We follow most of Angulo's transcription conventions: $v = [\beta]$, $tc = [\widehat{t}]$, and $\ddot{u} = [y]$.

¹¹⁶ Other verbs like the ones treated in §2.1.5-2.1.7 require a different prefix for the 1st person plural and at times for the dual too.

capture a number of formal similarities that these three tenses have in common in the inflection of Chichimec.

	PRS	ANT.PST	FUT	REC.PST	IMM.PST	POT	SEQ	NEG
1SG	<i>e-nú</i>	<i>tu-nú</i>	<i>ga-nú</i>	<i>ku-nú</i>	<i>u-nú</i>	<i>nu-nú</i>	<i>ra-nú</i>	<i>su-nú</i>
2SG	<i>ki-nú</i>	<i>ki-nú</i>	<i>ki-nú</i>	<i>ki-nú</i>	<i>i-nú</i>	<i>mi-nú</i>	<i>gi-nú</i>	<i>si-nú</i>
3SG	<i>e-nú</i>	<i>u-nú</i>	<i>ga-nú</i>	<i>ku-nú</i>	<i>zu-nú</i>	<i>mu-nú</i>	<i>ru-nú</i>	<i>su-nú</i>
3PL	<i>e-nhú</i>	<i>u-nhú</i>	<i>ga-nhú</i>	<i>ku-nhú</i>	<i>zu-nhú</i>	<i>mi-nhú</i>	<i>ru-nhú</i>	<i>su-nhú</i>

Table 2. Basic paradigm of the verb *-nú* ‘see’.

18.1. Prefix classes

Not all verbs in Chichimec select the same prefix set as does *-nú* ‘see’ in Table 2. In total there are seven possible sets. Let us see them one by one.

18.1.1. Class I verbs

Most verbs in Chinantec behave like the verb *-nú* ‘see’ in Table 2: 85 out of 168 in our sample, 50.5%. We consider that such such verbs form Class I. The prefixes of this class are given in Table 3. The prefixes are not sensitive to the number of the subject, i.e. the same prefix is used for 1st person singular, plural or dual. The only exception involves the 3rd person in the potential, where there is a different prefix for 3rd singular and dual from 3rd person plural. We arrange the prefixes by person value in a way that highlights the occasional syncretism between the 1st person and the 3rd person. This is important to give an idea of the structure of the paradigm. Notice as well that the prefix *ki-* for the 2nd person is used in several tenses.

	1	3SG/DU	3PL	2
PRS	<i>e-</i>	<i>e-</i>	<i>e-</i>	<i>ki-</i>
ANT.PST	<i>tu-</i>	<i>u-</i>	<i>u-</i>	<i>ki-</i>
FUT	<i>ga-</i>	<i>ga-</i>	<i>ga-</i>	<i>ki-</i>
REC.PST	<i>ku-</i>	<i>ku-</i>	<i>ku-</i>	<i>ki-</i>
IMM.PST	<i>u-</i>	<i>zu-</i>	<i>zu-</i>	<i>i-</i>
POT	<i>nu-</i>	<i>mu-</i>	<i>mi-</i>	<i>mi-</i>
SEQ	<i>ra-</i>	<i>ru-</i>	<i>ru-</i>	<i>gi-</i>
NEG	<i>su-</i>	<i>su-</i>	<i>su-</i>	<i>si-</i>

Table 3: Prefixes selected by Class I verbs.

After Angulo (1933), Lastra (1984) is the next published reference on Chichimec verbal inflection. In her work, Lastra presents a schematic representation of the prefix classes, but sometimes it is not clear to what extent her data are based on an interpretation of Angulo's paradigms or on her own synchronic observations of Chichimec as it was spoken in the late 70's; about the time when she carried out most of her fieldwork activities for that paper. Our Class I is called ‘Class 1’ by Lastra too. The only difference between Angulo's account and Lastra's lies in the fact that Lastra reports a prefix *ra-* for the 3rd person consecutive rather than *ru-*. This appears to have resulted from a leveling from the 1st person.

18.1.2. Class II verbs

A first deviation from Class I is found in verbs that select the set in Table 4, which we call Class II. They differ from Class I verbs in several respects: (i) the prefixes for the anterior past are also used for the present, at least for the 1st and the 3rd person; (ii) there is a dedicated prefix for the 2nd person present; (ii) and there is a distinctive prefix for the 1st person future.

	1	3SG/DU	3PL	2
PRS	<i>tu-</i>	<i>u-</i>	<i>u-</i>	<i>su-</i>
ANT.PST	<i>tu-</i>	<i>u-</i>	<i>u-</i>	<i>ki-</i>
FUT	<i>gu-</i>	<i>ga-</i>	<i>ga-</i>	<i>ki-</i>
REC.PST	<i>ku-</i>	<i>ku-</i>	<i>ku-</i>	<i>ki-</i>
IMM.PST	<i>u-</i>	<i>zu-</i>	<i>zu-</i>	<i>i-</i>
POT	<i>nu-</i>	<i>mu-</i>	<i>mi-</i>	<i>mi-</i>
SEQ	<i>ra-</i>	<i>ru-</i>	<i>ru-</i>	<i>gi-</i>
NEG	<i>su-</i>	<i>su-</i>	<i>su-</i>	<i>si-</i>

Table 4: Prefixes selected by Class II verbs.

A sample paradigm from Class II is given in Table 5 with the verb *-tsá* ‘do’ (Angulo 1933:171). Notice that in contrast with the verb *-nú* ‘see’ in Table 2, this verb has three different stems: *-tshá* for the 3rd person plural, *-tcá* for the 2nd person and *-tsá* for the rest of the cells. As we will show in §3.8, however, not all Class II verbs behave like this.

	PRS	ANT.PST	FUT	REC.PST	IMM.PST	FUT	POT	SEQ	NEG
1SG	<i>tu-tsá</i>	<i>ku-tsá</i>	<i>ku-tsá</i>	<i>gu-tsá</i>	<i>u-tsá</i>	<i>gu-tsá</i>	<i>nu-tsá</i>	<i>ra-tsá</i>	<i>su-tsá</i>
2SG	<i>su-tcá</i>	<i>ki-tcá</i>	<i>ki-tcá</i>	<i>ki-tcá</i>	<i>i-tcá</i>	<i>ki-tcá</i>	<i>mi-tcá</i>	<i>gi-tcá</i>	<i>si-tcá</i>
3SG	<i>u-tsá</i>	<i>ku-tsá</i>	<i>ku-tsá</i>	<i>ga-tsá</i>	<i>zu-tsá</i>	<i>ga-tsá</i>	<i>mu-tsá</i>	<i>ru-tsá</i>	<i>su-tsá</i>
3PL	<i>u-tshá</i>	<i>ku-tshá</i>	<i>ku-tshá</i>	<i>ga-tshá</i>	<i>zu-tshá</i>	<i>ga-tshá</i>	<i>mi-tshá</i>	<i>ru-tshá</i>	<i>su-tshá</i>

Table 5: Basic paradigm of the verb *-tsá* ‘do’.

18.1.3. Class III verbs

A few verbs, which we treat as Class III, select a prefix set very similar to Class II, but with two further idiosyncrasies: (i) they completely neutralize the formal distinction between present and anterior past by using the same markers for all persons; and (ii) they require a specific prefix *e-* for a 3rd person plural of the anterior past and the present. The forms are given in Table 6, with a sample paradigm in Table 7 with the verb *-mén* ‘like, love’. Like the verb *-nú* ‘see’ in Table 2, this verb also has a special stem for the 3rd person plural, but it is not used in all tenses, (from Angulo 1933:172).

	1	3SG/DU	3PL	2
PRS	<i>tu-</i>	<i>u-</i>	<i>e-</i>	<i>su-</i>
ANT.PST	<i>tu-</i>	<i>u-</i>	<i>e-</i>	<i>su-</i>
FUT	<i>gu-</i>	<i>ga-</i>	<i>ga-</i>	<i>ki-</i>
REC.PST	<i>ku-</i>	<i>ku-</i>	<i>ku-</i>	<i>ki-</i>
IMM.PST	<i>u-</i>	<i>zu-</i>	<i>zu-</i>	<i>i-</i>
POT	<i>nu-</i>	<i>mu-</i>	<i>mi-</i>	<i>mi-</i>
SEQ	<i>ra-</i>	<i>ru-</i>	<i>ru-</i>	<i>gi-</i>

NEG *su-* *su-* *su-* *si-*

Table 6: Prefixes selected by Class III verbs.

	PRS	ANT.PST	FUT	REC.PST	IMM.PST	POT	SEQ	NEG
1SG	<i>tu-mén</i>		<i>gu-mén</i>	<i>ku-mén</i>	<i>u-mén</i>	<i>nu-mén</i>	<i>ra-mén</i>	<i>su-mén</i>
2SG	<i>su-mén</i>		<i>ki-mén</i>	<i>ki-mén</i>	<i>i-mén</i>	<i>mi-mén</i>	<i>gi-mén</i>	<i>si-mén</i>
3SG	<i>u-mén</i>		<i>ga-mén</i>	<i>ku-mén</i>	<i>zu-mén</i>	<i>mu-mén</i>	<i>ru-mén</i>	<i>su-mén</i>
3PL	<i>e-mén</i>		<i>ga-mbén</i>	<i>ku-mbén</i>	<i>zu-mbén</i>	<i>mi-mén</i>	<i>ru-mbén</i>	<i>su-mbén</i>

Table 7: Basic paradigm of the verb *-mén* ‘like, love’.

Lastra's (1984) account of the forms involved in both Classes II and III differs from Angulo's in two respects: (i) Class II verbs also have prefix *e-* for the 3rd plural present. This could be interpreted as a reflex of an analogical extension having affected Class II verbs to make them more like Class III; and (ii) for the consecutive of Class II verbs, Lastra reports a split involving the 3rd person by virtue of which the prefix *ra-* associated to 1st person singular is also used for the 3rd person plural. In contrast, verbs of Class III use the same prefixes for the consecutive as Class I verbs.

18.1.4. Class IV verbs

Another small group of verbs in Angulo's sample select the prefix set in Table 8. We treat such verbs as Class IV. These verbs also have a series of idiosyncrasies. They also have a complete neutralization between the present and the anterior past. Particular to this class is that all /CV/ prefixes for the 2nd person have /a/. The prefixes in this class do not make distinctions of number. Lastra's (1984) paradigm for this class differs in certain respects from Angulo's and it is worth mentioning it separately. We have highlighted the areas where both accounts contrast.

	Angulo (1933)			Lastra (1984)		
	1	3	2	1	3	2
PRS	<i>tu-</i>	<i>u-</i>	<i>ka-</i>	<i>tu-</i>	<i>u-</i>	<i>ka-</i>
ANT.PST	<i>tu-</i>	<i>u-</i>	<i>ka-</i>	<i>tu-</i>	<i>ku-</i>	<i>ka-</i>
FUT	<i>ga-</i>	<i>ga-</i>	<i>ka-</i>	<i>ga-</i>	<i>ga-</i>	<i>ka-</i>
REC.PST	<i>ku-</i>	<i>ku-</i>	<i>e-</i>	<i>ku-</i>	<i>ku-</i>	<i>ka-</i>
IMM.PST	<i>u-</i>	<i>zu-</i>	<i>i-</i>	<i>u-</i>	<i>e-</i>	<i>e-</i>
POT	<i>mu-</i>	<i>ma-</i>	<i>ma-</i>	<i>mu-</i>	<i>ma-</i>	<i>ma-</i>
SEQ	<i>ra-</i>	<i>ru-</i>	<i>ga-</i>	<i>ra-</i>	<i>ru-</i>	<i>ga-</i>
NEG	<i>su-</i>	<i>sa-</i>	<i>sa-</i>	<i>su-</i>	<i>sa-</i>	<i>sa-</i>

Table 8: Prefixes selected by Class IV verbs.

A sample paradigm of Class IV according to Angulo's account is given in Table 9 with the verb *-ngw^é* ‘strike’ (p.172).¹¹⁷ Like the verb *-tsá* ‘do’ in Table 5 above, this verb has also three different stems, but they are distributed in different zones of the paradigm: *-mb^é* is used for the 3rd person plural, *-mb^é* for the 3rd person singular y

¹¹⁷ According to Angulo (p.154), superscript vowels like the <é> in *ngw^é* ‘strike’ represent so-called ‘whispered vowels’ of which it is said that “[T]here is apparently a very forcible contraction of the whole larynx and pharynx. A vowel thus produced sounds somewhat like a harsh Arabic *h* plus the timbre of a vowel.”

-ngw^e for anything else. As we will show in §3.8, however, not all Class IV verbs behave like this.

	PRS	ANT.PST	FUT	REC.PST	IMM.PST	POT	SEQ	NEG
1SG	<i>tu-ngw^e</i>	<i>ga-ngw^e</i>	<i>ku-ngw^e</i>	<i>u-ngw^e</i>	<i>mu-ngw^e</i>	<i>ra-ngw^e</i>	<i>sa-ngw^e</i>	<i>su-ngw^e</i>
2SG	<i>ka-ngw^e</i>	<i>ka-ngw^e</i>	<i>e-ngw^e</i>	<i>i-ngw^e</i>	<i>ma-ngw^e</i>	<i>ga-ngw^e</i>	<i>sa-ngw^e</i>	<i>su-ngw^e</i>
3SG	<i>u-m^e</i>	<i>ga-m^e</i>	<i>ku-m^e</i>	<i>zu-m^e</i>	<i>ma-m^e</i>	<i>ru-m^e</i>	<i>su-m^e</i>	<i>su-m^e</i>
3PL	<i>u-mb^e</i>	<i>ga-mb^e</i>	<i>ku-mb^e</i>	<i>zu-mb^e</i>	<i>ma-mb^e</i>	<i>ru-mb^e</i>	<i>su-mb^e</i>	<i>su-mb^e</i>

Table 9: Basic paradigm of the verb -ngw^e ‘strike’.

18.1.5. Class V verbs

A few other verbs form the inflectional class we treat as Class V, which is also the most complex. Verbs of this class have a neutralization between recent past and immediate past. As Class I, the characteristic marker for the 2nd person is *ki-*. Likewise, there is syncretism between the 3rd person plural and the 2nd person in the potential. For the 3rd person, the verbs select the same markers as Class I in the present, future, and the past tenses. But what makes this class different from the previous ones is the number split in the 1st person. The prefixes for the 1st person plural and the 1st person singular are contrastive in all tenses, but for the expression of the dual one uses either the singular or the plural form, except in the future and the negative.

	1PL	1DU	1SG	3SG/DU	3PL	2
	<i>u-</i>	<i>e-</i>	<i>e-</i>	<i>e-</i>	<i>e-</i>	<i>ki-</i>
ANT.PST	<i>ti-</i>	<i>ti-</i>	<i>ta-</i>	<i>u-</i>	<i>u-</i>	<i>ki-</i>
FUT	<i>gu-</i>	<i>ti-</i>	<i>ta-</i>	<i>ga-</i>	<i>ga-</i>	<i>ki-</i>
REC.PST	<i>ki-</i>	<i>ki-</i>	<i>ka-</i>	<i>ku-</i>	<i>ku-</i>	<i>ki-</i>
IMM.PST	<i>ki-</i>	<i>ki-</i>	<i>ka-</i>	<i>ku-</i>	<i>ku-</i>	<i>ki-</i>
POT	<i>mu-</i>	<i>ma-</i>	PRS-	<i>ma-</i>	<i>mi-</i>	<i>mi-</i>
SEQ	<i>nu-</i>	<i>na-</i>	<i>na-</i>	<i>ra-</i>	<i>ra-</i>	<i>mi-</i>
NEG	<i>su-</i>	<i>si-</i>	<i>sa-</i>	<i>si-</i>	<i>si-</i>	<i>si-</i>

Table 10: Prefixes selected by Class V verbs.

Apart from a few minor differences, Lastra's report of this class varies very little from Angulo's account. The main differences are: (i) a prefix *ti-* instead of *ta-* is used for the 1st person singular future, suggesting that the split between singular and dual is disappearing for this tense; (ii) a prefix *ma-* instead of *nu-* for 1st person plural consecutive, *mi-* instead of *mu-* for the potential, and *gu-* as an alternative option of *ti-* for the anterior past; and finally (iii) alongside *ra-* the prefix *ma-* is also used for the 3rd person singular consecutive.

The verbs of this class are special in that they require a plural stem for all persons plus a dedicated suffix for 2nd person plural in certain tenses. An example paradigm is given in Table 11 with the verb -ndü ‘go’.

	PRS	ANT.PST	REC/IMM.PST	FUT	POT	SEQ	NEG
1SG	<i>é-ndü</i>	<i>tá-ndü</i>	<i>ká-ndü</i>	<i>tá-ndü</i>	<i>má-ndü</i>	<i>ná-ndü</i>	<i>sá-ndü</i>
2SG	<i>kí-ndü</i>	<i>kí-ndü</i>	<i>kí-ndü</i>	<i>kí-ndü</i>	<i>mí-ndü</i>	<i>mí-ndü</i>	<i>sí-ndü</i>
3SG	<i>é-ndü</i>	<i>ú-ndü</i>	<i>kú-ndü</i>	<i>gá-ndü</i>	<i>má-ndü</i>	<i>rá-ndü</i>	<i>sí-ndü</i>
1DU.INCL	<i>é-ndü-s</i>	<i>tí-ndü-s</i>	<i>kí-ndü-s</i>	<i>tí-ndü-s</i>	<i>má-ndü-s</i>	<i>ná-ndü-s</i>	<i>sá-ndü-s</i>
1PL.INCL	<i>ú-ndür-ín</i>	<i>tí-ndür-ín</i>	<i>kí-ndür-ín</i>	<i>gú-ndür-ín</i>	<i>mú-ndür-ín</i>	<i>nú-ndür-ín</i>	<i>sú-ndür-ín</i>

2PL	<i>kí-ndür-ín</i>	<i>kí-ndür-ín</i>	<i>kí-ndür-kún</i>	<i>kí-ndür-ín</i>	<i>mí-ndür-kún</i>	<i>mí-ndür-kún</i>	<i>sí-ndür-kún</i>
3PL	<i>é-ndür</i>	<i>ú-ndür</i>	<i>kú-ndür</i>	<i>gá-ndür</i>	<i>má-ndür</i>	<i>rá-ndür</i>	<i>sí-ndür</i>

Table 11: Basic paradigm of the motion verb -ndü ‘go’.

18.1.6. Class VI verbs

A group of four verbs in Angulo form ‘Class VI’ (there is no mention of this class in Lastra 1984). This class resembles Class V in that there is also a number split involving the 1st person as well as neutralization between the recent and the immediate past. The verbs further show neutralization between potential and consecutive and between anterior past and future, except for the 1st person plural. The forms for the 1st person singular and those for the 3rd person are completely identical.

	1PL	1DU	1SG	3	2
PRS	<i>u-</i>	<i>e-</i>	<i>e-</i>	<i>e-</i>	<i>ki-</i>
ANT.PST	<i>ti-</i>	<i>ti-</i>	<i>ta-</i>	<i>ta-</i>	<i>sa-</i>
FUT	<i>gi-</i>	<i>ti-</i>	<i>ta-</i>	<i>ta-</i>	<i>sa-</i>
REC.PST	<i>si-</i>	<i>si-</i>	<i>sa-</i>	<i>sa-</i>	<i>sa-</i>
IMM.PST	<i>si-</i>	<i>si-</i>	<i>sa-</i>	<i>sa-</i>	<i>sa-</i>
POT	<i>ni-</i>	<i>ni-</i>	<i>na-</i>	<i>na-</i>	<i>za-</i>
SEQ	<i>ni-</i>	<i>ni-</i>	<i>na-</i>	<i>na-</i>	<i>za-</i>
NEG	<i>su-</i>	<i>si-</i>	<i>si-</i>	<i>si-</i>	<i>si-</i>

Table 12: Prefixes selected by Class VI verbs.

A representative Class VI verb is -ngw^é ‘be, sit’ (homophonous to -ngw^é ‘strike’ in Table 9), whose basic paradigm is given in Table 13. Apart from the two stems in the singular -ngw^é vs. -m^é, as a typical verb of Class V, plural cells receive a different stem, with subsequent allomorphy: *gá*, *ká* and *nga*.¹¹⁸

	PRS	ANT.PST / FUT	REC/IMM.PST	POT/SEQ	NEG
1SG	<i>é-m^é</i>	<i>ta-ngw^é</i>	<i>na-ngw^é</i>	<i>sa-ngw^é</i>	<i>sí-ngw^é</i>
2SG	<i>kí-m^é</i>	<i>sa-ngw^é</i>	<i>za-ngw^é</i>	<i>sa-ngw^é</i>	<i>sí-ngw^é</i>
3SG	<i>é-m^é</i>	<i>ta-ngw^é</i>	<i>na-ngw^é</i>	<i>sa-ngw^é</i>	<i>sí-ngw^é</i>
1DU.INCL	<i>é-m^é-s</i>	<i>ti-ngw^é-s</i>	<i>na-ngw^é</i>	<i>si-ngw^é</i>	<i>sí-ngwe-g-ós</i>
1PL.INCL	<i>u-g^á-n</i>	<i>ti-k^á-n/gi-k^á-n</i>	<i>si-k^á-n</i>	<i>ni-k^á-n</i>	<i>su-ngá-n</i>
3PL	<i>e-gá</i>	<i>ta-ká</i>	<i>na-ká</i>	<i>sa-ká</i>	<i>si-ngan</i>

Table 13: Basic paradigm of the verb -ngw^é ‘be, sit’.

18.1.7. Class VII verbs

The last prefix class of verbs in Chichimec is shown in Table 14. Like verbs of Classes V and VI, the verbs of this class also have a number split but only contrasting singular with dual and plural. Like verbs of Class VI, they show neutralization between anterior past and future, and between potential and consecutive. For the immediate past and the negative, there are no person distinctions, and there is a high degree of syncretism between the 1st person singular and the 3rd person.

¹¹⁸ There are other idiosyncrasies involving the marking of 2nd person plural. For further details see Angulo (1933: 175).

	1PL/DU	1SG	3	2
PRS	<i>ti</i>	<i>ti</i>	<i>i</i>	<i>si</i>
ANT.PST	<i>ti</i>	<i>ta</i>	<i>ta</i>	<i>sa</i>
FUT	<i>ti</i>	<i>ta</i>	<i>ta</i>	<i>sa</i>
REC.PST	<i>si</i>	<i>sa</i>	<i>sa</i>	<i>sa</i>
IMM.PST	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
POT	<i>ni</i>	<i>na</i>	<i>na</i>	<i>za</i>
SEQ	<i>ni</i>	<i>na</i>	<i>na</i>	<i>za</i>
NEG	<i>si</i>	<i>si</i>	<i>si</i>	<i>si</i>

Table 14: Prefixes selected by Class VII verbs.

Lastra's (1984) report differs from Angulo's only in the elision of the number split involving the 1st person. This makes Class VII more similar to the default of Class I. In Table 15 we give the reflexive paradigm of the transitive verb *-nú* 'see' with the meaning of 'see oneself' as a typical example of a Class VII verb (Angulo 1933: 173-174).

	PRS	ANT.PST/FUT	REC.PST	IMM.PST	POT/SEQ	NEG
1SG	<i>ti-nú</i>	<i>ta-nú</i>	<i>sa-nú</i>	<i>i-nú</i>	<i>na-nú</i>	<i>si-nú</i>
2SG	<i>si-nú</i>	<i>sa-nú</i>	<i>sa-nú</i>	<i>i-nú</i>	<i>za-nú</i>	<i>si-nú</i>
3SG	<i>i-nú</i>	<i>ta-nú</i>	<i>sa-nú</i>	<i>i-nú</i>	<i>na-nú</i>	<i>si-nú</i>
3PL	<i>i-nhú</i>	<i>ta-nhú</i>	<i>sa-nhú</i>	<i>i-nhú</i>	<i>na-nhú</i>	<i>si-nhú</i>
1PL.INCL	<i>ti-nú-n</i>	<i>ti-nú-n</i>	<i>si-nú-n</i>	<i>i-nú-n</i>	<i>ni-nú-n</i>	<i>si-nú-n</i>

Table 15: Basic paradigm of the reflexive verb *nú* 'see oneself'.

18.2. Principles of prefix class assignment

The population of the inflectional classes we have introduced in the previous sections in Angulo's sample of 168 verbs is uneven. Although this sample is small, we take it, nonetheless, to be representative of the general trends organizing the Chichimec verbal lexicon as a whole.¹¹⁹

Class I	85	50.5%
Class II	15	9.0%
Class III	6	3.6%
Class IV	11	6.5%
Class V	15	9.0%
Class VI	4	2.4%
Class VII	32	19.0%
Total	168	100%

Table 16: Distribution of the prefix classes.

We take Class I to be an inflectional default. This is because it is the most populated one and as we will see in §3.8, it is also the one that co-occurs with most types of stem

¹¹⁹ There are three verbs which do not fall into these classes but constitute two classes on their own. They are patientive verbs, and we have opted not to include them here for the information at hand is at present very scanty.

alternation patterns. Being an inflectional default means for us that the adscription of a given verb to Class I is not a piece of information that needs to be listed in the lexicon. In other words, as lexemes need to be inflected somehow, in lieu of further specifications, a verbal lexeme is inflected as a Class I verb. Some of the other inflectional classes could also be understood as variations from the formal core of Class I.

In contrast, the assignment to Classes V-VII could be accounted for as partially motivated by the semantics. For example, all verbs of these classes are intransitive. Class VII holds all reflexive counterparts of transitive verbs belonging to other classes. For this reason, we consider reflexive verbs as productively derived by changing the prefix class of a given verb to Class VII.¹²⁰ Because of this, we have not included in our sample the few reflexive verbs appearing in Angulo, and have only considered those that are basic to Class VII. However, despite the predictability of the prefix series a reflexive verb would take, the verb in question may independently involve changes at the stem level which are not predictable from the active counterpart and need to be specified in the lexical entry. However, our data are unfortunately too scanty to be informative about such correlations. This can be seen in Table 17 where the stem alternation pattern of the reflexive verb *-ngwon* ‘pierce oneself’ of prefix Class VII can be compared with that of the transitive counterpart *-pón* ‘pierce’ of Class I in Table 18 (Angulo 1933:173).

	PRS	ANT.PST	REC.PST	IMM.PST	FUT	POT	SEQ	NEG
1SG	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>
2SG	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>
3SG	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>	<i>-ngwon</i>
3PL	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>	<i>-pon</i>

Table 17: Distribution of stems of the reflexive verb *-ngwon* ‘pierce oneself’.

	PRS	ANT.PST	REC.PST	IMM.PST	FUT	POT	SEQ	NEG
1SG	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>	<i>-pón</i>
2SG	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>	<i>-ngwón</i>
3SG	<i>-món</i>	<i>-món</i>	<i>-món</i>	<i>-món</i>	<i>-món</i>	<i>-món</i>	<i>-món</i>	<i>-món</i>
3PL	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>	<i>-mbón</i>

Table 18: Distribution of stems of the transitive verb *-pón* ‘pierce’.

Besides reflexive verbs, Class VII also includes basic verbs. In our sample there are 32 such verbs. Ten of these verbs, given in (2), have stative semantics.¹²¹

- (2)
- | | | | |
|----------------|---------------|-------------|---------------|
| <i>-lè</i> | ‘be angry’ | <i>-ppó</i> | ‘be together’ |
| <i>-kú</i> | ‘be together’ | <i>-sü</i> | ‘be sick’ |
| <i>-ngwé’n</i> | ‘be loaded’ | <i>-te</i> | ‘be equal’ |
| <i>-ngwo</i> | ‘be ashamed’ | <i>-tsü</i> | ‘be angry’ |
| <i>-pqr</i> | ‘be rich’ | <i>-üts</i> | ‘be always’ |

Bearing all reflexive verbs, Class VII could be broadly characterized as including verbs with middle semantics. In this light, the remaining 22 verbs of Class VII fall into

¹²⁰ Reflexive verbs have an alignment split in the immediate past where they are patientive.

¹²¹ The other two verbs with stative semantics in the sample are found in Class I (*m’lan* ‘be finished’) and Class III (*mé* ‘be late’).

the typical middle situation types associated with middle voice in Kemmer (1993). This is illustrated in Table 19, where it is shown that the verbs of Class VII in our sample include reciprocal verbs; verbs with self-centered semantics which could be taken to be deponent reflexives; a few motion verbs (of both the translational and the non-translational type including manner verbs); posture verbs; body function actions; affective and cognitive verbs. However, as the class also includes verbs whose semantic type is hard to characterize and as other classes include other verbs which could easily be equally characterized as having middle semantics, class membership to Class VII is not motivated by just the semantics.

Reciprocal	<i>-sé'n</i>	(intr)	'agree'
	<i>-tér</i>	(intr)	'marry'
	<i>-ngwo</i>	(intr)	'meet'
Self-centered	<i>-hú</i>	(intr)	'put oneself'
	<i>-tü</i>	(intr)	'speak to oneself'
Body function	<i>-pü</i>	(intr)	'defecate'
	<i>-ser</i>	(intr)	'have diarrhoea'
	<i>-tsün</i>	(intr)	'urinate'
Motion	<i>-póó</i>	(intr)	'enter'
	<i>-yúts</i>	(intr)	'jump'
	<i>-há</i>	(intr)	'lag behind'
	<i>-ʔü</i>	(intr)	'pass'
	<i>-mér</i>	(intr)	'roll'
	<i>-kú'</i>	(intr)	'run a race'
Postural	<i>-nda</i>	(intr)	'lie on the ground'
	<i>-hunts</i>	(intr)	'sit'
Cognitive	<i>-tér</i>	(intr)	'laugh'
	<i>-mír</i>	(intr)	'wish'
Other	<i>-múr</i>	(intr)	'get lost'
	<i>-áʔi</i>	(intr)	'appear'
	<i>-sa</i>	(intr)	'measure'

Table 19: Middle verbs of Class VII.

Like Class VII, Classes V and VI also have a semantic core. As shown in Table 20, Class V mainly includes motion verbs. Class VI has four members only: two are postural and two motion verbs. However, despite this semantic core, the membership to classes is a lexical matter because neither all motion nor all postural verbs are found in such classes, making class membership again not entirely predictable by the semantics.

Class V	Motion	<i>-ndü</i>	(intr)	'go'
		<i>-te</i>	(intr)	'go hither'
		<i>-'un</i>	(intr)	'go in the lead'
		<i>-kun</i>	(intr)	'go with'
		<i>-ʔi</i>	(intr)	'come'
		<i>-pó</i>	(intr)	'return'
		<i>-kkü</i>	(intr)	'return, do something again'
		<i>-hǫ</i>	(intr)	'ascend'
		<i>-tsó</i>	(intr)	'fall down'

	Actions with a motion component (physical or fictive)	<i>-ndë</i>	(intr)	'grow up'
		<i>-rú</i>	(intr)	'pass away, die'
		<i>-ʔün</i>	(intr)	'sleep off'
Class VI	Motion	<i>-tú</i>	(intr)	walk about
		<i>-te</i>	(intr)	go out
	Postural	<i>-ngwé</i>	(intr)	sit, be
		<i>-me</i>	(intr)	stand up

Table 20: Verbs of Class V and VI.

In contrast, the members of the remaining classes (II-IV) lack any discernible semantic property in common and the membership to such classes is again lexically determined. For example, membership is not based on the phonological properties of stems. This is shown in Table 21, where we give the 32 verbs of classes II, III and IV arranged attending to similar onsets. In the Table, we have also given examples of verbs of Classes I and VI for comparison. The data reveal that we cannot pin point to the existence of a phonological conditioning for membership except for verbs with onsets in /g/, /r^h/, /k^h/ and /mb/, which only occur in either Classes II or III. However, we take this to be an accident of the sample rather than as a phonological property of the classes.

Onset	CLASS I		CLASS II		CLASS III		CLASS IV	
/ʔ/	-ʔi	(tr) 'want/wish'	-ʔi	(tr) 'hang'	-ʔor	(tr) 'appear'	-ʔi	(tr) 'care'
/m/	-máf	(tr) 'show'	-mé	(intr) 'be late'	-mén	(tr) 'love/like'	-má	(tr) 'think'
/p/	-pá	(tr) 'find'	-póf	(intr) 'dwindle'	-pór	(tr) 'remove'	-por	(tr) 'change'
/k/	-ká	(tr) 'take out'	-kó	(intr) 'deny'				
/n/	-nú	(tr) 'see'	-no	(intr) 'arrive'				
			-nov	(intr) 'last'				
/ts/	-tsá	(tr) 'try'	-tsá	(tr) 'do'				
			-tsé	(tr) 'pay'				
			-tsen	(tr) 'pierce'				
			-tsoh	(tr) 'spend'				
/s/	-sa	(tr) 'read'			-sení	(tr) 'remember'	-sá	(tr) 'win'
/t/	-ta	(tr) 'touch'					-tets	(tr) 'bet'
CLASS VI								
/r/	rú	(intr) 'die'	-rɛ	(tr) 'breed'				
/ng/	ngwé	(intr) 'sit, be'					-ngwá	(intr) 'go'
							-ngwiʔi	(tr) 'charge'
							-ngwá ^á s	(tr) 'pay'
							-ngwé	(tr) 'strike with'
							-ngwin	(intr) 'adhere'
							-ngwe	(intr) 'cry'
/g/			-gá	(tr) 'salute'				
/r ^h /			-rho	(tr) 'finish'				
/k ^h /					khar	(tr) 'possess'		
/mb/					mbe	(tr) 'stone'		
					n			

Table 21: Stem onset of stems in Classes II-IV

In this section, we have presented how the Chichimec verbal lexicon is divided up into at least seven inflectional classes attending to the prefix series they select for the realization of TAM categories and the person of the subject. The morphological complexity of Chichimec does not stop with prefix selection as verbs also require stem alternations, at times rather complex, to inflect for the different person and TAM values. We present these patterns in the following section where we also study the interconnection between such stem alternation patterns and the different prefix classes we have studied so far.

19. Stem alternation patterns

In the previous section, we saw that verbs in Chichimec may have stem alternation patterns. We propose to classify such patterns here in six types. The patterns in question may be morphosyntactically motivated, as for example in Table 22 where the verb *ngw^e* ‘strike with something’ requires three stems: one which functions as a lexical default and two other ‘special’ stems found in the realization of specific person values, namely the 3rd person singular and dual, and the 3rd person plural. In other cases, the patterns are morphological in character because the selection of the ‘special’ stems does not respond to any coherent morphosyntactic feature value. As an illustration take the verb *-pín* ‘narrate’ in Table 23. This verb has up to four different stems. Even if we took the stem *-pín* as the base form, the other three remaining stems *-mbín*, *-ngwín* and *-mín* do not seem to have a clear-cut morphosyntactic distribution and one is forced to propose complex morphological conditions, such as, for example, stem *-ngwín* is used for 2nd person for all TAM values except the present. But the same stem is also found in other cells.

	1SG	1DU	1PL	2SG	2 DU	2PL	3SG	3 DU	3PL
PRES	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
ANT.PST	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
FUT	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
REC.PST	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
IMM.PST	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
POT	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
SEQ	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>
NEG	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-ngw^e</i>	<i>-m^e</i>	<i>-m^e</i>	<i>-mb^e</i>

Table 22: Stem alternation pattern of the verb *-ngw^e* ‘strike with something’.

	1SG	1DU	1PL	2SG	2 DU	2PL	3SG	3 DU	3PL
PRES	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>
ANT.PST	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-mín</i>	<i>-mín</i>	<i>-mbín</i>
FUT	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-mín</i>	<i>-mín</i>	<i>-mbín</i>
REC.PST	<i>-mbín</i>	<i>-mbín</i>	<i>-mbín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-mín</i>	<i>-mín</i>	<i>-pín</i>
IMM.PST	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-pín</i>
POT	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-pín</i>
SEQ	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-pín</i>	<i>-pín</i>	<i>-pín</i>
NEG	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-ngwín</i>	<i>-pín</i>

Table 23: Stem alternation pattern of the verb *-pín* ‘narrate’.

In the following sub-sections, we study the properties of such alternation patterns one by one. We claim that some of the patterns are restricted to the phonological properties of the base stems in question, while others are not. Towards the end of the section we link the different patterns to the different prefix classes in search of morphological correlations between the two different sub-systems in order to understand how independent they are from each other. Our overall conclusion is that the stem patterns are mainly associated with the default prefix class.

Finally, the stem alternation patterns in Tables 22 and 23 involve the distributional properties of stems that contrast segmentally, but which are identical in tone. On top of such segmental differences, verbs may also involve tone alterations. An illustrative example is given in Table 24 for the verb *-her* ‘have’, which has three different stems: *-her*, which we consider to be the base form, *-nhér* and *-rhér*. Notice that the base stem *-her* has a lexical low tone (not represented in the orthography), but in the potential the low tone is overwritten by a high tone (indicated by the acute accent). We study such cases in §4.

	1SG	1DU	1PL	2SG	2 DU	2PL	3SG	3 DU	3PL
PRES	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-rhér</i>
ANT.PST	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-rhér</i>
FUT	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-rhér</i>
REC.PST	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-rhér</i>
IMM.PST	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-her</i>	<i>-her</i>	<i>-rhér</i>
POT	<i>-hér</i>	<i>-hér</i>	<i>-hér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-hér</i>	<i>-hér</i>	<i>-rhér</i>
SEQ	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-her</i>	<i>-her</i>	<i>-rhér</i>
NEG	<i>-her</i>	<i>-her</i>	<i>-her</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-nhér</i>	<i>-her</i>	<i>-her</i>	<i>-rhér</i>

Table 24: Stem alternation pattern of the verb *-her* ‘have’.

19.1. The minimal stem pattern: A-stems vs. B-stems

Most verbs in Chichimec have at least two stems. In the default case, one of these stems is used for all 3rd person plural forms. We will call this stem ‘B-stem’. The other stem has a wider distribution and we will refer to it as the ‘A-stem’ treating it as a base form which is available for further morphophonological operations.

In the vast majority of cases, a B-stem is derived from an A-stem by the mutation of the consonantal onset of the A-stem to its aspirated correlate. Table 25 gives an indication of the mechanisms in question organized according to onset. Notice that for some of the verbs, there is further lexical specification, i.e. speakers need to remember what the correct outcome is given that there are different options at play. For such cases, we first specify the number of verbs found in the sample with that specific onset. The different options are represented by the small roman numbers together with an indication of their frequency. Cases in the sample without lexically specified options are treated as ‘All’.¹²²

¹²² Verbs like *sus* ‘split’ in (e/iii) and *máf* ‘show’ in (i/iii) have B-stems such as *tsus* and *páf*, respectively. These cases suggest some constraint on root syllables to have one and only one [spread glottis] feature. As the respective codas in the verb are already specified with [+spread glottis] this

A-STEM Onset	B-STEM Onset	Prefix Class	Total	Outcomes	Total			
/p/	-pá	/m ^b /	-mbá	‘see, find’	(I)	22	(i)	17
	-pan	/p ^h /	-phan	‘heat’	(I)		(ii)	5
/t/	-tá	/r ^h /	-rhá	‘buy’	(I)	21	(i)	16
	-tq	/t ^h /	-thq	‘search for’	(I)		(ii)	3
	-tav	/r’/	-rʔav	‘cut’	(I)		(iii)	2
/ts/	-tsá	/ts ^h /	-tshá	‘hurt’	(I)	All		
/ʃ/	-tcon	/z ^h /	-zhon	‘bite’	(I)	All		
/s/	-sú	/ts ^h /	-tshú	‘take out’	(I)	8	(i)	5
	-sus		-tsus	‘split’	(I)		(ii)	2
	-sú	/s ^h /	-shú	‘peel’	(I)		(iii)	1
/k/	-kq	/k ^h /	-khq	‘take out’	(I)	All		
/ʔ/	-ʔe	/r’/	-r-ʔe	‘put’	(I)	All		
/h/	-há	/r ^h /	-rhá	‘drink’	(I)	All		
/r ^h /	-rho	/nt ^h /	-ntho	‘finish’	(II)	All		
/m/	-má	/m ^h /	-mhá	‘speak’	(I)	8	(i)	7
	-máf	/p/	-páf	‘show’	(I)		(ii)	1
/n/	-nef	/n ^h /	-nhef	‘run’	(I)	All		
/n ^d /	-ndun	/n ^{dh} /	-ndhun	‘buzz’	(I)	All		
/ng ^w /	-ngwé	/m ^b /	-mbé	‘strike with’	(IV)	6	(i)	4
	-ngwa ^a s	/p ^h /	-pha ^a s	‘pay’	(IV)		(ii)	2
/g/	-gá	/ng/	-ngar	‘salute’	(II)	All		

Table 25: The building of B-stems.

Verbs of prefix Classes VI and VII do not commonly build their B-stems following the mechanisms in Table 26. For such verbs, a B-stem is formed by suffixing *-r* to the A-stem. The same happens to the Class II verb *rë* ‘breed’ and to some Class V verbs.

A-STEM	B-STEM	Expected			
-me	-me-r	*-mhe	‘stand up’	(VI)	3 (out of 4)
-tsó	-tsó-r	*-tshó	‘fall’	(VII)	11 (out of 15)
-hú	-hú-r	*-r-hú	‘put oneself’	(V)	3 (out of 32)
-rę	-rę-r	*-rhe	‘breed’	(II)	1

Table 26: Exceptions for the building of B-stems.

At least 5 verbs of Class V, one of Class VI and another of Class IV have B-stems whose forms could be treated as suppletive when compared to their respective A-stems. For the verbs in (a) in Table 27, the B-stem also bears the suffix *-r*. The form of the B-stems for the verbs in (c) in reality is that of another stem we call a ‘C-stem’,

prevents aspiration from the initial consonant: *shus or *mháf. The outcome /ts/ and /p/ in the B-stem is, however, not predictable.

which we will treat further below in §3.2. Despite their form, the stems in (c) have the same distribution as B-stems.

	A-STEM	B-STEM	Expected		
a.	<i>-pó^o</i>	<i>-nhur</i>	* <i>-mbó^o</i>	‘enter’	(V)
	<i>-ʔü</i>	<i>-nür</i>	* <i>-rʔü</i>	‘pass’	(V)
b.	<i>-hunts</i>	<i>-kants</i>	* <i>-rhunts</i>	‘sit’	(V)
	<i>-ngwé</i>	<i>-ká</i>	* <i>-mbé</i>	‘sit, be’	(VI)
c.	<i>-ngwéʔn</i>	<i>-péʔn</i>	* <i>-mbéʔn</i>	‘be loaded with’	(V)
	<i>-ngwo</i>	<i>-vo</i>	* <i>-mbo</i>	‘meet’	(V)
	<i>-á-ʔi</i>	<i>-á-ti</i>	* <i>-á-rʔi</i>	‘appear’	(V)
	<i>-ʔi</i>	<i>-ti</i>	* <i>-rʔi</i>	‘care’	(IV)

Table 27: Suppletive B-stems.

19.2. The distribution of B-stems

For most verbs, the B-stem is used for all cells in the paradigm involving a 3rd person plural. Many verbs have only two stems, and they are distributed according to the stem alternation pattern in Table 28, which we treat as pattern 1a. A typical example of such pattern is the verb *-nú* ‘see’.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
ANT.PST	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
FUT	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
REC.PST	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
IMM.PST	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
POT	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
SEQ	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>
NEG	A	A	A	B	<i>-nú</i>	<i>-nú</i>	<i>-nú</i>	<i>-nhú</i>

Table 28: Pattern 1a: verb *-nú* ‘see’.

Pattern 1.a is used with most verbs regardless of prefix class, but verbs of prefix Class VI and VII are special in that they require a two-way split involving all plural cells regardless of person. For such verbs, the B-stem is used in all cells involving a plural subject. This is pattern 1b, and it is illustrated in Table 29 with the verb *-há* ‘ascend’ of Class VII.

	SG/DU	PL	SG/DU	PL
PRES	A	B	<i>-há</i>	<i>-há-r</i>
ANT.PST	A	B	<i>-há</i>	<i>-há-r</i>
FUT	A	B	<i>-há</i>	<i>-há-r</i>
REC.PST	A	B	<i>-há</i>	<i>-há-r</i>
IMM.PST	A	B	<i>-há</i>	<i>-há-r</i>
POT	A	B	<i>-há</i>	<i>-há-r</i>
SEQ	A	B	<i>-há</i>	<i>-há-r</i>

NEG	A	B	-hą	-hą-r
-----	---	---	-----	-------

Table 29: Pattern 1b: verb -hą ‘ascend’.

19.2.1. Verbs without a B-stem

Some verbs do not have a B-stem. For the verbs in (a) in Table 30, one could say that the morphophonology fails to provide a phonologically distinct B-stem. For example, all verbs of Class VII build their B-stem by adding the suffix *-r*, but as the sequence /CVnr/ is banned by the phonotactics, the B-stem surfaces as identical to the A-stem.¹²³ This ban affects four verbs of Class VII. The ones in (b), however, should be accounted for as being irregular because in principle they could have a B-stem by means of the default mechanisms proper of their class, but instead they do not.

	Stem A	Stem B		
a.	<i>-khar</i>	—	‘possess’	(III)
	<i>-mben</i>	—	‘stone’	(III)
	<i>-sení</i>	—	‘remember’	(III)
	<i>-tun</i>	—	* <i>-tun-r</i>	‘go in the lead’ (VII)
b.	<i>-hu</i>	—	* <i>-r-hú</i>	‘put, place’ (I)
	<i>-ʔi</i>	—	* <i>-r-ʔi</i>	‘hang’ (II)
	<i>-ʔe</i>	—	* <i>-r-ʔe</i>	‘do thus’ (II)
	<i>-ʔor</i>	—	* <i>-r-ʔor</i>	‘appear’ (III)
	<i>-tshé</i>	—	* <i>-tshé</i>	‘pay’ (II)
	<i>-tú</i>	—	* <i>-thú/ *rhú</i>	‘finish’ (II)
	<i>pór</i>	—	* <i>phór/ *mbór</i>	‘remove’ (III)

Table 30: Verbs without a B-stem.

In contrast to the cases in Table 30, 22 out of the 32 deponent middle verbs of prefix Class V do not have a B-stem. This specific lack of a B-stem we treat as a morphological characteristic of Class V. In this light, for these 22 verbs of Class V one could propose that they instantiate a pattern with no paradigmatic splits, illustrated in Table 31 with the verb *tér* ‘laugh’. We refer to this pattern as 1c. Alternatively, for such verbs one could still treat such cases as irregular instances of patterns 1a or 1b – depending on their prefix class – by arguing that their B-stem is phonologically identical to their A-stem. On the other hand, reflexive verbs of Class V derived from transitive verbs inherit some of the stem alternation properties of their source verbs in keeping a B-stem contrast, like for example in Tables 17 and 18 above.

	1	2	3	1	2	3
PRES	A	A	A	-tér	-tér	-tér
ANT.PST	A	A	A	-tér	-tér	-tér
FUT	A	A	A	-tér	-tér	-tér
REC.PST	A	A	A	-tér	-tér	-tér
IMM.PST	A	A	A	-tér	-tér	-tér

¹²³ We tested this observation in contemporary Chichimec. A hypothetical form **ekharer* ‘they possess’ is ungrammatical. The only form accepted is the one similar to the A-stem.

POT	A	A	A	-tér	-tér	-tér
SEQ	A	A	A	-tér	-tér	-tér
NEG	A	A	A	-tér	-tér	-tér

Table 31: Pattern 1c (without stem alternations): verb -tér ‘laugh’

Finally, the verb *-mén* ‘love’ of prefix Class III displays a hybrid stem alternation pattern between 1a and 1c, as illustrated in Table 32, which could be called pattern 1d. Notice that this verb is doubly irregular because its B-stem is also suppletive, i.e. the regular outcome should have been **-mhén*.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	A	-mén	-mén	-mén	-mén
ANT.PST	A	A	A	A	-mén	-mén	-mén	-mén
FUT	A	A	A	B	-mén	-mén	-mén	-mbén
REC.PST	A	A	A	B	-mén	-mén	-mén	-mbén
IMM.PST	A	A	A	B	-mén	-mén	-mén	-mbén
POT	A	A	A	A	-mén	-mén	-mén	-mén
SEQ	A	A	A	B	-mén	-mén	-mén	-mbén
NEG	A	A	A	B	-mén	-mén	-mén	-mbén

Table 32: Pattern 1d: verb mén ‘love’.

The distribution of verbs displaying the different sub-types of pattern 1 is given in Table 33.

Pattern 1a	36
Pattern 1b	19
Pattern 1c	30
Pattern 1d	1
Total	86

Table 33: Distributions of sub-types of stem pattern 1.

There are 30 verbs of pattern 1 that do not have a B-stem. The rest (56) abide by a basic stem alternation pattern involving just two stems. Other verbs in Chichimec have more complex stem alternation patterns, which are hierarchical in structure as all of them also require a B-stem in their paradigm structure. We study them in the following sections.

19.3. Stem alternation pattern 2

Besides the minimal stem alternation pattern involving two stems, 19 verbs in our sample display a second stem alternation pattern we call ‘Pattern 2’, which involves a third stem we treat as the ‘C-stem’. Pattern 2 requires a C-stem for the 3rd person singular and dual. The pattern was already illustrated in Table 22 above by the verb *-ngwé* ‘strike with something’, but it is given again for convenience in Table 34 with the verb *-pan* ‘heat’. Here, like in patterns 1.a and 1.b, the distribution of the stems attends to person values only; that is, it is not sensitive to TAM. Notice that pattern 2 is hierarchically built on pattern 1.a.

	1	2	3SG/DU	3PL		1	2	3SG/DU	3PL
PRES	A	A	C	B		-pan	-pan	-mban	-phan
ANT.PST	A	A	C	B		-pan	-pan	-mban	-phan
FUT	A	A	C	B		-pan	-pan	-mban	-phan
REC.PST	A	A	C	B		-pan	-pan	-mban	-phan
IMM.PST	A	A	C	B		-pan	-pan	-mban	-phan
POT	A	A	C	B		-pan	-pan	-mban	-phan
SEQ	A	A	C	B		-pan	-pan	-mban	-phan
NEG	A	A	C	B		-pan	-pan	-mban	-phan

Table 34: Pattern 2: verb -pan ‘heat’

We propose that the building of a C-stem involves an abstract process of ‘sonorization’ of the consonantal onset of an A-stem. The relevant outcomes are given in Table 35 with examples. Notice that for the onsets in /p/ and /ng^w/ there is further lexical specification as to the type of outcome.

A-STEM	C-STEM	Prefix Class	Total	Outcomes	Total
/p/	-pii	/m/ -mii ‘spy’	(I)	(i)	14
	-pen	/b/ [β] -ven ‘weigh’	(I)	(ii)	3
	-pan	/ ^m b/ > [mb] -mban ‘heat’	(I)	(iii)	1
/t/	-tav	/ɾ/ -rav ‘cut’	(I)	All	
/ts/	-tsó	/z/ -zó ‘scold’	(I)	All	
/ʃ/	-tcon	/z/ -zon ‘bite’	(I)	All	
/s/	-sa	/dz/ -dza ‘read’	(I)	All	
/k/	-kets	/g/ -gets ‘pick up’	(I)	All	
/ʔ/	-ʔán	/n-ʔ/ > [nd] -n-dán ‘ask’	(I)	All	
/h/	-há	/n-h/ -n-há ‘drink’	(I)	All	
/ɾ ^h /	-rho	/t ^h / -tho ‘finish’	(II)	All	
/ɾ/	-rɛ	/n-ɾ/ > [nd] -ndɛ ‘breed’	(II)	All	
/ng ^w /	-ngwé	/m/ -mé ‘strike with’	(IV)	(i)	4
	-ngwá ^a s	/p/ -pa ^a s ‘pay’	(IV)	(ii)	2

Table 35: The building of C-stems.

As illustrated in Table 35, for the simplest of cases a voiceless consonant becomes voiced, e.g. A -kets > C -gets ‘ask’, but in other cases, the voiced outcome is not the expected counterpart of the base form: e.g. A -tsó vs. C -zó ‘scold’. We interpret this as a phonological constraint because there are no contrastive labial and coronal voiced stops in the inventory of Chichimec: these are produced with prenasalization, fricativization or flaps. Thus, the alternations just described could still be taken to form a natural class at some abstract level. We could offer a preliminary hypothesis to include the default case whereby an onset *ngw* > *m*. Increasing the sonority of a labialized velar may be achieved by enlarging its inherent rather short vocal tract, as in a plain labial, in which case the overall aerodynamic conditions for sonority (voicing) will be enhanced. Verbs with glottic and rhotic consonants receive a nasal

which could be taken as the segmental reflex of a historical nasal prefix that may have once served as the marker to build a C-stem, now disappeared, which was responsible for the sonorization we observe today. Besides the verbs in Table 35, there are two verbs in our sample with an unexpected C-stem built with nasalization of a long nucleus, i.e. A *-paa* > C *-paan* ‘look hither’ or A *-pii* > C *-piin* ‘carry’.

19.4. Stem alternation pattern 3

Three verbs in Angulo's sample display yet another stem pattern that we treat as ‘pattern 3’. This new pattern involves another stem we call the ‘D-stem’. In pattern 3, the D-stem is used for the 2nd person, as illustrated in Table 36.¹²⁴

	1	2	3SG/DU	3PL		1	2	3SG/DU	3PL
PRES	A	D	A	B		-tsá	-tcá	-tsá	-tshá
ANT.PST	A	D	A	B		-tsá	-tcá	-tsá	-tshá
FUT	A	D	A	B		-tsá	-tcá	-tsá	-tshá
REC.PST	A	D	A	B		-tsá	-tcá	-tsá	-tshá
IMM.PST	A	D	A	B		-tsá	-tcá	-tsá	-tshá
POT	A	D	A	B		-tsá	-tcá	-tsá	-tshá
SEQ	A	D	A	B		-tsá	-tcá	-tsá	-tshá
NEG	A	D	A	B		-tsá	-tcá	-tsá	-tshá

Table 36: Pattern 3: verb *-tsá* ‘do’.

A D-stem is found in other patterns besides pattern 3, but all in all the number of verbs that can have a D-stem is restricted to the ones that have the onsets in Table 36. However, not all verbs having such onsets also have a D-stem or pertain to the patterns requiring a D-stem, so within this limited set of possibilities, speakers still need to remember which lexemes undergo which pattern. In contrast to the building mechanisms for B-stems and C-stems, the mutated onset of D-stems is more difficult to be accounted for in terms of a unified morphophonological process. For the verbs with onset /*ts*/ is palatalization, but for those with a glottal stop onset, the D-stem takes the form of a dental onset, suggesting coronalization turning the glottal gesture into a dentoalveolar one. But verbs with a glottal fricative /*h*/ lack a dedicated D-stem, and instead use a prenasalization stem that surfaces identical to their C-stem.¹²⁵

A-STEM	D-STEM					
/p/	<i>-pa</i>	/ng ^w /	<i>-ngwa</i>	‘cover’	(I)	All
/ts/	<i>-tsó</i>	/ʃ/	<i>-tcó</i>	‘scold’	(I)	All
/ʔ/	<i>-ʔán</i>	/t/	<i>-tán</i>	‘ask’	(I)	All
/h/	<i>-hü</i>	/n-h/	<i>-n-hü</i>	‘burn’	(I)	All

Table 37: The building of D-stems.

¹²⁴ Besides *-tsá* ‘do’, the two other verbs are *-tsoh* ‘scold’ (B *-thoh*/C *-tcoh*) and *-tsen* ‘pierce’ (B *-tshen*/C *-tceen*).

¹²⁵ The only noticeable exception is the verb *-ʔi* ‘suck’ (D *-n ʔi*).

All stem alternation patterns we have observed so far (except subtype 1.d) involve distributional properties of special stems as conditioned by the person or number of the subject. This means that the distribution of the stem allomorphy can be accounted for as responding to a simple conditioning factor easily spelled out by the morphosyntax, such as for example ‘a B-stem is required in 3rd person plural’ or ‘a D-stem is required for the 2nd person’, etc. There remain, however, three other types of stem alternations which are morphologically more complex in structural terms because the distribution of stems does not respond to a unified morphosyntactic value, hence it is not motivated, and involves a list of specifications, sometimes long, such as for example ‘a C-stem is required for the 3rd person singular or dual in all past tenses and the future’, etc, which point to the intrinsic morphological character of such patterns.¹²⁶ We treat such more complex patterns in the following sections.

19.5. Stem alternation pattern 4

22 two verbs in Angulo's sample abide by the stem alternation pattern we treat as ‘pattern 4’. This pattern requires four stems: An A-stem, a B-stem, a C-stem and a fifth type of stem we call the ‘E-stem’, which is used exclusively in this pattern for the recent past of the 1st person. Like pattern 2, this new pattern also requires a C-stem for the 3rd person, but restricted to specific TAM values, as shown in Table 38. There are two subtypes. In one, the C-stem is found in the future and in all past tenses; in the other, the immediate past is excluded.¹²⁷

		1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
Subtype 4.a (19 verbs)	PRES	A	A	A	B	-pen	-pen	-pen	-phen
	ANT.PST	A	A	C	B	-pen	-pen	-ven	-phen
	FUT	A	A	C	B	-pen	-pen	-ven	-phen
	REC.PST	E	A	C	B	m ^b en	-pen	-ven	-phen
	IMM.PST	A	A	C	B	-pen	-pen	-ven	-phen
	POT	A	A	A	B	-pen	-pen	-pen	-phen
	SEQ	A	A	A	B	-pen	-pen	-pen	-phen
	NEG	A	A	A	B	-pen	-pen	-pen	-phen
Subtype 4.b (3 verbs)	PRES	A	A	A	B	-ker	-kér	-ker	-khér
	ANT.PST	A	A	C	B	-ker	-kér	-gér	-khér
	FUT	A	A	C	B	-ker	-kér	-gér	-khér
	REC.PST	E	A	C	B	-nger	-kér	-gér	-khér
	IMM.PST	A	A	A	B	-ker	-kér	-ker	-khér
	POT	A	A	A	B	-kér	-kér	-kér	-khér
	SEQ	A	A	A	B	-ker	-kér	-ker	-khér
	NEG	A	A	A	B	-ker	-kér	-ker	-khér

Table 38: Pattern 4: verbs -pen ‘weigh’ and -ker ‘learn’.

E-stems only appear in pattern 4 and are phonologically derived from C-stems by means of a nasal prefix.

¹²⁶ Morphemes in the sense by Aronoff (1993).

¹²⁷ The A-stem of the verb -ker ‘learn’ receives high tone for the 2nd person. This tonal pattern and others will be studied in more detail in §4.

A-STEM	C-STEM	E-STEM					
/p/	-pen	/b/ [β]	-ven	/n-b/	-mben	‘weigh’	(I) All
/ʃ/	-tcon	/z/	-zon	/n-z/	-nzon	‘bite’	(I) All
/k/	-kɨn	/g/	-gɨn	/n-g/	-ngɨn	‘look hard’	(I) All
/t/	-tó	/ɾ/	-ró	/n-ɾ/ > [nd]	-ndó	‘watch’	(I) All

Table 39: The building of E-stems.

19.6. Stem alternation pattern 5

Other 22 verbs in our sample display yet another stem alternation pattern we treat as ‘pattern 5’. This pattern results as a combination of patterns 3 and 4.b: like pattern 3, a D-stem is used for the 2nd person; like pattern 4.b, a C-stem is used for the 3rd person singular and dual of the anterior past, the recent past and the future. The pattern has various subtypes involving the marking of present (subtypes *a* and *b*) and the marking of the 3rd person singular and dual of the consecutive (subtypes *i* and *ii*). Verbs of pattern 5 are characterized phonologically by having an A-stem with a glottal consonant, but there are also three with an alveolar affricative /ts/.¹²⁸

		1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
Subtype 5.a.i (10 verbs)	PRES	A	A	A	B	-ʔi	-ʔi	-ʔi	-rʔi
	ANT.PST	A	D	C	B	-ʔi	-tí	-ndí	-rʔi
	FUT	A	D	C	B	-ʔi	-tí	-ndí	-rʔi
	REC.PST	A	D	C	B	-ʔi	-tí	-ndí	-rʔi
	IMM.PST	A	D	A	B	-ʔi	-tí	-ʔi	-rʔi
	POT	A	D	A	B	-ʔi	-tí	-ʔi	-rʔi
	SEQ	A	D	A	B	-ʔi	-tí	-ʔi	-rʔi
	NEG	A	D	A	B	-ʔi	-tí	-ʔi	-rʔi
Subtype 5.a.ii (4 verbs)	PRES	A	A	A	B	-hü	-hü	-hü	-rhü
	ANT.PST	A	D	C	B	-hü	-nhü	-nhü	-rhü
	FUT	A	D	C	B	-hü	-nhü	-nhü	-rhü
	REC.PST	A	D	C	B	-hü	-nhü	-nhü	-rhü
	IMM.PST	A	D	A	B	-hü	-nhü	-hü	-rhü
	POT	A	D	A	B	-hü	-nhü	-hü	-rhü
	SEQ	A	D	C	B	-hü	-nhü	-nhü	-rhü
	NEG	A	D	A	B	-hü	-nhü	-hü	-rhü
Subtype 5.b.i (3 verbs)	PRES	A	D	A	B	-tsó	-tco	-tsó	-tshó
	ANT.PST	A	D	C	B	-tsó	-tco	-zó	-tshó
	FUT	A	D	C	B	-tsó	-tco	-zó	-tshó
	REC.PST	A	D	C	B	-tsó	-tco	-zó	-tshó
	IMM.PST	A	D	A	B	-tsó	-tco	-tsó	-tshó
	POT	A	D	A	B	-tsó	-tco	-tsó	-tshó
	SEQ	A	D	A	B	-tsó	-tco	-tsó	-tshó
	NEG	A	D	A	B	-tsó	-tco	-tsó	-tshó

¹²⁸ As pointed out above and indicated in Table 32, verbs with an onset /h/ have homophonous C and D-stems. This makes a verb such as -hü ‘burn’ appear to have a C-stem for the stem zones where other verbs would have a phonologically distinct D-stem. Positing stem homophony saves the analysis from having to propose another stem alternation pattern, specific for verbs with an onset /h/.

Subtype 5.b.ii (5 verbs)	PRES	A	D	A	B	-ʔer	-tir	-ʔer	-rʔer
	ANT.PST	A	D	C	B	-ʔer	-tir	-ndir	-rʔer
	FUT	A	D	C	B	-ʔer	-tir	-ndir	-rʔer
	REC.PST	A	D	C	B	-ʔer	-tir	-ndir	-rʔer
	IMM.PST	A	D	A	B	-ʔer	-tir	-ʔer	-rʔer
	POT	A	D	A	B	-ʔer	-tir	-ʔer	-rʔer
	SEQ	A	D	C	B	-ʔer	-tir	-ndir	-rʔer
	NEG	A	D	A	B	-ʔer	-tir	-ʔer	-rʔer

Table 40: Pattern 5: verbs -ʔi 'want', -hü 'burn', -tsó 'scold' and -ʔer 'throw'.

19.7. Stem alternation pattern 6

The last and most complex stem alternation pattern is given in Table 41. This pattern shares some structure with the previous patterns. Pattern 6 is like pattern 5.i in having a D-stem for the 2nd person and a C-stem for the 3rd person singular and dual of the anterior past, the recent past and the future. Like pattern 5.a it also requires an A-stem for the present, but this time for all persons. Like pattern 4, the verbs of pattern 6 also require an E-stem for the 1st person of the recent past, but as the verbs in question have a B-stem of class (i) (see Table 25 above), thus both stems surface as homophonous.

Pattern 6 has its own idiosyncrasies: (a) the verbs do only require a B-stem for the 3rd person plural in the anterior past, the future, and for some verbs also in the consecutive; (b) the forms for the 1st person are built involving three different stems; and (c) the immidiative past, the potential and the negative all select a D-stem, except for the 3rd person plural.

		1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
Subtype 6.a (9 verbs)	PRES	A	A	A	A	-pín	-pín	-pín	-pín
	ANT.PST	A	D	C	B	-pín	-ngwín	-mín	-mbín
	FUT	A	D	C	B	-pín	-ngwín	-mín	-mbín
	REC.PST	E	D	C	A	-mbín	-ngwín	-mín	-pín
	IMM.PST	D	D	D	A	-ngwín	-ngwín	-ngwín	-pín
	POT	D	D	D	A	-ngwín	-ngwín	-ngwín	-pín
	SEQ	A	D	A	A	-pín	-ngwín	-pín	-pín
	NEG	D	D	D	A	-ngwín	-ngwín	-ngwín	-pín
Subtype 6.b (7 verbs)	PRES	A	A	A	A	-pi	-pi	-pi	-pi
	ANT.PST	A	D	C	B	-pí	-ngwi	-mí	-mbí
	FUT	A	D	C	B	-pí	-ngwi	-mí	-mbí
	REC.PST	E	D	C	A	-mbí	-ngwi	-mí	-pí
	IMM.PST	D	D	D	A	-ngwí	-ngwi	-ngwí	-pí
	POT	D	D	D	A	-ngwí	-ngwi	-ngwí	-pí
	SEQ	A	D	A	B	-pí	-ngwi	-pí	-mbí
	NEG	D	D	D	A	-ngwí	-ngwi	-ngwí	-pí

Table 41: Pattern 6: verbs -pín 'narrate' and -pi 'wait'.

19.8. Other stem alternation patterns

Three verbs in the sample of pattern 1.b have stem alternation patterns which are idiosyncratic because they require a suppletive stem for either present or the negative or both. These suppletive stems are built on the stem zones proper of pattern 1.b, so that the verbs in question end up having two A-stems and two B-stems. The verbs in question are given in Table 42.

	SG/DU	PL	SG/DU	PL	SG/DU	PL
PRES	-tú	-túr	-né	-nér	-me/-ni	-gá/-gün
ANT.PST	-tú	-túr	-te	-ter	-ngwé	-ká
FUT	-tú	-túr	-te	-ter	-ngwé	-ká
REC.PST	-tú	-túr	-te	-ter	-ngwé	-ká
IMM.PST	-tú	-túr	-te	-ter	-ngwé	-ká
POT	-tú	-túr	-te	-ter	-ngwé	-ká
SEQ	-tú	-túr	-te	-ter	-ngwé	-ká
NEG	-nú	-núr	-né	-nér	-ngwé	-ká

Table 42. Irregular verbs -tú 'walk about', -te 'go out' and -ngwé 'sit, be'.

After having seen the six different stem alternation patterns observable in our corpus of Chichimec verbs with their different subtypes, a question still remains about predictability: To what extent is it predictable that a given verb will be inflected according to a given pattern? We try to give an answer to this question in the following section.

19.9. Membership to the stem alternation patterns

In order to tackle the issue of predictability of verbs to specific patterns, let us start by considering first verbs of pattern 6. All verbs of pattern 6 have an A-stem with an onset /p/. However, a quick look at the verbs in our small sample reveals that verbs with a similar onset /p/ pertain to other patterns. This can be seen if we compare the stem patterns of the verbs in Table 43: -pár 'be rich' of pattern 1.c (with no stem alternations); -pan 'heat' of pattern 2 with B-stem of Class (ii) and C-stem of Class (iii); -pq 'see, find' of pattern 4.a with B-stem of class (i) and C-stem of class (ii); and -pa 'cover' of pattern 6.b with B-stem of class (i) and -stem of class (i).

	1	2	3SG/DU	3PL		1	2	3SG/DU	3PL
PRES	1c -pár	-pár	-pár	-pár	2	-pan	-pan	-mban	-phan
ANT.PST	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
FUT	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
REC.PST	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
IMM.PST	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
POT	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
SEQ	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	
NEG	-pár	-pár	-pár	-pár	-pan	-pan	-mban	-phan	

	4a			6b				
PRES	-pá	-pá	-pá	-mbá	-pa	-pa	-pa	-pa
ANT.PST	-pá	-pá	-vá	-mbá	-pa	-ngwa	-ma	-mba
FUT	-pá	-pá	-vá	-mbá	-pa	-ngwa	-ma	-mba
REC.PST	-mbá	-pá	-vá	-mbá	-mba	-ngwa	-ma	-pa

IMM.PST	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-v\acute{q}}$	$\overline{-mb\acute{q}}$	$\overline{-ngwa}$	$\overline{-ngwa}$	$\overline{-ngwa}$	$\overline{-pa}$
POT	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-mb\acute{q}}$	$\overline{-ngwa}$	$\overline{-ngwa}$	$\overline{-ngwa}$	$\overline{-pa}$
SEQ	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-mb\acute{q}}$	$\overline{-pa}$	$\overline{-ngwa}$	$\overline{-pa}$	$\overline{-mpa}$
NEG	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-p\acute{q}}$	$\overline{-mb\acute{q}}$	$\overline{-ngwa}$	$\overline{-ngwa}$	\overline{ngwa}	$\overline{-pa}$

Table 43: Comparison of the stem alternation patterns of verbs with an onset /p/.

The existence of verbs like the ones in Table 43 indicates that the stem alternation behaviour of a certain verb cannot be predicted from the phonological shape of its base stem and it needs to be listed in the lexicon. However, from our small sample one can still observe strong tendencies of membership to a given pattern based on the phonological shape of the onsets of the base form. This is shown in Table 44, where we correlate the frequency of type of onset with type of pattern.

	1a	1b	1c	1d	2	3	4a	4b	5ai	5aii	5bi	5bii	6a	6b	Total
/p/	4	1	3		1		3						9	7	28
/t/	3	4	5		1										13
/ts/	2	1			1	3	13	2			3				25
/ʃ/							1								1
/s/	6		6		2										14
/g/					2										2
/k/	1	2	2		4		2	1							12
/k ^h /			1												1
/ʀ/	3	4	4						5	1		6			23
/h/	2	1	1						5	3					12
/m/	8	1	4	1											14
/n/	4														4
/nd/	1	2	1												4
/ngw/	2	1	1		6										10
/r/		1			1										2
/r ^h /					1										1
Other ¹²⁹			2												2
Total	36	18	30	1	19	3	19	3	10	4	3	6	9	7	168

Table 44: Type of onset per stem alternation pattern.

If the figures in Table 44 are observed vertically, one can see that some patterns accept more types of onsets than others, for example the membership to patterns 1 and 2 is clearly not phonologically conditioned, whether that to patterns 5 and 6 clearly is. In contrast, if the figures are observed horizontally, one may conclude that predictability for pattern membership is clearer for some type of onsets than for others. If the base form of a verb has onsets such as /p/, /ts/, /ʀ/ or /h/, it is difficult to predict membership. But the number of options is significantly reduced with onsets /t/ and /k/ or even /ngw/. Verbs with a /m/ pertain to pattern 1, mainly due to phonological restrictions in producing other suitable stems to feed the stem zones characteristic of other patterns. On the other hand, we take onset types instantiated in the sample by just one verb as being uninformative.

Another way to tackle the distribution of the patterns across verbs is to study the relation they hold with the prefix classes. Again, our sample is too small to be informative about the Chichimec verbal lexicon in a definite manner, but it nonetheless shows illustrative tendencies, shown in Table 45.

¹²⁹ The verbs *-üts* ‘be always’ and *-yüts* ‘jump’.

	1a	1b	1c	1d	2	3	4a	4b	5a.i	5a.ii	5b.i	5b.ii	6a	6b	Total
I	17				8		19	3	10	4	3	5	9	7	85
II	5		3		4	3									15
III			5	1											6
IV	4				7										11
V		15													15
VI		4													4
VII	10		22												32
Total	36	18	30	1	19	3	19	3	10	4	3	6	9	7	168

Table 45. Prefix class per stem alternation pattern.

If the figures are observed horizontally, one can see that verbs of prefix Class I undergo all types of stem alternation patterns (except pattern 3, possibly an accident of the sample). It also means that verbs of other prefix classes are constrained to pattern 1, except for Class II, whose verbs may also belong to patterns 2 or 3, or those of Class IV which may select pattern 2. This also means that verbs of prefix classes III, V, VI and VII can only be of stem alternation pattern 1. As we have already pointed out in §2.2, two thirds of the verbs of Class VII do not have a stem alternation pattern. In a similar way, if the figures are observed vertically one may see that pattern 1 is found in verbs that pertain to a greater number of different prefix classes. The phonological diversity of base stems of pattern 1 as given in Table 44, together with the figures in Table 45, suggest that both pattern 1 and prefix Class I work as inflectional defaults for the inflection of verbs in Chichimec.

In the previous sections we have studied prefix classes, stem alternation patterns and their interconnection. In the next section, we study the role of tone in the making of the inflectional paradigm of verbs in Chichimec.

20. Tone and inflection

Chichimec has two tones, low (L) and high (H) (marked with an acute accent). While the tone of inflected forms is clearly registered in Angulo (1933), there is not an explicit analysis of tone, so ours is novel although entirely based on Angulo's materials. In the default case, in an inflected form of a verb, tone is lexically assigned to the stem and then prosodically assigned to the prefix as part of the prosodic word. This is done in such a way that the tone of the prefix ends up receiving the opposite tone of the stem, as illustrated in Table 46, where the same set of prefixes (for example of the 1st person of Class I) when associated to two stems bearing a different tone, end up having opposite tone values.

PRES	<i>e-tú</i>	L-H	<i>é-me</i>	H-L
ANT.PST	<i>tu-tú</i>	L-H	<i>tú-me</i>	H-L
FUT	<i>ga-tú</i>	L-H	<i>gá-me</i>	H-L
REC.PST	<i>ku-tú</i>	L-H	<i>kú-me</i>	H-L
IMM.PST	<i>u-tú</i>	L-H	<i>ú-me</i>	H-L
POT	<i>nu-tú</i>	L-H	<i>nú-me</i>	H-L
SEQ	<i>ra-tú</i>	L-H	<i>rá-me</i>	H-L

Table 46: 1st person of -tú ‘finish’ vs. -me ‘place upright’.

20.1. Distribution and idiosyncracies of inflectional tone

In our sample of 168 verbs, 137 of them (81%) have stems with lexical tone. The verbs in question may undergo stem changes, but tone remains the same across the segmental changes. For the remaining 31 verbs tone has a grammatical function, sometimes as an exponent of person only, sometimes in a combination of person and TAM. Broadly speaking there are three general tendencies: (i) 2nd person receives low tone; (ii) present receives low tone; and (iii) potential receives high tone. However, the role of tone is linked to each verb in such a way that these general tendencies are easily overwritten by lexical tone or other possibilities, making it impossible to conjure a general morphosyntactic rule for tone that works for the system. Because of this, in our presentation of the data, we almost need to proceed verb by verb.

There are four verbs where the forms for the 2nd person have a low tone throughout the paradigm. Two of the verbs (-*sá* ‘win’ and -*sempf* ‘open’) are of pattern 1 and the other two (-*tsó* ‘scold’ and -*tsá* ‘try’) of pattern 3 whose D-stem is used for the 2nd person.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
ANT.PST	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
FUT	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
REC.PST	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
IMM.PST	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
POT.	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
SEQ	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H
NEG	A	A	A	B	<i>sá</i>	<i>sa</i>	<i>sá</i>	<i>tshá</i>	H	L	H	H

PRES	A	D	A	B	<i>tsó</i>	<i>tco</i>	<i>tsó</i>	<i>tshó</i>	H	L	H	H
ANT.PST	A	D	C	B	<i>tsó</i>	<i>tco</i>	<i>zó</i>	<i>tshó</i>	H	L	H	H
FUT	A	D	C	B	<i>tsó</i>	<i>tco</i>	<i>zó</i>	<i>tshó</i>	H	L	H	H
REC.PST	A	D	C	B	<i>tsó</i>	<i>tco</i>	<i>zó</i>	<i>tshó</i>	H	L	H	H
IMM.PST	A	D	A	B	<i>tsó</i>	<i>tco</i>	<i>tsó</i>	<i>tshó</i>	H	L	H	H
POT.	A	D	A	B	<i>tsó</i>	<i>tco</i>	<i>tsó</i>	<i>tshó</i>	H	L	H	H
SEQ	A	D	A	B	<i>tsó</i>	<i>tco</i>	<i>tsó</i>	<i>tshó</i>	H	L	H	H
NEG	A	D	A	B	<i>tsó</i>	<i>tco</i>	<i>tsó</i>	<i>tshó</i>	H	L	H	H

Table 47: Inflectional low tone for 2nd person: verbs -*sá* ‘win’ and -*tsó* ‘scold’.

There are eight verbs that have a different tone conditioned by TAM. One such a verb is -*tú* ‘walk about’ with low tone assigned to the present. As already stated in §3.7 above, this verb is irregular.

	SG/DU	PL	SG/DU	PL	SG/DU	PL
PRES	A	B	<i>tu</i>	<i>tur</i>	L	L
ANT.PST	A	B	<i>tú</i>	<i>túr</i>	H	H

FUT	A	B	tú	túr	H	H
REC.PST	A	B	tu	tur	H	H
IMM.PST	A	B	tú	túr	H	H
POT.	A	B	tú	túr	H	H
SEQ	A	B	tú	túr	H	H
NEG	A'	B'	nú	núr	H	H

Table 48: Tone and TAM for verbs -tú 'walk about' and -te 'go out'.

The remaining seven verbs belong to pattern 5 (-*ɔe* 'give', -*ɔi* 'want', -*ɔo* 'hear', -*ɔu* 'kill', -*he* 'leave', -*her* 'have' and -*ho* 'love'). For such verbs, the potential receives a high tone, as in Table 49. On the lexical tier, the A-stem of such verbs is always low, whereas the B-stem is always high. It is likely that these verbs also have low tone for the present, but this is obscured by the fact that their A-stem has a low tone.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	-ɔi	-ɔi	-ɔi	-rɔi	L	L	L	H
ANT.PST	A	D	C	B	-ɔi	-tí	-ndí	-rɔi	L	H	H	H
FUT	A	D	C	B	-ɔi	-tí	-ndí	-rɔi	L	H	H	H
REC.PST	A	D	C	B	-ɔi	-tí	-ndí	-rɔi	L	H	H	H
IMM.PST	A	D	A	B	-ɔi	-tí	-ɔi	-rɔi	L	H	L	H
POT.	A	D	A	B	-ɔi	-tí	-ɔi	-rɔi	H	H	H	H
SEQ	A	D	A	B	-ɔi	-tí	-ɔi	-rɔi	L	H	L	H
NEG	A	D	A	B	-ɔi	-tí	-ɔi	-rɔi	L	H	L	H

PRES	A	A	A	B	-he	-he	-he	rhé	L	L	L	H
ANT.PST	A	C	C	B	-he	nhí	nhí	rhé	L	H	H	H
FUT	A	C	C	B	-he	nhí	nhí	rhé	L	H	H	H
REC.PST	A	C	C	B	-he	nhí	nhí	rhé	L	H	H	H
IMM.PST	A	C	A	B	-he	nhí	-he	rhé	L	H	L	H
POT.	A	C	A	B	-hé	nhí	-hé	rhé	H	H	H	H
SEQ	A	C	A	B	-he	nhí	-he	rhé	L	H	L	H
NEG	A	C	A	B	-he	nhí	-he	rhé	L	H	L	H

Table 49: Inflectional tone for Potential: verbs -ɔi 'want' and -he 'leave'.

Of the remaining 16 verbs that use tone to mark some inflectional category, they do so displaying a combination of person and TAM. The verbs in question show again idiosyncratic behaviours and we have to treat them separately. For example, the verb -*se* 'say' also has low tone in the present, but it is unique in having a low tone for the 1st person.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	-se	-se	-se	-tshé	L	L	L	H
ANT.PST	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
FUT	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
REC.PST	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
IMM.PST	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
POT.	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
SEQ	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H

NEG	A	A	A	B	-se	-sé	-sé	-tshé	L	H	H	H
-----	---	---	---	---	-----	-----	-----	-------	---	---	---	---

Table 50: Inflectional tone for the verb -se 'say'.

The verbs *-tsá* 'hurt' and *-mế* 'defend' both use low tone for the 2nd person. Both verbs have low tone for the potential, but while *-tsá* 'hurt' has low tone for the present and the recent past in all persons (and also the future for the 3rd person singular and dual), *-mế* 'defend' has high tone for the present.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	D	A	B	-tsa	-tca	-tsa	-tsha	L	L	L	L
ANT.PST	A	D	C	B	-tsá	-tca	-zá	-tshá	H	L	H	H
FUT	A	D	C	B	-tsá	-tca	-za	-tshá	H	L	L	H
REC.PST	A	D	C	B	-tsa	-tca	-za	-tsha	L	L	L	L
IMM.PST	A	D	A	B	-tsá	-tca	-tsá	-tshá	H	L	H	H
POT.	A	D	A	B	-tsa	-tca	-tsa	-tsha	L	L	L	L
SEQ	A	D	A	B	-tsá	-tca	-tsá	-tshá	H	L	H	H
NEG	A	A	A	B	-tsá	-tca	-tsá	-tshá	H	L	H	H

PRES	A	A	A	B	-mế	-mế	-mế	-mhế	H	H	H	H
ANT.PST	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H
FUT	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H
REC.PST	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H
IMM.PST	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H
POT.	A	A	A	B	-mế	-mế	-mế	-mhế	L	L	L	H
SEQ	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H
NEG	A	A	A	B	-mế	-mế	-mế	-mhế	H	L	H	H

Table 51: Inflectional tone for the verbs -tsá 'hurt' and -mế 'defend'.

Other verbs, such as *-tán* 'work', *-tá* 'touch' and *-kér* 'lean on' of pattern 4, but also *-hun* 'throw' and *-hu* 'put' of pattern 5, have a high tone for the 2nd person. All these verbs have high tone for the potential. For verbs of pattern 4, the present has low tone, and their C-stem high tone.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	-tan	-tán	-tan	-rhán	L	H	L	H
ANT.PST	A	A	C	B	-tan	-tán	-rán	-rhán	L	H	H	H
FUT	A	A	C	B	-tan	-tán	-rán	-rhán	L	H	H	H
REC.PST	E	A	C	B	-ndan	-tán	-rán	-rhán	L	H	H	H
IMM.PST	A	A	A	B	-tan	-tán	-tan	-rhán	L	H	L	H
POT.	A	A	A	B	-tán	-tán	-tán	-rhán	H	H	H	H
SEQ	A	A	A	B	-tan	-tán	-tan	-rhán	L	H	L	H
NEG	A	A	A	B	-tan	-tán	-tan	-rhán	L	H	L	H

PRES	A	A	A	A	-hun	-hun	-hun	-hún	L	L	L	H
ANT.PST	A	D	C	A	-hun	-nhún	-nhún	-hún	L	H	L	H
FUT	A	D	C	A	-hun	-nhún	-nhún	-hún	L	H	L	H
REC.PST	A	D	C	A	-hun	-nhún	-nhún	-hún	L	H	L	H
IMM.PST	A	D	A	A	-hun	-nhún	-hun	-hún	L	H	L	H
POT.	A	D	A	A	-hún	-nhún	-hún	-hún	H	H	H	H

SEQ	A	D	A	A	-hun	-nhín	-hun	-hún	L	H	L	H
NEG	A	D	A	A	-hun	-nhín	-hun	-hún	L	H	L	H

Table 52: Inflectional tone for the verbs *-tán* ‘work’ and *-hun* ‘throw’

Such verbs contrast with verbs with no tone changes but with similar stem alternation patterns, like for example the one in Table 53.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	-tan	-tan	-tan	-rhan	L	L	L	L
ANT.PST	A	A	C	B	-tan	-tan	-ran	-rhan	L	L	L	L
FUT	A	A	C	B	-tan	-tan	-ran	-rhan	L	L	L	L
REC.PST	E	A	C	B	-ndan	-tan	-ran	-rhan	L	L	L	L
IMM.PST	A	A	A	B	-tan	-tan	-tan	-rhan	L	L	L	L
POT.	A	A	A	B	-tan	-tan	-tan	-rhan	L	L	L	L
SEQ	A	A	A	B	-tan	-tan	-tan	-rhan	L	L	L	L
NEG	A	A	A	B	-tan	-tan	-tan	-rhan	L	L	L	L

Table 53: Lack of inflectional tone for the verb *-tan* ‘bathe’.

For another small set of verbs, including *-pen* ‘play music’, *-pe* ‘hit/fight’, *-pa* ‘cover’, and *-pin* ‘wait’, the 2nd person is again marked with low tone. The present and the recent past have low tone for all persons, but also the consecutive and the negative of the 3rd person singular and dual.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	A	-pen	-pen	-pen	-pen	L	L	L	L
ANT.PST	A	D	C	B	-pén	-ngwen	-mén	-mbén	H	L	H	H
FUT	A	D	C	B	-pén	-ngwen	-mén	-mbén	H	L	H	H
REC.PST	B	D	C	A	-ngwen	-ngwen	-men	-pen	L	L	L	L
IMM.PST	D	D	D	A	-ngwén	-ngwen	-ngwén	-pén	H	L	H	H
POT.	D	D	D	A	-ngwén	-ngwen	-ngwén	-pén	H	L	H	H
SEQ	A	D	A	B	-pén	-ngwen	-pen	-mbén	H	L	L	H
NEG	D	D	D	A	-ngwén	-ngwen	-ngwen	-pén	H	L	L	H

Table 54: Lack of inflectional tone for the verb *-pen* ‘play music’.

In contrast, the set of verbs including *-pén* ‘kill’, *-páa* ‘look hither’ and *-píi* ‘bring/carry’ have an almost mirror image of the verbs represented by *-pen* ‘play music’ by virtue of having both the 2nd person and the present marked by high tone.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	A	-pén	-pén	-pén	-pén	H	H	H	H
ANT.PST	A	D	C	B	-pen	-ngwén	-mén	-mbén	L	H	H	H
FUT	A	D	C	B	-pen	-ngwén	-mén	-mbén	L	H	H	H
REC.PST	B	D	C	A	-mben	-ngwén	-men	-pén	L	H	L	H
IMM.PST	D	D	D	A	-ngwen	-ngwén	-ngwen	-pen	L	H	L	L
POT.	D	D	D	A	-ngwen	-ngwén	-ngwen	-pen	L	H	L	L
SEQ	A	D	A	B	-pen	-ngwén	-pen	-mbén	L	H	L	H
NEG	D	D	D	A	-ngwen	-ngwén	-ngwen	-pen	L	H	L	L

Table 55: Lack of inflectional tone for the verb *-pén* ‘kill’.

Finally, the verb *-me* ‘stand up’ also has high tone for the 2nd person, but both present and negative receive a low tone that overwrites the high tone of the 2nd person.

	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL	1	2	3SG/DU	3PL
PRES	A	A	A	B	<i>-me</i>	<i>-me</i>	<i>-me</i>	<i>-mer</i>	L	L	L	L
ANT.PST	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
FUT	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
REC.PST	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
IMM.PST	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
POT.	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
SEQ	A	A	A	B	<i>-me</i>	<i>-mé</i>	<i>-me</i>	<i>-mer</i>	L	H	L	L
NEG	A	A	A	B	<i>-me</i>	<i>-me</i>	<i>-me</i>	<i>-mer</i>	L	L	L	L

Table 56: Lack of inflectional tone for the verb *-me* ‘stand up’.

21. Conclusions

In this paper, we have revised the conjugation classes proposed by Jaime de Angulo in 1933 for the Chichimec language of Mexico and have proposed new ways to organize the complexities of its verbal inflection. In general, verb inflection in Chichimec displays a high degree of morphological complexity which cannot be reduced to phonological/phonetic processes nor responds to any morphosyntactic requirements, and thus only respond to an autonomous morphological level in the language.

While there are a number of tendencies in the selection of a given stem alternation pattern and in the selection of a certain set of prefix class, there are many exceptions, making the tendencies probabilistic in character so that all verbs need to be listed for class membership and tone pattern.

Tone inflection seems to operate in number of levels: At a lexical level, verbs are assigned an underlying tone. However, in some conjugations, TAM and person categories have associated tone specifications. Complex tone patterns occur in a small set of verbs, but these coincide with morphological complexity at other levels, such as stem alternation patterns.

The Chichimec data is theoretically relevant as it is an instance in which tone assignment depends on multiple factors that have to be computed simultaneously to yield the correct surface form. It is a matter of further research to understand the processes by which a native speaker factors out the relevance of the categories expressed by tone for each verb conjugation.

Overall, regardless of the paucity of the data analysed, our findings show a general pattern of inflection consistent with other Oto-Manguean languages; most conspicuously the inflectional idiosyncrasies and the complexity of different morphological subsystems.

Acknowledgements

We presented a first version of this paper at the special session organized by Heriberto Avelino on the morphological complexity of the languages of the Americas at the Linguistic Society of America Annual Meeting in Portland, Oregon in January 2012.

We are very grateful for all comments received then. We also want to thank Greville Corbett and Dunstan Brown for their invaluable comments to an earlier written version of this paper and to Matthew Baerman and Timothy Feist for their respective reviews. Errors and deficiencies remain our own responsibility.

Abbreviations

ANT anterior past; DU dual; EXCL exclusive; FUT future; IMM immediate; INCL inclusive; NEG negative; PL plural; POT potential; PRS present; PST past; REC recent; SEQ sequential; SG singular.

References

- Angulo, Jaime de. 1933. The Chichimeco Language (Central Mexico). *International Journal of American Linguistics* 7-3/4: 152-194.
- Feist, Timothy and Enrique L. Palancar. 2015. Oto-Manguean Inflectional Class Database: Chichimec. University of Surrey. <http://dx.doi.org/10.15126/SMG.28/1.20>
- Kemmer, Suzanne. 1993. *The Middle Voice*. [Typological Studies in Language 23]. Amsterdam/ Philadelphia: John Benjamins.
- Lastra, Yolanda. 1984. Chichimeco Jonaz, in Munro S. Edmonson S. (ed.), *Supplement to the Handbook of Middle American Languages: Linguistics*, pp. 20-42. Austin: University of Texas Press.

12. Inflectional class interaction in Matlatzinca

Enrique L. Palancar
SeDyL (UMR8202), CNRS

Leonardo Carranza

Abstract

*In this chapter, we propose that verbs in Matlatzinca fall into six different inflectional classes. The interest of the study is on class membership and class interaction. Many verbs in Matlatzinca come in pairs (some in triplets) reflecting typical valence pairings, but each member in such pairings is inflected as a member of a different inflectional class without additional changes at stem level. For example, a form such as *to-'uši* means 'I pushed it', it belongs to one of the transitive classes and evokes a causative scenario, while a form *ta-'uši* is from a different inflectional class, an intransitive one, and means 'I pushed' as an intransitive atelic activity. Some of the pairings appear to be productive some are not. In the chapter, we discuss this phenomenon at length in a sample of 600 verbs.*

22. Introduction

Matlatzinca is a language that belongs to the Oto-Pamean branch of Oto-Manguan. In this chapter, we first propose an analysis of Matlatzinca verbal inflection in terms of six inflectional classes. In this, we elaborate on the proposal articulated in Carranza (2014) by taking it further in depth, in detail and in theoretical implications. But the emphasis of the chapter is on showing that verbs in Matlatzinca may and often do display multiple class membership, which is a property that we consider to be typologically remarkable. For example, a transitive verb conjugated in one class often has an intransitive correlate conjugated in another. When this happens, the verbs in question display slight differences in meaning which are of the type commonly observed in valence changing mechanisms in other languages. But in Matlatzinca verbs do this without additional changes at stem level, only by a change in class, e.g. the form *to-'uši* means 'I pushed it' while *ta-'uši* means 'I pushed' as an atelic activity. Do these forms instantiate two lexical entries or two different aspects of the same verb? This is an important theoretical question. If we were to treat them as two different aspects of the same verb, it would mean that for every verb we would have to postulate that every verb has at least two paradigms, a transitive one and an intransitive one, but as there are many verbs that do not display the same behaviour as *-'uši*, we would be forced to say something about those instances where one of the paradigms is not filled. Because of this and other arguments that we will treat at greater length in section 4.2, we prefer to view them as instantiating two different, but semantically related, verbs. Such verbs enter, in turn, into interesting patterns of inflectional class interaction.

The closest language to Matlatzinca is Tlahuica (a.k.a. Ocuiltec), whose inflectional system has been described as having inflectional classes but with a different distribution. These two languages are the only surviving members of a common ancestor language once spoken around the 5th century AD by the different Matlatzinca tribes living in the Toluca Valley in Central Mexico. This valley was known as ‘Matlatzinco’ to Nahuatl speakers. This is how we obtain the term ‘Matlatzinca’ for both the inhabitants and their language.

Matlatzinca and Tlahuica form the Atzincan branch of Otomian, which also includes Mazahua and the Otomi languages as the most closely related languages, but at a considerable distance. The language is nowadays spoken in the village of San Francisco Oxtotilpan (see Figure 1), in the municipality of Temascaltepec, about 40 km from Toluca, the capital city of the State of Mexico. It is also the first language of a few hundred migrants living in Mexico City and Toluca, who rarely use it outside the village. The overall number of speakers is 1500; roughly half of them still live in San Francisco Oxtotilpan the whole year round.



Figure 1: Location of Chichimec in Mexico.

Apart from the missionary work by Andrés de Castro, Diego Basalenque and Miguel de Guevara, and the brief treatment in Soustelle (1937), the modern emphasis in the study of Matlatzinca has been on the phonology, e.g. Bartholomew (1966a), Cazes (1971), and more recently Pascacio (2006) and Pérez (2007); the latter containing a useful list of a thousand cognate terms with Tlahuica. Apart from phonologies, Pascacio (2011) proposes a historical reconstruction of the pronominal system and there is a grammatical sketch in Bartholomew (1966b), but the grammar of the language remains unexplored.

Taking a step in this direction, Carranza (2013) is the first modern exploration of some morphosyntactic aspects of the verbal inflection such as the marking of TAM. However, the approach overlooks inflectional allomorphy. In this regard, inspired by the work on Tlahuica by Martinez (2012), Carranza (2014) puts forward an analysis of verbs as organized in six inflectional classes which we retake here in section 3.2. This work led us to create an initial database of 374 verbal entries (Feist, Palancar and Carranza, 2015), which is now publicly accessible at <<http://www.oto-manguean.surrey.ac.uk/Search/MAT>>. The present study is based on a sample from that database.

The article has a simple structure and is divided in sections as follows. We first present general notions about the verbal inflection of Matlatzinca. In Section 3, we

introduce the different inflectional classes by focusing on their formal make-up and then in section 4 we concentrate on the phenomenon of class interaction in detail, followed by a discussion. Section 5 concludes.

23. Verbal inflection in Matlatzinca

Verbs inflect for at least seven tense-aspect-mood (TAM) values by means of prefixes or prefixal chains. The prefixes also encode information about person and number of subject in a cumulative way. Table 1 shows the various inflected forms for the 3rd person singular of the transitive verb *-táni* ‘buy’ in all seven TAM values.

TAM	<i>-táni</i> ‘buy’	Loose translation
INCPL	ron-táni	‘they buy it’ / ‘they’re buying it’
CPL	tu-táni-hə ☒	‘they’ve bought it’
PST	roma’-táni	‘they bought it’
IMPF	rona’-táni	‘they were buying it’
POT	tátu-táni-hə ☒	‘they’d buy it’
FUT	táho’-táni	‘they’ll buy it’
FUT.PRG	táron-táni	‘they’ll be buying it’

Table 1: TAM values in Matlatzinca verbs.

As noted, information about person of subject is encoded along with TAM. In the default case, prefixes also encode a distinction of number of subject involving singular, dual and plural. This is shown in the incomplete of the verb *-táni* ‘buy’ in Table 2, where it may be seen that dual and plural prefixes for the 1st and 2nd person contrast in that the plural ones have aspirated onsets. For the potential of transitive verbs, the pattern changes and number of subject is encoded by means of suffixes (i.e., *-wewí* for dual; *-hə☒* for plural). The same happens in the completive of transitive verbs, but only in the forms of the 3rd person.

<i>-táni</i> ‘buy’	INCPL	POT	CPL
1SG	tuh-táni	ru-táni	to-táni
1DU.INCL	kwen-táni	ru-táni -wewí	ko-táni
1DU.EXCL	kwén-táni	ru-táni -we-bí	ko-bú-táni
1PL.INCL	khwen-táni	ru-táni -hə☒	kho-táni
1PL.EXCL	khwén-táni	ru-táni -hə☒-bí	kho-bú-táni
2SG	’i-táni	ri-táni	’o-táni
2DU	čen-táni	ri-táni -wewí	čo-táni
2PL	čhen-táni	ri-táni -hə☒	čo-táni
3SG	kuh-táni	tá-tu-táni	tu-táni
3DU	kwén-táni	tá-tu-táni -wewí	tu-táni -wew í
3PL	ron-táni	tá-tu-táni - hə☒	tu-táni - hə☒

Table 2: Number of subject in Matlatzinca verbs.

The data in Table 2 also show that there is a clusivity opposition, where the inclusive form is the least marked member. When number is encoded by prefixes, depending on the subparadigm the exclusive value is encoded suprasegmentally by a high tone or by the prefix *bú-*.¹³⁰ When number is encoded with a suffix, exclusive is marked by *-bí* (in the dual, *-wewí-bí* surfaces as *-we-bí*).

Finally, intransitive verbs with a semantic motion component may also inflect for deixis by means of a translocative subparadigm. The translocative is used when the action is presented as occurring away from the deictic centre or while moving away from it. In contrast, when the default subparadigm is used with similar verbs, speakers may interpret that the action involves a cislocative component (i.e., movement towards the deictic centre). The translocative is encoded by the prefix *h-*, which replaces the glottal coda of the prefix. Examples appear in Table 3 with the verb *-čapi* ‘jump’, only in the singular. The translocative is only available in the completive and the potential, so other subparadigms without a corresponding translocative, like for example the incomplete, are unmarked for deixis.

<i>-čapi</i> ‘jump’		unmarked	CISLOC ‘hither’ or ‘away from’	TRANSLOC ‘thither’
CPL	1SG		ta’-čapi	ta- h -čapi
	2SG		’a’-čapi	’a- h -čapi
	3SG		ma’-čapi	ma- h -čapi
POT	1SG		ruh-čapi	ru- h -čapi
	2SG		ri’i’-čapi	ri’i’- h -čapi
	3SG		tá’-čapi	tá- h -čapi
INCPL	1SG	tu-čapi		
	2SG	’i-čapi		
	3SG	kučapi		

Table 3: The cislocative and the translocative in Matlatzinca.

24. Inflectional classes in Matlatzinca verbs

Verbs in Matlatzinca fall into six inflectional classes. There is a very strong correlation between transitivity and class membership: all transitive verbs of the language fall into three of the six classes (and only in those three), and all intransitive verbs fall in the three remaining ones. This is a distributional pattern which is not seen elsewhere in Oto-Manguean. For example, the closest we can get to this situation is found in Otomi languages with inflectional classes. There an inflectional class treated as Class II in Tilapa Otomi and Acapulco Otomi (see Palancar 2012 and Hernández Green, this volume) only comprises intransitive verbs. But in these Otomi languages, intransitive verbs are also found in other classes, so the correlation breaks once we step out of Class II verbs.

¹³⁰ There is also reduplication in the future, which can be observed in Table 3.

Another interesting characteristic of the inflectional class system of Matlatzinca is that at least some of the classes retain semantic cores. This may be seen when one observes the semantics of the members of the intransitive classes in isolation from a certain distance, but especially so when one compares members across classes. We can do this in Matlatzinca because it often happens that a certain verb inflected in one class also appears inflected in another. When this holds there is always a change in meaning, typical of valence alternations. For example, the verb *-’iwi* is found in one of the transitive classes and it has a causative meaning ‘make somebody go to sleep’, but the same verb or a closely related one sharing the same stem *-’iwi* is found as a member of one of the intransitive classes to refer to the activity of ‘sleeping’. One way to talk about this phenomenon would be to say that there is only one lexical entry *’IWI* (i.e. one verb with one stem *-’iwi*) that is a member of different classes. Another way would be to think that there are two different, but related verbs, that share a core lexical meaning and the same stem *-’iwi*: one that is transitive and means ‘make go to sleep’ and another one that is intransitive and means ‘sleep’. As it happens, in Matlatzinca these two verbs are members of different inflectional classes, that is, they inflect according to two different paradigms. We discuss this type of verb in more detail in Section 4. To understand inflectional class interactions involving such items, it is first convenient to pay attention to relevant aspects of the formal distinctions involved in the make-up of the classes as genuine inflectional classes. This is what we do next. We shall start with transitive verbs.

24.1. Transitive classes

The three transitive classes of Matlatzinca present an inflectional contrast in only four subparadigms: incomplete, imperfect, future and future progressive. This means that all three classes share the same inflections for what is left of the paradigm. We use Roman numerals to refer to the transitive classes and Arabic ones to refer to the intransitive ones. From a strict morphological point of view, Class I could be said to serve as default. Classes II and III are more marked and they are based either on Class I or on the default class of intransitive verbs, to which we will refer as Class 1. Marking contrasts are shown in Table 4. We have used shading to highlight inflectional markers shared by transitive Class I and intransitive Class 1. We have taken such markers to work as morphological defaults.

		‘buy’ Class I (tr)		‘leave behind’ Class II (tr)		‘kick’ Class III (tr)		‘weep’ Class 1 (intr)	
INCP	1SG	tuh- ¹³¹	táni	tuh-t-	habi	tu-tú-	nhəti	tu-	kwe’e
	2SG	’ih-	táni	’ih-t-	habi	’i-tú-	nhəti	’i-	kwe’e
	3SG	kuh-	táni	kuh-t-	habi	ku-tú-	nhəti	ku-	kwe’e
	1DU.INCL	kwen-	táni	kwen-t-	habi	kwen-tú-	nhəti	kwen-	kwe’e
	1DU.EXCL	kwén-	táni	kwén-t-	habi	kwén-tú-	nhəti	kwén-	kwe’e
	2DU	čén-	táni	čén-t-	habi	čén-tú-	nhəti	čén-	kwe’e
	3DU	kwén-	táni	kwén-t-	habi	kwén-tú-	nhəti	kwén-	kwe’e
	1PL.INCL	khwen-	táni	khwen-t-	habi	khwen-tú-	nhəti	khwen-	kwe’e
	1PL.EXCL	khwén-	táni	khwén-t-	habi	khwén-tú-	nhəti	khwén-	kwe’e
	2PL	čhen-	táni	čhen-t-	habi	čhen-tú-	nhəti	čhen-	kwe’e
3PL	ron-	táni	ron-t-	habi	ron-tú-	nhəti	ron-	kwe’e	
IMPF	1SG	tuh-	táni	tuh-t-	habi	tu-tú-	nhəti	tu-	kwe’e
	2SG	’ih-	táni	’ih-t-	habi	’i-tú-	nhəti	’i-	kwe’e
	3SG	kuh-	táni	kuh-t-	habi	ku-tú-	nhəti	ku-	kwe’e

¹³¹ Glottal in prefixes are only allowed with stems with voiceless plosive onsets (i.e., /p,t,k/).

	1DU.INCL	kwena'- táni	kwena'-t- habi	kwena'-tú- nhàti	kwena'- kwe'e
	1DU.EXCL	kwéna'- táni	kwéna'-t- habi	kwéna'-tú- nhàti	kwéna'- kwe'e
	2DU	čena'- táni	čena'-t- habi	čena'-tú- nhàti	čena'- kwe'e
	3DU	kwéna'- táni	kwéna'-t- habi	kwéna'-tú- nhàti	kwéna'- kwe'e
	1PL.INCL	khwena'- táni	khwena'-t- habi	khwena'-tú- nhàti	khwena'- kwe'e
	1PL.EXCL	khwéna'- táni	khwéna'-t- habi	khwéna'-tú- nhàti	khwéna'- kwe'e
	2PL	čhena'- táni	čhena'-t- habi	čhena'-tú- nhàti	čhena'- kwe'e
	3PL	rona'- táni	rona'-t- habi	rona'-tú- nhàti	rona'- kwe'e
FUT	1SG	ruho'- táni	ruho'-rú- habi	ruho'-rú- nhàti	ruho'- kwe'e
	2SG	ri'o'- táni	ri'o'-rú- habi	ri'o'-rú- nhàti	ri'o'- kwe'e
	3SG	táho'- táni	táho'-rú- habi	táho'-rú- nhàti	táho'- kwe'e
	1DU.INCL	ruko'- táni	ruko'-rú- habi	ruko'-rú- nhàti	ruko'- kwe'e
	1DU.EXCL	ruko'ko'- táni	ruko'ko'-rú- habi	ruko'ko'-rú- nhàti	ruko'ko'- kwe'e
	2DU	ričo'- táni	ričo'-rú- habi	ričo'-rú- nhàti	ričo'- kwe'e
	3DU	táko'- táni	táko'-rú- habi	táko'-rú- nhàti	táko'- kwe'e
	1PL.INCL	rukho'- táni	rukho'-rú- habi	rukho'-rú- nhàti	rukho'- kwe'e
	1PL.EXCL	rukho'kho'- táni	rukho'kho'-rú- habi	rukho'kho'-rú- nhàti	rukho'kho'- kwe'e
	2PL	ričo'- táni	ričo'-rú- habi	ričo'-rú- nhàti	ričo'- kwe'e
3PL	táho'- táni	táho'-rú- habi	táho'-rú- nhàti	táho'- kwe'e	
FUT.PRG	1SG	ruh- táni	ruh-t- habi	ruh- nhàti	ruh- kwe'e
	2SG	ri'ih- táni	ri'ih-t- habi	ri'ih- nhàti	ri'ih- kwe'e
	3SG	táh- táni	táh-t- habi	táh- nhàti	táh- kwe'e
	1DU.INCL	rukwen- táni	rukwen-t- habi	rukwen-tú- nhàti	rukwen- kwe'e
	1DU.EXCL	rukwén- táni	rukwén-t- habi	rukwén-tú- nhàti	rukwén- kwe'e
	2DU	ričen- táni	ričen-t- habi	ričen-tú- nhàti	ričen- kwe'e
	3DU	tákwén- táni	tákwén-t- habi	tákwén-tú- nhàti	tákwén- kwe'e
	1PL.INCL	rukhwén- táni	rukhwén-t- habi	rukhwén-tú- nhàti	rukhwén- kwe'e
	1PL.EXCL	rukhwén- táni	rukhwén-t- habi	rukhwén-tú- nhàti	rukhwén- kwe'e
	2PL	ričhen- táni	ričhen-t- habi	ričhen-tú- nhàti	ričhen- kwe'e
	3PL	táron- táni	táron-t- habi	táron-tú- nhàti	táron- kwe'e

Table 4: The morphological make-up of the transitive classes, part 1.

Some observations are in order with regard to Table 4. Shaded areas indicate that Class I and Class 1 contrast only in the singular portion of the incomplete and the imperfect; the rest are default exponents. For that portion of the paradigm, Class II is clearly based on Class I plus *t-*, while Class III is based on Class 1 plus *tú-*. Class III has the marker *tú-* throughout. The rhotic variant *-rú-* in the future can be accounted for in morphophonological terms as triggered by the rule in (1).

$$(1) /...V_{MID}'-tV/ \rightarrow [...V_{MID}'-rV]$$

The occurrence of *tú-* suggests derivation, possibly with a causative function. This would cancel any analysis of Class III as a genuine inflectional class because its members could be better treated as derived verbs from intransitive verbs of Class 1 by means of *tú-*. The same would apply for Class II with its prefix *t-*. Although such an analysis would be adequate from a historical point of view, there are, however, two very strong arguments in favor of treating Class III (and II for that matter) as a genuine inflectional class.

The first argument is a morphological one. If *tú-* and *t-* were true derivational markers, we would expect them to be associated with the stem in an invariable fashion, and appear in all inflected forms of the alleged derived lexeme. This does not happen. Already shown in Table 4 above, *tú-* is missing in the singular cells of the future progressive of Class III, where we find the default exponent. Similarly, shown in Table 5 below, neither *tú-* nor *t-* occur in the complete and the potential, where all transitive classes have the same exponents (notice that transitive verbs contrast in turn with the inflection of intransitive verbs, represented by Class 1), or in the past, where all four classes have the same markers. The presence of default exponent makes an analysis of *tú-* and *t-* as derivational markers less satisfying.

		'buy' Class I (tr)		'leave behind' Class II (tr)		'kick' Class III (tr)		'weep' Class 1 (intr)	
CPL	1SG	to-	táni	to-	habi	to-	nhəti	ta'-	kwe'e
	2SG	'o-	táni	'o-	habi	'o-	nhəti	'i'-	kwe'e
	3SG	tu-	táni	tu-	habi	tu-	nhəti	ka'-	kwe'e
	1DU.INCL	ko-	táni	ko-	habi	ko-	nhəti	kwe'-	kwe'e
	1DU.EXCL	kobú-	táni	kobú-	habi	kobú-	nhəti	kwé'-	kwe'e
	2DU	čo-	táni	čo-	habi	čo-	nhəti	čé'-	kwe'e
	3DU	tu-	táni -wewí	tu-	habi -wewí	tu-	nhəti -wewí	kwé'-	kwe'e
	1PL.INCL	kho-	táni	kho-	habi	kho-	nhəti	khwe'-	kwe'e
	1PL.EXCL	khobú-	táni	khobú-	habi	khobú-	nhəti	khwé'-	kwe'e
2PL	čo-	táni	čo-	habi	čo-	nhəti	ché'-	kwe'e	
3PL	tu-	táni -həǂ	tu-	habi -həǂ	tu-	nhəti -həǂ	ro'-	kwe'e	
POT	1SG	ru-	táni	ru-	habi	ru-	nhəti	ruh-	kwe'e
	2SG	ri-	táni	ri-	habi	ri-	nhəti	ri'i'-	kwe'e
	3SG	tátu-	táni	tátu-	habi	tátu-	nhəti	tá'-	kwe'e
	1DU.INCL	ru-	táni -wewí	ru-	habi -wewí	ru-	nhəti -wewí	rukwe'-	kwe'e
	1DU.EXCL	ru-	táni -webí	ru-	habi -webí	ru-	nhəti -webí	rukwe'-	kwe'e
	2DU	ri-	táni -wewí	ri-	habi -wewí	ri-	nhəti -wewí	riče'-	kwe'e
	3DU	tátu-	táni -wewí	tátu-	habi -wewí	tátu-	nhəti -wewí	tákwé'-	kwe'e
	1PL.INCL	ru-	táni -həǂ	ru-	habi -həǂ	ru-	nhəti -həǂ	rukhwé'-	kwe'e
	1PL.EXCL	ru-	táni -həǂbí	ru-	habi -həǂbí	ru-	nhəti -həǂbí	rukhwé'-	kwe'e
2PL	ri-	táni -həǂ	ri-	habi -həǂ	ri-	nhəti -həǂ	riče'-	kwe'e	
3PL	tátu-	táni -həǂ	tátu-	habi -həǂ	tátu-	nhəti -həǂ	táro'-	kwe'e	
PST	1SG	ta'-	táni	ta'(-)	habi	ta'-	nhəti	ta'-	kwe'e
	2SG	'a'-	táni	'a'(-)	habi	'a'-	nhəti	'a'-	kwe'e
	3SG	ma'-	táni	ma'(-)	habi	ma'-	nhəti	ma'-	kwe'e
	1DU.INCL	kwa'-	táni	kwa'(-)	habi	kwa'-	nhəti	kwa'-	kwe'e
	1DU.EXCL	kwa'-	táni	kwa'(-)	habi	kwa'-	nhəti	kwa'-	kwe'e
	2DU	ča'-	táni	ča'(-)	habi	ča'-	nhəti	ča'-	kwe'e
	3DU	kwa'-	táni	kwa'(-)	habi	kwa'-	nhəti	kwa'-	kwe'e
	1PL.INCL	khwa'-	táni	khwa'(-)	habi	khwa'-	nhəti	khwa'-	kwe'e
	1PL.EXCL	khwa'-	táni	khwa'(-)	habi	khwa'-	nhəti	khwa'-	kwe'e
2PL	ča'-	táni	ča'(-)	habi	ča'-	nhəti	ča'-	kwe'e	
3PL	roma'-	táni	roma'(-)	habi	roma'-	nhəti	roma'-	kwe'e	

Table 5: The morphological make-up of the transitive classes, part 2.

The second argument is a lexical one. Class III is the only productive transitive class. All transitive verbs which are borrowed from Spanish, already established or just spontaneous, are inflected as Class III verbs. No such verbs can then be inflected as intransitives of Class 1 if one would like to convey an intransitive meaning, so whatever the relation there might have been between Class III and Class 1, it is now dead. Examples are shown in Table 6 with forms of the 1st person singular of the incomplete. Notice that transitive Class I, although being the morphological default, cannot be used for this purpose; so it is a closed class.¹³² Spanish verbs are integrated into the Matlatzinca system via the Spanish past participle (in its default, masculine form). The coda /do/ is accommodated phonologically to /du/, which is further flapped to /ru/ to make it sound more native, which is a sign of more integrated loanwords, as in (b), included in our sample.

	Class III	Class I	Spanish PTPL/ADJ	(verb)
a.	tutú-meskladu	*tu(h)-meskladu	'mix'	mezclado ← MEZCLAR
	tutú-'apoyadu	*tu(h)-'apoyadu	'lean'	apoyado ← APOYAR
	tutú-kambiadu	*tuh-kambiadu	'change'	cambiado ← CAMBIAR
	tutú-serradu	*tu(h)-serradu	'close'	cerrado ← CERRAR
	tutú-'abridu	*tu(h)-'abridu	'open'	abrido ¹³³ ← ABRIR
	tutú-dormidu	*tu(h)-dormidu	'make sleep'	dormido ← DORMIR
b.	tutú-gastaru	*tu(h)-gastaru	'spend'	gastado ← GASTAR

¹³² No novel intransitive verb can be created and used as a Class 1 verb.

¹³³ The form *abrido* is a regularized outcome; the actual participial form in Spanish is irregular *abierto*.

tutú-pensaru	*tuh-pensaru	‘think’	pensado	←	PENSAR
tutú-mandaru	*tu(h)-mandaru	‘send’	mandado	←	MANDAR
tutú-čansearu	*tu(h)-čansearu	‘tease’	chanceado	←	CHANCEAR (Mex. ‘flirt’)
tutú-manxu	*tu(h)-manxu	‘tame’	manso		

Table 6: Transitive loan verbs in Matlatzinca from Spanish.

Apart from loanwords, class membership of native verbs to Classes I-III is lexically determined. However, membership of Class II is restricted to (transitive) verbs that have a stem with a glottal onset /h/ or /ʔ/. This phonological property, although characteristic of Class II, is not exclusive to it. In our sample, there are two native verbs in Class I with a similar glottal onset (eg. *-’api* ‘dig’ and *-’eki* ‘carve’), and many in Class III (eg. *-’huhti* ‘inflate’, *-’hite* ‘bathe’, *-’eya* ‘heal’, *-’eri* ‘empty’, etc.).

24.2. Intransitive classes

As noted, we will refer to intransitive classes using Arabic numbers. Class 1 is the least marked of the three. The morphological make-up of Class 2 involves the prefix *te-*; and of Class 3 the prefix *n-*. We will first study the relation between Class 1 and 2. This is shown in Table 7, where we also use shading to indicate default exponence (i.e., cells with identical material for Class 1 and Class I). Table 7 shows that the inflectional material for the past, the future, and the future progressive of Class 2 verbs is regularly based on the default exponence plus *te-*. This is true for all cells, except for the 3rd person singular of the future progressive where we find the sequence *táre-*. This is obviously related to the default exponent *táh-* plus *te-*, but it is not derived from it synchronically by any non-ad-hoc morphophonological rule.

		‘weep’ Class 1 (intr)	‘get lost’ Class 2 (intr)	‘buy’ Class I (tr)
PST	1SG	ta’- kwe’e	ta’-te- baati	ta’- táni
	2SG	’a’- kwe’e	’a’-te- baati	’a’- táni
	3SG	ma’- kwe’e	ma’-te- baati	ma’- táni
	1DU.INCL	kwa’- kwe’e	kwa’-te- baati	kwa’- táni
	1DU.EXCL	kwa’- kwe’e	kwa’-te- baati	kwa’- táni
	2DU	ča’- kwe’e	ča’-te- baati	ča’- táni
	3DU	kwa’- kwe’e	kwa’-te- baati	kwa’- táni
	1PL.INCL	khwa’- kwe’e	khwa’-te- baati	khwa’- táni
	1PL.EXCL	khwa’- kwe’e	khwa’-te- baati	khwa’- táni
	2PL	ča’- kwe’e	ča’-te- baati	ča’- táni
3PL	roma’- kwe’e	roma’-te- baati	roma’- táni	
FUT	1SG	ruho’- kwe’e	ruho’-re- baati	ruho’- táni
	2SG	ri’o’- kwe’e	ri’o’-re- baati	ri’o’- táni
	3SG	táho’- kwe’e	táho’-re- baati	táho’- táni
	1DU.INCL	ruko’- kwe’e	ruko’-re- baati	ruko’- táni
	1DU.EXCL	ruko’ko’- kwe’e	ruko’ko’-re- baati	ruko’ko’- táni
	2DU	ričo’- kwe’e	ričo’-re- baati	ričo’- táni
	3DU	táko’- kwe’e	táko’-re- baati	táko’- táni
	1PL.INCL	rukho’- kwe’e	rukho’-re- baati	rukho’- táni
	1PL.EXCL	rukho’kho’- kwe’e	rukho’kho’-re- baati	rukho’kho’- táni
	2PL	ričo’- kwe’e	ričo’-re- baati	ričo’- táni
3PL	táho’- kwe’e	tároho’-re- baati	táho’- táni	
FUT.PRG	1SG	ruh- kwe’e	ruh-te- baati	ruh- táni
	2SG	ri’íh- kwe’e	ri’í’-te- baati	ri’íh- táni
	3SG	táh- kwe’e	tá-re- baati	táh- táni

1DU.INCL	rukwen-	kwe'e	rukwen-te-	baati	rukwen-	táni
1DU.EXCL	rukwén-	kwe'e	rukwén-te-	baati	rukwén-	táni
2DU	ričen-	kwe'e	ričen-te-	baati	ričen-	táni
3DU	tákwén-	kwe'e	tákwén-te-	baati	tákwén-	táni
1PL.INCL	rukhwén-	kwe'e	rukhwén-te-	baati	rukhwén-	táni
1PL.EXCL	rukhwén-	kwe'e	rukhwén-te-	baati	rukhwén-	táni
2PL	ričhen-	kwe'e	ričhen-te-	baati	ričhen-	táni
3PL	táron-	kwe'e	táron-te-	baati	táron-	táni

Table 7: The morphological make-up of Class 2, part 1.

Table 7 is uninformative about the source of Class 2. In Table 8, however, the singular cells of the incomplete and the imperfect suggest that the morphological source of Class 2 is Class 1, and not Class I.

		'weep'		'get lost'		'buy'	
		Class 1 (intr)		Class 2 (intr)		Class I (tr)	
INCPL	1SG	tu-	kwe'e	tu-te-	baati	tuh-	táni
	2SG	'í-	kwe'e	'í-te-	baati	'ih-	táni
	3SG	ku-	kwe'e	ku-te-	baati	kuh-	táni
	1DU.INCL	kwen-	kwe'e	kwen-te-	baati	kwen-	táni
	1DU.EXCL	kwén-	kwe'e	kwén-te-	baati	kwén-	táni
	2DU	čen-	kwe'e	čen-te-	baati	čen-	táni
	3DU	kwén-	kwe'e	kwén-te-	baati	kwén-	táni
	1PL.INCL	khwen-	kwe'e	khwen-te-	baati	khwen-	táni
	1PL.EXCL	khwén-	kwe'e	khwén-te-	baati	khwén-	táni
	2PL	čhen-	kwe'e	čhen-te-	baati	čhen-	táni
3PL	ron-	kwe'e	ron-te-	baati	ron-	táni	
IMPF	1SG	tu-	kwe'e	tu-te-	baati	tuh-	táni
	2SG	'í-	kwe'e	'í-te-	baati	'ih-	táni
	3SG	ku-	kwe'e	ku-te-	baati	kuh-	táni
	1DU.INCL	kwena'-	kwe'e	kwena'-te-	baati	kwena'-	táni
	1DU.EXCL	kwéna'-	kwe'e	kwéna'-te-	baati	kwéna'-	táni
	2DU	čena'-	kwe'e	čena'-te-	baati	čena'-	táni
	3DU	kwéna'-	kwe'e	kwéna'-te-	baati	kwéna'-	táni
	1PL.INCL	khwena'-	kwe'e	khwena'-te-	baati	khwena'-	táni
	1PL.EXCL	khwéna'-	kwe'e	khwéna'-te-	baati	khwéna'-	táni
	2PL	čhena'-	kwe'e	čhena'-te-	baati	čhena'-	táni
3PL	rona'-	kwe'e	rona'-te-	baati	rona'-	táni	

Table 8: The morphological make-up of Class 2, part 2.

Had the derivation been from Class I, in principle we would have expected the sequence /CVhtV/. Such a sequence is allowed by the morphotactics as it is found in singular forms of Class I (e.g. /tuhtáni/; /'íhtáni/ and /kuhtáni/).¹³⁴ But as such forms involve a stressed vowel on the stem which is absent in the prefix sequence, we cannot be sure. In contrast, the completive and the potential show stronger evidence for a derivation from Class 1. This is shown in Table 9 in the areas marked with dark

¹³⁴ Notice as well the Class 2 forms *tahte-* and *ruhte-* for the 1st singular completive and potential, respectively, in Table 9.

shading, where the inflectional marking for Class 2 is clearly based on Class 1 plus prefix *te-* as Class 1 and Class I do not have formal overlaps.¹³⁵

		‘weep’	‘get lost’	‘buy’	
		Class 1 (intr)	Class 2 (intr)	Class I (tr)	
CPL	1SG	ta’- kwe’e	tah- te- baati	to- táni	
	2SG	’i’- kwe’e	’e’- baati	’o- táni	
	3SG	ka’- kwe’e	re- baati	tu- táni	
	1DU.INCL	kwe’- kwe’e	kwe’- baati	ko- táni	
	1DU.EXCL	kwé’- kwe’e	kwé’-be- baati	kobú- táni	
	2DU	če’- kwe’e	če’- baati	čo- táni	
	3DU	kwé’- kwe’e	kwé’-re- baati	tu- táni -wewí	
	1PL.INCL	khwe’- kwe’e	khwe’- baati	kho- táni	
	1PL.EXCL	khwé’- kwe’e	khwé’-be- baati	khobú- táni	
	2PL	če’- kwe’e	če’- baati	čo- táni	
	3PL	ro’- kwe’e	ro’-re- baati	tu- táni -həʔ	
	POT	1SG	ruh- kwe’e	ruh- te- baati	ru- táni
		2SG	ri’i’- kwe’e	ri’e- baati	ri- táni
3SG		tá’- kwe’e	tá’-re- baati	tátu- táni	
1DU.INCL		rukwe’- kwe’e	rukwe’- baati	ru- táni -wewí	
1DU.EXCL		rukwé’- kwe’e	rukwé’-be- baati	ru- táni -webí	
2DU		riče’- kwe’e	riče’- baati	ri- táni -wewí	
3DU		tákwé’- kwe’e	tákwé’-re- baati	tátu- táni -wewí	
1PL.INCL		rukhwe’- kwe’e	rukhwe’- baati	ru- táni -həʔ	
1PL.EXCL		rukhwé’- kwe’e	rukhwé’-be- baati	ru- táni -həʔbí	
2PL		riče’- kwe’e	riče’- baati	ri- táni -həʔ	
3PL		táro’- kwe’e	táro’-re- baati	tátu- táni -həʔ	

Table 9: The morphological make-up of Class 2, part 3.

Nevertheless, the prefix *te-* is not found throughout the paradigm of Class 2 verbs, so we cannot ascribe it a derivational function just like we can't to the prefixes *t-* and *tú-* of transitive Classes II and III. For example, indicated in light shading in Table 9, there are many cells in the completive and the potential of Class 2 verbs where we observe default markers from Class 1.¹³⁶ Also, and most crucially, Class 2 has a set of idiosyncratic markers (indicated in boldface in Table 9) which cannot be ascribed to Class 1 by any synchronic morphophonological rule that makes sense.

This formal evidence points to an independent Class 2, which is based on Class 1. Further evidence that Class 2 constitutes a class of primary intransitives comes from the lexicon. Like Class III for transitive verbs, Class 2 is the productive intransitive class in Matlatzinca. This is shown in Table 10, where verbs are again divided between spontaneous (a) and established loanwords (b).

	Class 2	Class 1		Spanish PTPL	(verb)
a.	tute-fihadu	*tu-fihadu	‘look at/pay attention’	fijado	← FIJAR=SE
	tute-kortadu	*tu-kortadu	‘cut oneself’	cortado	← CORTAR=SE

¹³⁵ The rhotic variant *re-* of prefix *te-* can also be accounted for as an outcome of the same rule in (1) above.

¹³⁶ The evidence that these default markers are from Class 1 is that they do not show any traces of *te-*

tute-reidu	*tu-reidu	‘laugh’	reído	←	REÍR(=SE)
tute-bisitadu	*tu-bisitadu	‘pay visits’ (activity)	visitado	←	VISITAR
b. tute-pensaru	*tu-pensaru	‘be a believer’	pensado	←	PENSAR

Table 10: Intransitive loan verbs in Matlatzinca from Spanish.

As we show in the next section, Class 2 has a semantic core: It conveys middle voice semantics. In this respect, it is expected that all loanwords in Matlatzinca coming from middle verbs in Spanish (i.e., inflected with the reflexive pronominal paradigm) should also be found as members of Class 2. However, as Class 2 is the only option in the borrowing process (i.e., all intransitive loans whatever their semantics end up in this class), the fact that they are found in Class 2 is uninformative. What is more interesting is the fact that most loans from Spanish can be both inflected as Class III, when transitive, and as Class 2 when intransitive. Through those alternations in class, one obtains the typical semantic nuances associated with active *vs.* middle semantic alternations. This in turn suggests that the paradigms of Class III and Class 2 in Matlatzinca are working in the system as equivalents to the active *vs.* middle/passive paradigms of better-known languages like Latin or Greek. We will come back to this in Section 4.2.

Finally, the last intransitive class we treat as Class 3 is a small class containing only 18 verbs. The class is characterized by a nasal prefix that occurs at the stem boundary. We treat it as inflectional because, shown in Table 11, the nasal prefix does not occur in the past and in the singular forms of the imperfect,¹³⁷ and represents an absence that cannot be explained as motivated from a rule operating within the morphosyntax or the morphophonology.

		Class 1 ‘weep’	Class 3 ‘walk’
INCPL	1SG	tu- kwe’e	tu-n-tru
	2SG	’i- kwe’e	’i-n-tru
	3SG	ku- kwe’e	ku-n-tru
	1DU.INCL	kwen- kwe’e	kwe(n)-n-tru
	1DU.EXCL	kwén- kwe’e	kwé(n)-n-tru
	2DU	čen- kwe’e	če(n)-n-tru
	3DU	kwén- kwe’e	kwé(n)-n-tru
	1PL.INCL	khwen- kwe’e	khwe(n)-n-tru
	1PL.EXCL	khwén- kwe’e	khwé(n)-n-tru
	2PL	čhen- kwe’e	čhe(n)-n-tru
3PL	ron- kwe’e	ro(n)-n-tru	
CPL	1SG	ta’- kwe’e	ta’-n-tru
	2SG	’i’- kwe’e	’i’-n-tru
	3SG	ka’- kwe’e	ka’-n-tru
	1DU.INCL	kwe’- kwe’e	kwe’-n-tru
	1DU.EXCL	kwé’- kwe’e	kwé’-n-tru
	2DU	če’- kwe’e	če’-n-tru
	3DU	kwé’- kwe’e	kwé’-n-tru
	1PL.INCL	khwe’- kwe’e	khwe’-n-tru
	1PL.EXCL	khwé’- kwe’e	khwé’-n-tru
	2PL	čhe’- kwe’e	čhe’-n-tru
3PL	ro’- kwe’e	ro’-n-tru	
PST	1SG	ta’- kwe’e	ta’-tru
	2SG	’a’- kwe’e	’a’-tru

¹³⁷ When the base prefix already has a nasal coda, nasal gemination is avoided.

	3SG	ma'-	kwe'e	ma'-tru
	1DU.INCL	kwa'-	kwe'e	kwa'-tru
	1DU.EXCL	kwa'-	kwe'e	kwa'-tru
	2DU	ča'-	kwe'e	ča'-tru
	3DU	kwa'-	kwe'e	kwa'-tru
	1PL.INCL	khwa'-	kwe'e	khwa'-tru
	1PL.EXCL	khwa'-	kwe'e	khwa'-tru
	2PL	ča'-	kwe'e	ča'-tru
	3PL	roma'-	kwe'e	roma'-tru
IMPF	1SG	tu-	kwe'e	tu-tru
	2SG	'i-	kwe'e	'i-tru
	3SG	ku-	kwe'e	ku-tru
	1DU.INCL	kwena'-	kwe'e	kwena'-n-tru
	1DU.EXCL	kwéna'-	kwe'e	kwéna'-n-tru
	2DU	čena'-	kwe'e	čena'-n-tru
	3DU	kwéna'-	kwe'e	kwéna'-n-tru
	1PL.INCL	khwena'-	kwe'e	khwena'-n-tru
	1PL.EXCL	khwéna'-	kwe'e	khwéna'-n-tru
	2PL	čhena'-	kwe'e	čhena'-n-tru
	3PL	rona'-	kwe'e	rona'-n-tru

Table 11: The paradigm of a Class 3 verb compared with Class 1.

In this section, we have studied relevant aspects of the form of the inflectional classes of Matlatzinca. We have seen that transitive and intransitive verbs are given a different grammatical treatment and that there is one morphological default class for each type. However, the classes that are morphological defaults do not serve as lexical defaults and only the most marked classes are productive. The marked classes are based on the more basic ones in ways that suggest we are observing the remnants of derivational operations that were once productive in the history of the language. Such derivational processes are also behind the phenomenon of inflectional class interaction, which we deal with in the next section.

25. Inflectional class interaction

In the previous section we proposed that three out of the five inflectional classes of Matlatzinca verbs (i.e., Class II, Class III and Class 2) are formally based on the two remaining ones (i.e., Class I and Class 1). The derived classes have extra morphological elements (*t-*, *tú-* and *te-*), which add no additional meaning and carry no other morphological function but to serve as builders of inflectional class distinctions. Our view of Matlatzinca verbal classes has focused on form, but when one looks at how these classes are populated, we are confronted with a challenging situation.

The situation involves verbs that share a given stem and are found as members of different classes. When that happens, there is always an accompanying change in meaning and sometimes, but not always, a change in transitivity value. The examples in (2) illustrate the case at hand with the verbal stem *-bet'áni* 'work'.

- (2) a. *ku-khá-bet'áni* (Class 1)
 3SG.INCPL-ugly-work
 'He doesn't work well.' (Lit. 'He works uglily.')

- b. *kutú-bet'áni núniháni* (Class III)

3SG.INCPL-work field
 'He's tilling the field.' (Lit. 'He's working the field.')

In (2a), we have an instance of the intransitive verb *-bet'áni* 'work', which is a Class 1 verb. Here *-bet'áni* conveys the action of working as an atelic activity. In contrast, in (2b) we have an instance of the transitive verb *-bet'áni*, which is a Class III verb that has core semantics similar to the verb in (2a) but which presents the action of working as an action directed to another participant, i.e. the product of the work, which is treated as an object.

One way to regard this phenomenon in Matlatzinca is to think that both occurrences in (2) instantiate the same lexical entry. Under this view, one would need to explain how a verb can be a member of more than one inflectional class at the same time and why. That is, how the two different aspects (semantic, syntactic and inflectional) are stored in a unified lexical entry. An alternative way is to think that the two occurrences in (2) instantiate two different, although obviously related, verbs. Under such a view, one needs to spell out the ways they are related. The phenomenon at hand reminds us of so-called 'labile' or ambitransitive verbs in the typological literature. In other words, do the examples in English in (3) involve two verbs or just one?

- (3) a. I am shaking badly (intransitive)
 b. Don't shake the bottle! (transitive)

In Matlatzinca, the phenomenon involves not only changes in inflectional class but also changes in meaning which may be further reflected on changes in the verb's syntactic valence. One of the main disadvantages of taking an approach that considers there is only one verb is that a verbal entry in question would be expected to have at least two paradigms devoted to the expression of different meanings, but as we will see in this section there are many verbs that do not show this behavior. Under such an approach one is forced to say something about those instances where one of the paradigms is not filled. We believe that the existence of a different morphological inflection can be interpreted more coherently if the relation that holds between the classes in such pairings is seen as having a derivational nature, rather than being an inflectional property of lexemes.

The phenomenon shows restrictions too. For example, membership in a given inflectional class is fixed and dictated by the lexicon. In other words, for the activity 'reading' of the intransitive verb in (2a) one cannot chose to inflect the verb as a Class 2 verb. The same can be said of the transitive counterpart in (2b), which cannot be inflected as a Class I verb. This is shown by the ungrammaticality of the two examples in (4).

- (4) a. **kute-khá-bet'áni* (Class 2)
 3SG.INCPL-ugly-work
 Intended reading: 'He doesn't work well.'
- b. **kuh-bet'áni núniháni* (Class I)
 3SG.INCPL-work field
 Intended reading: 'He's tilling the field.'

Similarly, there are lexical restrictions regarding productivity: one cannot take any transitive action verb of Class III similar in semantics to *-bet'áni* 'work' in (2b), such as *-huhti* 'inflate', as in (5a), and obtain its semantic antipassive equivalent by means of an intransitive verb that inflects as a Class 1 verb, like (5b), which is ungrammatical.

- (5) a. *kutú-huhti* (Class III)
 3SG.INCPL-inflate[3OBJ]
 'It (the pump) inflates it (the tyre).'
- b. **ku-huhti* (Class 1)
 3SG.INCPL-inflate
 Intended reading: 'It (the pump) inflates.'

In other words, pairs such as the one in (2) are not the outcome of a productive derivational rule in Matlatzinca that involves valence changing mechanisms which, depending on directionality, could either produce antipassive or extended transitive verbs. Lexical and inflectional restrictions associated with such pairs suggest that the members that take part in them are learned independently as different lexical entries. Under such a view, verbs that show multiple class membership could be seen as forming lexical networks of cognate verbs that share a stem and have interrelated semantics. This is the approach we take here. We refer to such networks as 'pairings'. Such pairings normally involve two verbs, i.e. a proper pair, but sometimes they involve triplets.

However, the phenomenon is not straightforward. When Class 2 verbs are brought into the picture, it appears that speakers have much greater flexibility in producing new meanings, especially when the two productive classes are involved as for example in the pair in (6). This suggests that the rules linking two verbs in an inchoative/causative alternation (Haspelmath 1993) involving such classes is more productive.

- (6) a. *tutú-čibe* *nín-tešna* (Class III)
 1SG.INCPL_{TR}-get.dirty 3POSS.SG-clothes
 'I'm making his shirt dirty.'
- b. *tute-čibe* (Class 2)
 1SG.INCPL_{INTR}-get.dirty
 'I'm getting dirty.'

A Class 2 verb in such pairings often renders more than one meaning. For example, in (7a) the transitive verb *kánči* of Class I conveys the action meaning of 'hitting something or somebody', while its Class 2 counterpart conveys a situation typical of middle voice semantics: the verb may have a meaning involving a self-centered action (i.e. reflexive), like in (7b), where it has a meaning equivalent to 'flagellate', or it may have a reciprocal meaning, like in (7c).

- (7) a. *kuh-kánči* *Šuba* (Class I)
 3SG.INCPL-hit John
 'He's hitting John.'

- b. *kute-kánči be tetíwi* (Class 2)
 3SG.INCPL-hit P church
 ‘He’s flogging himself in the church.’
- c. *ronte-kánči* (Class 2)
 3PL.INCPL-hit
 ‘They are hitting each other.’

25.1. Interaction of classes in pairings

In contrast to verbs which belong to these pairings, there are other verbs that do not and are thus members of only one inflectional class. Table 12 gives an indication of the distribution of both types of verbs in total and per class in our sample of 614 verbs.

Interaction	I	%	II	%	III	%	1	%	2	%	3	%	Total	%
Unpaired	32	50%	2	10%	13	10%	105	63%	31	16%	18	100%	201	33%
Paired	32	50%	17	90%	133	90%	63	37%	168	84%	0	–	413	67%
Total	64	100%	19	100%	146	100%	168	100%	199	100%	18	100%	614	100%

Table 12: Inflectional class interaction ratio in a sample of 614 Matlatzinca verbs.

Table 12 reveals that two thirds of the verbs in our sample enter into some type of pairing network, but their distribution is different across the classes. The general preference is for morphologically derived classes to show a much greater propensity to have verbs involved in pairings. This is consistent with a historical scenario where such classes are seen as involved in derivational morphology (see 4.2 below). Intransitive Class 3 is the class with the lowest degree of interaction with other classes, followed by Class 1.

The sample of 614 verbs was selected by paying careful attention to all semantic possibilities. The relevant figures that emerge from pairings are given in Table 13, where we have arranged the different classes in columns. The 413 verbs entering into some type of class interaction display almost 190 pairings, which may involve just pairs of verbs or triplets of verbs. The rows in Table 13 elaborate on the different logical patterns that could have emerged involving all possibilities, but not all of them are manifested in the data:

- (a) interactions involving transitive verbs only;
- (b-d) interactions involving one transitive verb plus one or two intransitive verbs;
- (e) interactions involving two transitive verbs and one intransitive verb;
- (f) interactions involving intransitive verbs only.

We have stopped at patterns involving more than three members because such patterns are not attested in the data.

Interaction	Pairing pattern	N° of pairings	N° of verbs involved
a. Only tr.	I-II	Unattested	Unattested
	I-III	1	2

	II-III	Unattested	Unattested
	I-II-III	Unattested	Unattested
b. tr & intr.	I-1	1	2
	I-2	22	44
	I-1-2	7	21
c.	II-1	2	4
	II-2	Unattested	Unattested
	II-1-2	14	42
d.	III-1	17	34
	III-2	101	202
	III-1-2	12	36
e.	I-II-1	Unattested	Unattested
	I-II-2	Unattested	Unattested
	I-III-1	Unattested	Unattested
	I-III-2	1	3
	II-III-1	Unattested	Unattested
	II-III-2	1	3
f. Only intr.	1-2	10	20
	Total	189	413

Table 13: Types and frequency of pairings in inflectional class interactions

One may see in Table 13 that some pairing patterns are more populated than other; some simply are not attested. Also, with the exception of pattern III-1 in (d), attested in 17 pairings, most involve at least a verb of Class 2 in some way or another. This we take to be significant in telling us about the special inflectional status of Class 2 (see 4.2).

Pairing patterns in (a) of Table 13 involve transitive verbs only. The figures reveal that this is the least favored type of interaction in Matlatzinca. In our sample it is only attested in the pair in (8). The difference between the verbs is one of valence: the verb of Class I is a monotransitive and the verb of Class III a ditransitive (example 8b is an instance of an external possessor construction where the possessor has been promoted to object).¹³⁸

- (8) a. *kuh-péwi n-boyi* (Class I)
 3SG.INCPL-steal DET.SG-cow
 ‘He’s stealing the cow.’
- b. *kutú-pe¹³⁹-ki the-boyi* (Class III)
 3SG.INCPL-steal.AS-1OBJ 1POSS-cow
 ‘He’s stealing my cow.’ (Lit. ‘He’s stealing me my cow.’)

The patterns in (b-e) involve pairings of transitive verbs with intransitive ones. In all such pairings, the transitive verb encodes an agentive scenario with two

¹³⁸ This is the expected repartition because, except for a couple of verbs, ditransitive verbs in the language belong to Class III.

¹³⁹ The segment *-wi* in *péwi* is a stem formative that is dropped when the verb receives affixes.

asymmetrical participants treated as subject and object. The interest lies in the intransitive verbs.

When the pattern involves pairings with an intransitive verb of Class 1, the default case is for the verb of Class 1 to present the action of its transitive counterpart as an atelic, agentive activity, i.e. it serves as its semantic antipassive equivalent that defocuses the O argument of the transitive verb. Examples of this are given in Table 14.

		Transitive	Intransitive of Class 1
<i>-tani</i>	I	'buy sth.'	'go shopping'
<i>-šeepi</i>	I	'harvest sth.' (eg. corn)	'do the harvest'
<i>-he</i>	II	'see sth./sb.'	'see' (have sight)
<i>-'iwi</i>	II	'make go to sleep'	'sleep'
<i>-nhəti</i>	III	'kick sth./sb.'	'kick'
<i>-čana</i>	III	'shout at sth./sb.'	'scream'
<i>-soots'i</i>	III	'scold sb./tell sb. off'	'be prone to tell people off'

Table 14: The activity semantics of Class 1 verbs.

However, in some of the pairings the semantic relation that holds between the members is more of an inchoative/causative type, where the argument of the intransitive verb of Class 1 is the one undergoing the same state of affairs that the O of the transitive member undergoes.

		Transitive	Intransitive of Class 1
<i>-nuki</i>	III	'make sth. fall'	'fall'
<i>-nuwi</i>	III	'wake sb. up'	'wake up'
<i>-tahts'i</i>	III	'move sth. up'	'get up'
<i>-k'ayosa</i>	III	'make sb. lose weight'	'lose weight'

Table 15: The inchoative semantics of Class 1 verbs.

Only in the three cases in Table 16, the verb of Class 1 denotes a state.

		Transitive	Intransitive of Class 1
<i>-yə</i>	I	'have' (existential)	'exist'
<i>-čoši</i>	III	'tire sb. out'	'be tired'
<i>-čunta</i>	III	'wake sb. up'	'be awake'

Table 16: The stative semantics of Class 1 verbs.

The distribution of the different semantic readings in the pairs involving a Class 1 verb is given in Table 17.

Semantic relation	Pairs	Types
--------------------------	--------------	--------------

Action/antipassive	2 II-1
	9 III-1
	31 I/I/III-1-2
Inchoative/causative	8 III-1
Action/state	1 I-1
	2 III-1-2
Total	53

Table 17: Distribution of semantic types in Class 1 verbs.

As indicated in Table 16, a Class 1 verb has an antipassive reading by default. This is particularly the case when a Class 2 verb is also involved; in such pairs it is the Class 2 verb that conveys the inchoative meaning. For example, the verb *-'eya* means 'heal somebody' when transitive, 'be a healer' or 'heal as a profession' when it is a Class 1 verb, and 'get better' or 'be healthy' when in Class 2. This brings us to Class 2.

The great majority of pairings involve intransitive verbs of Class 2, but these pairings also display the greater heterogeneity of meanings, which are, nonetheless typical of systems with middle voice (Kemmer 1993, Palancar 2004, etc.). The class subsumes the five semantic types in (9), which are illustrated in turn in Table 18 with examples:

- (9)
- a. An atelic routine-like activity that is characteristic of Class 1 verbs.
 - b. An autonomously performed action. Here the action of the transitive verb is presented as a typical reflexive or self-centered situation that commonly involves human beings. In such a situation, the subject is viewed as performing the action independently from external causes while holding some control. Such actions can also involve motion and manner of motion.
 - c. A spontaneous action or state of affairs that mainly involves inanimate participants exerting no control.
 - d. A reciprocal action.
 - e. An impersonal/passive reading, which presents the result of the action after a given unspecified agent has acted upon it.

	Transitive	Intransitive of Class 2
a. <i>-nəči</i>	III 'give sb. lunch'	agentive activity 'provide lunches'
<i>-yoši</i>	III 'give sb. dinner'	'provide dinners'
b. <i>-baati</i>	III 'lose sth./sb.'	autonomously performed 'lose one's way'
<i>-surihmhi</i>	III 'suffocate sb.'	'drown'
	'kill by drowning'	
<i>-kubi</i>	III 'make sb. go away'	'run'
<i>-šaye</i>	III 'wet sth./sb.'	'get wet'
c. <i>-šuk'a</i>	III 'burn sth.'	spontaneous 'get burned (food)'
<i>-čunta</i>	III 'wake sb. up'	'wake up'
<i>-čo'či</i>	III 'break sth. down'	'get broken down'
<i>-'ebi</i>	III 'spill sth.'	'get spilled'
d. <i>-kanči</i>	I 'hit sth./sb.'	reciprocal 'hit each other'
<i>-panti</i>	I 'join sth./sb.'	'join'

	<i>-supi</i>	I	'kiss sb.'		'kiss each other'
	<i>-sanuta</i>	III	'greet sb.'		'greet each other'
e	<i>-kunči</i>	I	'crash sth.'	impersonal/passive	'get crashed'
	<i>-toči</i>	I	'break sth.'		'get broken'
	<i>-noni</i>	III	'plough sth.' (a field)		'be ploughed (field)'
	<i>-huti</i>	III	'bury sth./sb.'		'be buried'

Table 18. The middle voice semantic types of Class 2 verbs.

Often, the verb is polysemous. For example the verb *-tentani* means 'help somebody' or 'defend something/something' when transitive, and when a Class 2 verb it may mean 'defend/help oneself' as an autonomously performed action or 'defend/help each other' with a reciprocal meaning. Such behavior is also common in middle voice systems.

Patterns involving three members have two intransitive verbs, as seen in (a) in Table 19. The only exceptions are the two cases in (b) which involve two transitive verbs. As expected the transitive verb of Class III is ditransitive, like the one in the pairing in (8b) above.

a.	<i>-tani</i>	I	'buy sth.'	1	'go shopping'	2	'get sth. for oneself'
	<i>-'api</i>	I	'dig sth.'	1	'be digging'	2	'get dug'
	<i>-nabi</i>	III	'dance sth.' (a dance)	1	'dance'	2	'dance in a fête'
b.	<i>-čobi</i>	I	'fertilize sth.' (a field)	III	'spread sth.'	2	'get spilled', 'get fertilized'
	<i>-heči</i>	II	'loosen up', 'release'	III	'release sth. for sb.'	2	'get loosened'

Table 19: The semantics of typical triplets.

Finally pattern (f) in Table 13 involves pairings of intransitive verbs only. For the 10 existing cases in our sample, given in Table 20, the verb of Class 1 refers to an agentive activity, like the many cases in Table 14, while the meaning in Class 2 is again as heterogeneous as the instances in Table 18.

	Class 1	Class 2
<i>-hunči</i>	'shake' (e.g. a dog shaking off water)	spontaneous 'shudder'
<i>-košbi</i>	'slide'	'slip'
<i>-həʒts'i</i>	'fly'	'fly off (papers)'
<i>-tenu</i>	'gore'	reciprocal 'gore each other'
<i>-šeči</i>	'scratch (oneself)'	'scratch each other'
<i>-we'e</i>	'write'	'write to each other'
		impersonal/passive 'be written'
<i>-sopi</i>	'have a go at something, taste'	'be tried, tasted'
<i>-ta'či</i>	'milk'	'milk'
<i>-tuni</i>	'sow'	'be sowed'
<i>-'eti</i>	'hear'	'be heard'

Table 20: The semantics of pairings involving only intransitive verbs.

25.2. Discussion

In the previous section we have shown that many verbs in Matlatzinca enter into what we have treated as ‘pairings’ of related verbs that share both a lexical core meaning and the same stem. The verbs in such pairings inflect as members of different inflectional classes. In the default case, the pairings involve members with different transitivity values. The verbs involved also differ in aspects of their shared lexical meaning and the differences in meaning are associated with their change in class. Typically in these pairings, a verb of Class 1 presents the action of the transitive verb as an atelic activity and a verb of Class 2 has one or more of the senses typically associated with middle voice. When two verbs of the transitive classes pair (something that happens only rarely), the verb of Class III is ditransitive.

Transitive Classes I and II and intransitive Class 1 are closed classes. Membership in them is dictated by the lexicon. In the pairings, while most verbs of Class 1 have antipassive semantics, one must learn them independently as different verbs, i.e. one does not use a Class 1 verb spontaneously to produce an activity meaning. What happens with Class 2 verbs is more complex.

We propose that one can produce a middle verb in Matlatzinca from an active one in a productive way. The restrictions that apply to the meaning that such a verb will have are semantic. In other words, speakers can convert a given active verb into a Class 2 verb if the outcome makes sense both semantically and pragmatically. In this light, we view Class 2 as a productive grammatical mechanism to produce middle verbs in Matlatzinca. For this language, the derivation of a middle verb does not involve an overt morphological marking, but a change in inflectional class. The same phenomenon has been observed for Tilapa Otomi (see Palancar 2012), where all reciprocal verbs in the language are found in the cognate class of Class 2 (see below).

Under this view, Class 2 in Matlatzinca would be equivalent in functional and typological terms to the medio-passive paradigms of Indo-European languages like Latin and Greek, which are also observed beyond Indo-European in other languages such as Mari (Finno-Ugric) (Salo, 2006). Whether passivization is an inflectional or a derivational mechanism remains a matter of debate, as it does not fall within the prototypical properties of either domain spelled out in Dressler (1989). Stump (2006) argues for its inflectional character, while Andersen (1992), Sadler and Spencer (1998), and more recently Walther (2011a, 2011b, 2013) are all in favor of a derivational approach. Other authors, like Keenan (1985), prefer to split the functions and keep middle morphology within the inflectional domain, while passive goes to the derivational end. However, for the grammar of languages like Latin, Modern Greek, Sorani Kurdish and Matlatzinca there is no morphological distinction between the mechanisms used to produce middle meanings and those used for passive meanings. This makes Keenan's position impractical and untenable.

In the light of this discussion, there could be two main views on the status of the Class 2 paradigm in Matlatzinca. In one view, which we can call the ‘inflectional view’, transitive verbs in Matlatzinca would be seen as having two paradigms: one paradigm would be used when the verb is used in the active voice, and the other paradigm would be used when the verb is used in the medio-passive. Under this view, the selection of the active paradigm for a given transitive verb is stored in the lexicon, while the paradigm of Class 2 would be seen as the only existing middle/passive paradigm for a transitive verb. In an alternative view, which we can call the ‘derivational view’, the paradigm of Class 2 would be seen as functioning as an

independent inflectional class to which active verbs of the transitive classes are converted in order to express middle semantics. A similar analysis has been advanced for the ‘binyanim’ of Hebrew grammar by Aronoff (1994). The binyanim are traditionally regarded as a set of phonological templates that operate at an inflectional level and carry the function of modulating the core lexical meaning of a given verb to produce the typical changes observed in valence changing mechanisms, but not all verbs can be inflected in all binyanim. Aronoff’s analysis of such structures is in terms of interconnected inflectional classes.

The inflectional view has its advantages. It would account for the pairings of transitive verbs with Class 2 verbs as resulting from a transparent inflectional operation: active to medio-passive. Such a view has a certain elegance to it: the 31 unpaired verbs of Class 2 in Table 12, which do not appear in class interactions, would have to be seen as deponent verbs, that is, as intransitive verbs with no transitive counterpart. However, the inflectional analysis would also have to account for the reason why some transitive verbs do not have a corresponding pair in Class 2, and indirectly for the existence of other intransitive paradigms such as those of Class 1 and 3. If one takes a derivational view, acquiring middle semantics by a transitive verb would be seen as attained from class membership change; a derivational mechanism, i.e. a transitive active verb is converted into a Class 2 verb in order to obtain the middle semantics proper of the semantic core of Class 2. Such a view would allow for the existence of all unpaired verbs, regardless of class, but more crucially it would account for a lexical property of Class 2 which the inflectional approach would be unable to explain.

We have pointed out that Class 2 is also the only productive class for introducing new words into the intransitive lexicon of the language. While this makes Class 2 a lexical default, it also strongly suggests that middle voice, while being its core meaning, is not a necessary condition for membership to the class. In other words, avoiding a circular argument, we will not consider that all verbs of Class 2 are middle verbs by virtue of being members of Class 2. In contrast, we propose to treat as true middles only those verbs that occur in pairs with active verbs. In that sense, what would be regarded as ‘deponent’ verbs in languages like Latin or Greek, should just be treated as primary verbs in Matlatzinca.

As noted in Table 13 above, a typical pair involves a transitive verb of Class III and its corresponding middle counterpart of Class 2. This we take to be a by-product of these two classes being the only productive ones in the language. But these are also the classes that are morphologically more marked. As a consequence of this markedness, the morphology is uninformative about the directionality of the derivation, and the pairs operate as fully equipollent counterparts. This is most obvious when the verb is inflected in marked subparadigms, such as the incomplete, as shown in Table 21.

		‘make sb. go away’ Class III (tr)	‘run’ Class 2 (intr)
INCPL	1SG	tu-tú-kubi	tu-te-kubi
	2SG	’í-tú-kubi	’í-te-kubi
	3SG	ku-tú-kubi	ku-te-kubi
	1DU.INCL	kwen-tú-kubi	kwen-te-kubi
	1DU.EXCL	kwén-tú-kubi	kwén-te-kubi

2DU	čen-tú-kubi	čen-te-kubi
3DU	kwén-tú-kubi	kwén-te-kubi
1PL.INCL	khwen-tú-kubi	khwen-te-kubi
1PL.EXCL	khwén-tú-kubi	khwén-te-kubi
2PL	čhen-tú-kubi	čhen-te-kubi
3PL	ron-tú-kubi	ron-te-kubi

Table 21: A typical pairing in Matlatzinca.

We may ask us how this situation has come about. We recall that Tlahuica, the sister language of Matlatzinca, has a very similar inflectional system with similar inflectional class interactions (see Martínez 2012). In Tlahuica, there are only two transitive classes and they are more blended. In this regard, we propose that the prefix *tú-* characteristic of Class III in Matlatzinca functioned as a causative marker at some earlier stage in the common language. This analysis is based on evidence from its morphological distribution because Class III appears to be based on Class 1, but also on the semantic pairings that still involve Class 1 in an inchoative/causative alternation as shown in Table 15 above. Similarly, the best analysis for the prefix *te-* of Class 2 is to treat it as a middle marker which was operational at a given historical stage of the language.

But the development of *te-* posits a further problem from a typological point of view. Middles are commonly thought to evolve from reflexive markers, as is known to have happened in Romance and Scandinavian languages, and in an independent way. Reflexive markers are in turn often associated with transitive verbs. The story of *te-* would make more sense in the light of comparative evidence if Class 2 were morphologically based on transitive Class I. But as we have shown in Table 9, that is not the case for it is based on intransitive Class 1. This raises a syntactic problem: it is not immediately obvious how a middle verb can develop from a reflexive that is associated with an intransitive verb.¹⁴⁰ However, note that this is only a problem for a model that sees middle voice as a derived category from reflexivity. In other words, it is a problem as long as we believe *te-* was a reflexive in historical times, but this is not necessary. For example, Palancar (2004, 2006) has proposed that a reflexive origin of middle voice has to be discarded for the middle nasal prefix of Northern Otomi (unrelated to *te-*). There it is claimed that middle morphology in Northern Otomi serves as a grammatical mechanism to construe a given action by focusing on its internal dynamism. This internal construal may give rise to interpreting the action as atelic and imperfective, but also as not oriented towards an external participant. The latter interpretation gives rise to a construal where the action in question is portrayed as centered towards the self; which is a construal achieved via the syntax of reflexives in other languages. At present we do not know what triggered *te-* to become a middle marker in Matlatzinca. We know that the typological possibility exists that it may not have been a reflexive. What we also know is that it is a very old marker. Its old age can be deduced from comparative evidence with Otomi languages.

¹⁴⁰ Intransitive verbs that receive reflexives are known in languages with a fully developed middle voice category, like Spanish *caer=se* ‘fall’ which contrasts with intransitive *caer* ‘fall’ adding a nuance of unexpectedness (see Maldonado 1988), but these are commonly interpreted as late extensions from the more canonical construction where the reflexive is associated with a transitive base, ej. *romper=se* ‘break’ from transitive *romper* ‘break sth.’

As it happens, Class III and Class 2 merged into one class in Otomi-Mazahua; the other branch of Otomian. In the merger, *tú-* and *te-* became *ti-*. As a result of the merging, proto-Otomi-Mazahua obtained an inflectional class whose class membership was very heterogeneous from a lexical point of view. This old class has only survived in Otomi languages with conservative morphology. It corresponds to the so-called Class III of Tilapa Otomi (Palancar 2012), Acapulco Otomi (Hernández-Green, this volume) and Eastern Otomi (Palancar 2011). In Tilapa Otomi, for example, the class contains a miscellanea of verbs. Some are shown in Table 22 with forms of the first person incomplete realis. Here we have some ditransitive verbs like in (a), many transitive verbs (b), ambitransitives (c), and many intransitives, which may be marked with the middle prefix *n-* (d), or unmarked (e).

a.	trá-tí	'ũni	(dtr)	'give sth. to a third person'
b.	trá-tí	k'wáni	(tr)	'boil'
	trá-tí	'mbédi	(tr)	'lose'
	trá-tí	'ot'i	(tr)	'dry'
c.	trá-tí	k'ats'i	(tr/intr)	'gather cattle'
	trá-tí	münt's'i	(tr/intr)	'gather, harvest'
	trá-tí	mundo	(tr/intr)	'gather, heap up'
d.	trá-tí	m-pódi	(intr)	'get to know each other'
	trá-tí	n-hufi	(intr)	'hug each other'
	trá-tí	n-thëts'i	(intr)	'hug each other'
	trá-tí	m-phëts'i	(intr)	'swim'
	trá-tí	m-pá ^h ki	(intr)	'get something twisted'
	trá-tí	n-tx'öni	(intr)	'hide'
	trá-tí	n-khüni	(intr)	'get cloudy'
e.	trá-tí	ñú	(intr)	'see each other' (each other)
	trá-tí	ne ^h ki	(intr)	'be dawn'
	trá-tí	zópho	(intr)	'talk to each other'
	trá-tí	khú't'i	(intr)	'drip, leak'
	trá-tí	phët'i	(intr)	'tremble'
	trá-tí	k'ént'i	(intr)	'get an ankle twisted'

Table 22: Some Tilapa Otomi verbs of Class III.

It is generally agreed that Proto-Otomi split from Proto-Mazahua around the VI century AD, possibly a little later. Cazes (1973) proposes about the same time frame for the separation of Matlatzinca and Tlahuica from Proto-Atzinca. Although it is more difficult to posit a comparably accurate time frame for the separation of Proto-Atzinca and Proto-Otomi-Mazahua, the degree of lexical and grammatical divergences between the two groups makes it certain that the separation occurred around the 1st century AD, at the latest.

Verbal inflections in Otomi-Mazahua, cognates for the most part with Matlatzinca, like the sequence *trátí* in Table 21, are phonological words rather than canonical prefixes (see Palancar and Baerman 2014 for a discussion). This strongly suggests that *tú-* and *te-* in Matlatzinca (also in Tlahuica) are the likely and only remnants of the roots of two old auxiliary verbs in Otomian. The auxiliary-like origin of these constructions additionally accounts for the amalgam of prefixal material we still

observe in Atzincan paradigms. The constructions in which **tú-* and **te-* were involved date back to Proto-Otomian times. But the phonological merger in proto-Otomi-Mazahua of the two verbal classes in which **tú-* and **te-* participated in Otomian suggests that the markers had to have already ceased to function as morphosyntactically transparent constructions in Otomian, at least by the time of the separation of Proto-Otomi-Mazahua. The function was absorbed by the inflectional classes as morphological entities. And as this happened in both branches, the phenomenon we are observing is a very old one.

One more puzzle remains. In the Otomi languages that have retained it, Class III is also the productive class. For example, in Tilapa Otomi all loanwords from Spanish are integrated into the Otomi lexicon by inflecting them as members of Class III. Three examples are given in Table 23.

trá- tí	hórbá	‘annoy’	(tr)	(Sp.)	<i>jorobar</i>
trá- tí	gósá	‘enjoy’	(tr/intr)	(Sp.)	<i>gozar</i>
trá- tí	súphri	‘suffer’	(intr)	(Sp.)	<i>sufrir</i>

Table 23: Loans from Spanish in Tilapa Otomi as Class III verbs.

The cognate classes of Otomi Class III are Class III and Class 2 in Matlatzinca, which also happen to be the productive classes of the language. This we find remarkable. On the one hand, they are morphologically marked classes in their respective inflectional systems. They are cognate and very old. As we discard the possibility of borrowing, we are compelled to conclude that the productive character of these classes is not only old, but very old, most probably going back to the time of the diversification of Otomian before the Christian era (i.e. it appears to be an inherited feature). This makes us wonder about the functional load of the classes which are morphologically more basic, like Class I and Class 1 in Matlatzinca or Class I in Otomi, which have survived through the centuries although not adding much to the lexicon.

26. Conclusions

In this chapter we have shown that the inflectional classes of verbs in Matlatzinca are outstanding from a cross-linguistic point of view from both morphologically and lexically. From the study of a well-informed sample of 614 verbs, we have shown that verbs divide into six different classes in an even fashion regarding transitivity values, i.e. transitive verbs fall into three classes and only those three, and intransitive verbs fall into the three remaining ones, and only into those ones. Unexpectedly from a morphological point of view, the two classes which are the most marked are also the only ones which are productive. Besides this, we have also extensively shown that changes in the meaning and in the transitivity value of a given verb, which in other languages are commonly attained by valence changing mechanisms such as affixation or stem alternations, are carried out in Matlatzinca only by changes in the inflectional class membership of a verb. For example the lexical entry for the semantic notion of ‘waking somebody up’ is conveyed by a transitive verb that has the stem *-čunta*, but to convey the spontaneous event of ‘waking up’, the verb is now intransitive and it is thus inflected in a different class without any changes at stem level. Some such

membership changes are more productive than others. In general, they reveal that two of the three intransitive classes have a semantic core: one class holds typical activity verbs proper of semantic antipassives (buy something > go shopping) while the other class holds typical middle voice semantics (scratch something > scratch). We have considered that this state of affairs resulted from old markers that once had a clear derivational function but which became opaque in both form and function and were effaced by other inflectional material, causing their derivational function to become absorbed by the inflectional class distinctions. All this gave rise to the remarkable inflectional system that we can observe nowadays in Matlatzinca.

Acknowledgments

We want to thank Matthew Baerman and to Timothy Feist for their comments which helped us to improve the work. We are grateful for all comments received in the workshop on inflectional classes in the languages of Mesoamerica celebrated at CIESAS in July 2013, and to comments received at the VI Coloquio Sobre Lenguas Otomangues y Vecinas, celebrated in Oaxaca in April 2014.

Abbreviations

ADJ adjective; AS adjusted stem; CISLOC cislocative; CPL completive; DET determiner; DU dual; FUT future; IMPF imperfect; INCPPL incompletive; OBJ object; P preposition; PL plural; POSS possessive; PRG progressive; PTPL participle; PST past; sb. somebody; sth. something; TRANSLOC translocative.

References

- Andersen, Paul K. 1992. Remarks on passive morphology. In Bela Brogyanyi (ed.), *Prehistory, history and historiography of language, speech, and linguistic theory*, pp. 207-242. Amsterdam: John Benjamins.
- Bartholomew, Doris. 1966a. Matlatzinca phonology based on informant work done in May, 1966. MS.
- Bartholomew, Doris. 1966b. Observations about Matlatzinca grammar based on informant work done in May, 1966. MS.
- Carranza, Leonardo. 2013. Flexión verbal en matlatzinca. Marcación de Tiempo-Aspecto-Modo y persona-número. Unpublished MA thesis. Universidad Autónoma Metropolitana Unidad Iztapalapa
- Carranza, Leonardo. 2014. Flexión y clases verbales en matlatzinca. Paper read at VI Coloquio Sobre Lenguas Otomangues y Vecinas. Oaxaca: Mexico, (24-27 April).
- Cazes, Daniel. 1971. La lengua Maclasinca de Nsampaanchu, San Francisco Oxtotilpan. *Journal de la Société des Américanistes* 60: 191-232.
- Cazes, Daniel. 1973. *Le microphylum linguistique amerindien Hña-Méko (Otopame). Étude de comparaison, reconstruction et glottochronologie*. Unpublished PhD Dissertation at Paris La Sorbonne Nouvelle (Paris 3).
- Dressler, Wolfgang U. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42(1): 3-10.

- Escalante, Roberto and Marciano Hernández. 1999. Matlatzinca de San Francisco Oxtotilpan, Estado de México. [Archivo de lenguas indígenas de México 23]. Mexico: El Colegio de Mexico.
- Feist, Timothy, Enrique L. Palancar and Leonardo Carranza. 2015. Oto-Manguenan Inflectional Class Database: Matlatzinca. University of Surrey. <<http://dx.doi.org/10.15126/SMG.28/1.19>>.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie and Maria Polinsky (eds.), *Causatives and transitivity*, pp. 87-120. Amsterdam: John Benjamins.
- Keenan, Edward. 1985. Passive in the world's languages. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 1, pp. 243-281. Cambridge: Cambridge University Press.
- Maldonado, Ricardo. 1988. Energetic reflexives in Spanish. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 153-165
- Martínez, Aileen. 2012. Clases verbales, transitividad y valencia verbal en el p'jyekakj'ó, tlahuica de San Juan Atzingo. Unpublished PhD Thesis at El Colegio de Mexico.
- Palancar, Enrique L. 2004. Middle Voice in Otomi, *International Journal of American Linguistics* 70(1): 52-85.
- Palancar, Enrique L. 2006. Middle Voice and Intransitivity in Otomi. *Linguistics* 43(1): 613-643.
- Palancar, Enrique L. 2011. Revisiting the conjugation classes of Eastern Highlands Otomi. *Language Typology and Universals* 64(3): 213-236.
- Palancar, Enrique L. 2012. The conjugations classes of Tilapa Otomi: An approach from canonical typology. *Linguistics* 50(4): 783-832-
- Palancar, Enrique L. and Matthew Baerman. 2014. Perífrasis y clases flexivas en otomí. Paper read at Coloquio sobre lenguas otomanguanas y vecinas 6. Oaxaca: Mexico, (24-27 April).
- Pascacio, Etna. 2006. Fonología de la lengua matlatzinca. Unpublished MA thesis at the National School for Anthropology and History, Mexico.
- Pascacio, Etna. 2011. Pronombres personales independientes del proto-atzinca: reconstrucción morfológica. Unpublished MA thesis at the Autonomous National University of Mexico.
- Pérez, Nancy. 2007. Synchronic and diachronic Matlatzinkan phonology. Unpublished PhD thesis at the University of Michigan.
- Plank, Frans. 1994. Inflection and derivation. In R.E. Asher (ed.), *Encyclopedia of Language and Linguistics*, vol. 3, 1671-1678. Oxford: Pergamon.
- Salo, Merja. 2006. The derivational passive and reflexive in Mari grammars. *Slavica Helsingiensia* 27: 328-240.
- Sandler, Louisa and Andrew Spencer. 1998. Morphology and argument structure. In Andrew Spencer and Arnold M. Zwicky (ed.), *Handbook of Morphology*, pp. 206-236. Oxford: Blackwell
- Soustelle, Jacques. 1937. *La famille Otomi-Pame du Mexique Central*. Paris: Institut d'ethnologie.
- Stump, Gregory T. 2006. Heteroclis and paradigm linkage. *Language* 82(2): 279-322.
- Walther, Géraldine. 2011. *A derivational account for Sorani Kurdish passives*. Presentation at the 4th International Conference on Iranian Linguistics (ICIL4), 17-19 June 2011, Uppsala, Sweden.

Timothy Feist and Enrique L. Palancar

11 Tracing the emergence of inflectional tone in Cuicatec

1 Introduction

In this paper, we study a number of relevant aspects of the aspectual inflection of Cuicatec, an Oto-Manguanean language of Mexico. We show that the inflection of Cuicatec, which involves tone, displays a great degree of morphological complexity as a result of the following factors:

- (a) Verbs can be said to fall into a large number of inflectional classes according to the affix patterns they select for their aspectual inflection.
- (b) The same verbs fall into other inflectional classes according to the tone pattern they select for the same purpose.
- (c) Affixal morphology and tonal morphology operate as two independent systems of inflection, but they show areas of interconnection. Some of these areas are traces of lexical tones associated with prefixes which we claim were then reanalyzed as inflectional formatives.
- (d) Similarly, the inherent structural intricacies of the system are such that they give way to other alternative analyses, but whatever analysis is taken at the end, the complexity is still not greatly reduced.
- (e) Whichever analysis is adopted as a means of explaining the inflectional system, it invariably only accounts for a portion of the inflectional system, leaving behind a plethora of ‘loose ends’ and exceptions.

Cuicatec belongs to the Mixtecan branch of Oto-Manguanean, together with the Mixtec languages and Triqui. Our analysis of Cuicatec verb inflection is pioneering, as no previous analysis exists. The literature is rich on the Mixtec family and Triqui. But even in Mixtec the interest in the analysis has not been on inflection but on the phonology of tone, starting in Pike’s (1948) seminal book and leading to a vast literature; recent relevant work includes Daly & Hyman (2007), DiCano, Amith & García Castillo (2014), McKendry (2013) and DiCano (2012a, 2012b). Besides tone, studies with emphasis on the grammar focus on aspects of the morphosyntax (e.g. Bradley & Hollenbach 1988, 1990, 1991, 1992). The only exception is Macaulay’s (1996) grammar of Chalcatongo Mixtec, which includes a few sections on the morphological properties of inflection. See also Palancar, Amith & García Castillo (this volume) for a first analysis of the verb inflection of Yoloxóchitl Mixtec. Our analysis of Cuicatec inflection is based on secondary

data: we have compiled a sample of 2,480 inflected forms of 620 verbs from the excellent dictionary by Anderson and Roque (1983).

As a language, Cuicatec (c. 8,600 speakers) can be subdivided into two main varieties – Tepeuxila Cuicatec and Teutila Cuicatec – with a high degree of mutual intelligibility existing between them. The variety registered in Anderson and Roque (1983) is from Santa María Pápalo, a socio-linguistically significant dialect of Tepeuxila Cuicatec (Lewis et al. 2013). The language is spoken in the northwest of the state of Oaxaca in south-central Mexico, as shown in Map 1.



Map 1: Location of Santa María Pápalo, Oaxaca State, Mexico.

In Section 2 we give a brief summary of a number of pertinent points relating to the phonology of Cuicatec, which are necessary to understand the following sections, and in Section 3 we provide an overview of the verbal inflection system. Section 4 is concerned with the aspectual/modal prefixes, as it will become clear that these are relevant in understanding the emergence of inflectional tone.

Following the necessary background information provided in Sections 2–4, Section 5 introduces the notion of tone as an inflectional exponent. In Section 6, we investigate the tonal patterns observed in four different groups of verbs (based on the shape of their prefixes) and finally, in Section 7, we summarize our findings and propose a hypothesis to explain the synchronic complexity observed in Cuicatec verbal inflection.

2 Overview of Cuicatec phonology¹

2.1 Tone

All syllables in Cuicatec are tone bearing and predominantly CV in shape. Tones may be either level or contour. Level tones display a four-way contrast between high, high-mid, low-mid and low, marked with superscript numerals from 1 (high) to 4 (low).

A subset of the possible combinations of the four level tones gives rise to at least five contour tones. These contour tones are mostly observed on monosyllabic words, but are rare on native disyllabic words; this suggests that monosyllabic words are, in fact, bimoraic and can therefore either display a long level tone or a contour tone, but unfortunately vowel length is not marked in the orthography.

Table 1 presents nouns showing the main tone patterns on both monosyllables and disyllables. We have intentionally not shown verb forms in this table since this paper proposes that tone plays an important role in the inflection of Cuicatec verbs and the purpose of Table 1 is to illustrate the possible lexical tones.

As we will show in Section 6, in verbal inflected forms we often find some other tonal combinations which result from the addition of an inflectional tone 2 or 4 to a syllable that carries lexical tone. From such combinations we obtain additional tone patterns such as: 42, 23, 232 and 242. We are, however, unable to determine if such combinations should be analyzed as two adjacent level tones (our preference) or as specific contour tones found only in inflection.

Table 1: Cuicatec nouns displaying different lexical tone patterns

Monosyllables (level tones)					
1	high	H	l [?] in ¹	/l [?] i ¹ /	‘boy’
2	high-mid	M1	nd [?] a ²	/ ⁿ d [?] a ² /	‘worm’
3	low-mid	M2	g ^w an ³	/g ^w a ³ /	‘between’
4	low	L	ka ⁴	/ka ⁴ /	‘clothes’

¹ Orthography: when *n* is syllable-final, this indicates nasalisation of the preceding vowel (or the preceding two vowels where they are separated by *h* or *ʔ*). Also note that we have made a number of small adjustments to the orthography used in Anderson and Roque (1983) for the sake of clarity: (1) we represent laryngealization with an IPA glottal symbol, (2) we represent *c* with *k*, and (3) where *u* indicates labialization of the preceding consonant, we represent this with a superscript *w*.

Table 1: (continued)

Monosyllables (contour tones)					
14	descending	H>L	den ¹⁴	/dē ¹⁴ /	‘a little of’
24	descending	M1>L	yan ²⁴	/jā ²⁴ /	‘village’
32	raising	M2>M1	min ³²	/mī ³² /	‘there’
43	raising	L>M2	ku ⁴³	/ku ⁴³ /	‘metal’
243	descend./raising	M1>L>M2	yen ²⁴³	/jē ²⁴³ /	‘foam’
Disyllables					
1–2			d ² i ¹ v ² a ²	/d ² i ¹ .v ² a ² /	‘sparrow hawk’
1–4			l ² a ¹ nda ⁴	/l ² a ¹ . ⁿ da ⁴ /	‘flabby’
2–1			le ² ngue ¹	/le ² . ⁿ ge ¹ /	‘lame’
2–2			yu ² ni ²	/ju ² .ni ² /	‘path’
2–4			sa ² an ⁴	/sa ² . ² ā ⁴ /	‘man’
3–2			k ^w a ³ ku ²	/k ^w a ³ .ku ² /	‘straight’
3–3			ka ³ ka ³	/ka ³ .ka ³ /	‘paper’
4–3			d ² u ⁴ ni ³	/d ² u ⁴ .ni ³ /	‘brains’
4–4			nda ⁴ cha ⁴	/ ⁿ da ⁴ .tja ⁴ /	‘seven’
2–24			ga ² nar ²⁴	/ga ² .nar ²⁴ /	‘win’

2.2 Laryngealized vowels

Vowels in Cuicatec may be laryngealized. Anderson and Roque refer to two degrees of laryngealization – ‘lightly interrupted vowels’ and ‘abruptly interrupted vowels’. They treat ‘lightly interrupted vowels’ as a phonation effect on the vowel, which they represent with an apostrophe before the affected vowel and which we represent with a superscript glottal symbol (e.g. **ch[?]i⁴ka³**).

On the other hand, they treat ‘abruptly interrupted vowels’ as being consonantal in nature and occurring between two vowels and therefore **nde^{3?2}e³** ‘dance’ (POT) would be seen as being disyllabic **nde³⁻²e³**. Nevertheless, we have diverged from their analysis and instead treat **V²V** as a complex nucleus consisting of a single, bimoraic, glottalized vowel. We have done this for the following three reasons:

1. complex, glottalized, nuclei are a well-known phenomenon in Oto-Manguean languages (Suárez 1983: 33) and it is therefore less likely that these are glottal-initial syllables as the authors state;
2. in all examples in the data, the vowel preceding the ‘abrupt interruption’ is identical in quality to the vowel which follows. Were the glottal stop a true consonant and the onset of a new syllable, we might expect the quality of both vowels to differ in at least some examples;

3. the authors state that nasalization and the phonation associated with ‘lightly interrupted vowels’ always ‘spreads’ over an ‘abrupt interruption’ and affects both vowels. This could be seen as evidence that both vowels in fact belong to the same nucleus, and therefore the explanation of ‘spreading’ is superfluous.

3 Overview of Cuicatec verbal inflection

Verbs in Cuicatec inflect for aspect and mood and for person and number of the subject. Table 2 gives the Progressive forms of the verb ca^3ta^3 ‘sing’. As this partial paradigm shows, the pronominal system displays a number of grammatical distinctions, including clusivity (1PL), familiarity (2SG and 3rd person), gender (3rd person) and animacy (3rd person). Person and number are realised both syntactically, by means of independent pronouns, as well as morphologically by means of pronominal enclitics and inflection. Person/number inflection may involve tone (1SG), suffixes (2SG.FAM and 1PL.INCL) or the bare stem (all non-shaded forms in Table 2).

Unfortunately, despite the interesting complexity person and number marking bring to verbal inflection in Cuicatec, the dictionary does not provide sufficient data to enable us to reconstruct all the forms of a paradigm and, therefore, the remainder of this paper is concerned only with aspectual inflection.

Table 2: Progressive forms of the verb ca^3ta^3 ‘sing’

		Independent pro.	Verb	Dependent pro.
1 st person	1SG	$u^1=ne^1$	yi^3-ta^1	
	1PL.EXCL	$nd^2u^1s^1un^{43}=ne^1$	yi^3-ta^3	$=n^7un^{14}$
	1PL.INCL	$s^7un^1=ne^1$	yi^3-t-o^{31}	
2 nd person	2SG.FAMILIAR	$di^2=ne^1$	yi^3-t-e^2	
	2SG.FORMAL	$nd^2i^4=ne^1$	yi^3-ta^3	$=ni^4$
	2PL	$nd^2ist^7i^4=ne^1$	yi^3-ta^3	$=ni^4$
3 rd person	3.M.FAMILIAR	$sa^{27}an^4=ne^1$	yi^3-ta^3	$=san^2$
	3.F.FAMILIAR	$ta^{17}an^4=ne^1$	yi^3-ta^3	$=tan^1$
	3.FORMAL	$?i^4yan^4=ne^1$	yi^3-ta^3	$=yan^1$
	3.ANIMAL	$i^3ti^3=ne^1$	yi^3-ta^3	$=ti^4$
	3.INANIMATE	–	yi^3-ta^3	

All verbs have four aspectual forms – the Potential (POT), Progressive (PROG), Completive (CPL) and Perfect (PRF). A fifth form, the Habitual, is only given for a small subset of verbs and we will therefore leave it out of this paper. The forms used in the dictionary correspond to the bare stems. Examples of the four aspectual forms of five verbs are presented below in Table 3.

Table 3: Examples of the four aspectual forms of five verbs

	POT	PROG	CPL	PRF	
(1)	ka ² ta ⁴	yi ² ta ⁴	chi ² ta ⁴	ndi ² ta ⁴	‘get ill’
(2)	ku ⁴ chi ⁴ ka ³	yi ⁴ chi ⁴ ka ³	ka ⁴ chi ⁴ ka ³	nchi ⁴ ka ¹	‘stroll’
(3)	k ² a ² ka ²	?i ² ka ²	ch ^{?i} ka ³	nd ^{?i} ka ²	‘clap’
(4)	ndu ³ vi ²	yi ³ ndu ³ vi ²	ndu ⁴ vi ⁴	nnu ² vi ¹	‘praise’
(5)	n ^w a ²⁴	yi ² n ^w a ²⁴	n ^w a ²⁴	n ^w a ²⁴	‘recover’

In this paper we make reference to verbal stems and aspectual prefixes and, thus, it is important to define what we mean by a stem and a prefix. In our analysis, we treat a stem as any part of an inflected form which is segmentally invariant across all cells of the paradigm.

We use the word *invariant* somewhat loosely, however, as we allow for phonological changes to affect the stem. Table 4 shows the verb forms of Table 3, but this time indicating our segmentation of the prefix and the stem. In example (3), the initial vowel of the stem in the POT is /[?]a/, while in all other forms it is /[?]i/. We treat this as a phonological change triggered by the addition of the potential prefix *ka-*. Likewise, in example (4), the stems of all forms except the perfect begin with *nd-*, while the perfect begins with *n-*, which we treat as a sound change triggered by the addition of the perfect prefix *n-*.

Table 4: Repetition of Table 3, showing the segmentation of prefixes and stems

	POT	PROG	CPL	PRF	
(1)	ka ² -ta ⁴	yi ² -ta ⁴	chi ² -ta ⁴	ndi ² -ta ⁴	‘get ill’
(2)	ku ⁴ -chi ⁴ ka ³	yi ⁴ -chi ⁴ ka ³	ka ⁴ -chi ⁴ ka ³	n-chi ⁴ ka ¹	‘stroll’
(3)	k- [?] a ² ka ²	∅- [?] i ² ka ²	ch- [?] i ⁴ ka ³	nd- [?] i ² ka ²	‘clap’
(4)	∅-ndu ³ vi ²	yi ³ -ndu ³ vi ²	∅-ndu ⁴ vi ⁴	n-nu ² vi ¹	‘praise’
(5)	∅-n ^w a ²⁴	yi ² -n ^w a ²⁴	∅-n ^w a ²⁴	∅-n ^w a ²⁴	‘recover’

Also note, in this regard, that by *invariant* we are not considering tone, thus segmentally-invariant stems may nevertheless display tonal changes.

With these points in mind, any part of a verb form which does not appear in all forms is treated as external to the stem and thus as an exponent of inflection.

As the examples in Table 4 illustrate, inflectional prefixes may appear in all four forms (1, 2), in three forms (3), in two forms (4) or in only one form (5).

As Table 4 shows, a verb can either be composed of (i) an inflectional stem alone (e.g. POT form of (4)), (ii) an inflectional stem and a non-syllabic prefix, in which case the number of tones will remain unaltered (e.g. PRF form of (2)), or (iii) an inflectional stem and a syllabic prefix (e.g. PROG form of (5)), in which case the number of tones will be greater than that observed in any other form in that verb's paradigm which either does not have a prefix or whose prefix is non-syllabic.

4 Prefix series

In the previous section we showed that verbs select different affixes for building their four aspectual forms. Verbs can be arranged into inflectional classes according to the prefixes they select. We will refer to the set of prefixes selected by a verb as a 'prefix series'.

The data used for this paper consists of a sample of 534 verbs. The most common prefix series is represented by 181 of these verbs (~34%), and we treat these verbs as belonging to *prefix series 1*. Verbs belonging to prefix series 1 do not take a prefix in the POT or CPL, while they take the prefix YI- in the PROG and N- in the PRF, as represented in Table 5.

Table 5: Prefix series 1

	POT	PROG	CPL	PRF
Prefix series 1	∅	YI	∅	N

The prefixes here are given in uppercase letters to indicate that they are the underlying prefixes, since prefixes in Cuicatec may be subject to allomorphy. When this type of morphological variance can be accounted for by morphophonology, we treat the prefix series as a macro-series that may be made up of more than one micro-series. By way of example, prefix series 3 is presented in Table 6. Here, the underlying prefixes (KU, YI, CHI and NDI) are given in uppercase letters, with the morphophonologically-motivated micro-series listed below with an example verb.

Table 6: Examples of verbs belonging to the four micro-series of prefix series 3

	POT	PROG	CPL	PRF	
3	KU	YI	CHI	NDI	
a	ku ³ -ndu ² ku ⁴	yi ³ -ndu ² ku ⁴	chi ⁴ -ndu ² ku ⁴	ndi- ² ndu ² ku ⁴	‘accompany’
b	k ^w - ² a ⁴ nen ¹	∅- ² a ⁴ nen ¹	ch- ² a ⁴ nen ¹	nd- ² a ⁴ nen ¹	‘shave’
c	k- ² u ³ chi ³	∅- ² i ³ chi ³	ch- ² i ⁴ chi ³	nd- ² i ² chi ³	‘bury’
d	k ^w -a ³ ku ²	g-a ³ ku ²	ch-a ⁴ ku ⁴	nd-a ² ku ³	‘cry’

We will now take each micro-series in turn and explain why we take them to represent the same prefix series. Series 3a displays the archetypal forms of prefix series 3 in that all four forms display the fully realised prefixes.

Series 3b displays reduced, non-syllabic prefixes in all forms except the PROG. This phonological reduction can easily be accounted for when we note that the same pattern is repeated on all verbs where the stem begins with a laryngealized vowel. With regard to the absence of a prefix in the PROG form, again, this holds for all stems beginning with a laryngealized vowel and thus we are able to say that ∅- is an allomorph of yi- with which it is in complementary distribution.

Series 3c is identical in most respects to series 3b, except that the vowel of the prefix *ku-* replaces the initial vowel of the stem. While this behaviour is typical of verbs beginning with a laryngealised /²i/, there are exceptions to this, thus we are unable to treat this as an automatic phonological change. Nevertheless, it is clear that it is the phonological shape of the prefix that is affecting the stem and therefore it can still be considered a micro-series of prefix series 3.

Series 3d follows the same principles as Series 3b, except that verbs belonging to this micro-series have a vowel-initial stem, rather than beginning with a laryngealized vowel. The PROG prefix in vowel-initial verbs is realised as *g-*, which is in complementary distribution with yi- and ∅-.

Having explained the methodology behind our division of Cuicatec verbs into different prefix series, in Table 7 we present the 14 different inflectional classes we identify, according to the prefix series which verbs select. We indicate the number of verbs which belong to each series in the final column. Note that some series are instantiated by a single verb.

Table 7: Example verbs from the 14 inflectional prefix series

	POT	PROG	CPL	PRF	
1	∅	YI	∅	N	(181)
	na ⁴ nde ⁴	yi ⁴ -na ⁴ nde ⁴	na ⁴ nde ⁴	n-na ² nde ⁴	'converse'
2	∅	YI	∅	∅	(138)
	nti ³ vi ²	yi ³ -nti ³ vi ²	nti ⁴ vi ⁴	nti ²³ vi ²	'turn on'
3	KU	YI	CHI	NDI	(52)
a	ku ³ -ndu ² ku ⁴	yi ³ -ndu ² ku ⁴	chi ⁴ -ndu ² ku ⁴	ndi ² -ndu ² ku ⁴	'accompany'
b	k ^w -a ⁴ nen ¹	?a ⁴ nen ¹	ch- ² a ⁴ nen ¹	nd- ² a ⁴ nen ¹	'shave'
c	k- ² u ³ chi ³	?i ³ chi ³	ch- ² i ⁴ chi ³	nd- ² i ² chi ³	'bury'
d	k ^w -a ³ ku ²	g-a ³ ku ²	ch-a ⁴ ku ⁴	nd-a ² ku ³	'cry'
4	KU	YI	CHI	N	(37)
a	ku ³ -nda ³ va ³	yi ³ -nda ³ va ³	chi ⁴ -nda ³ va ³	n-na ²³ va ³	'fall over'
b	k ^w -e ³ nu ³	?e ³ nu ³	ch- ² e ⁴ nu ³	n- ² e ² nu ³	'occur'
c	k- ² u ² ne ²	?i ² ne ²	ch- ² i ⁴ ne ⁴	n- ² i ² ne ²	'grind'
d	k ^w -e ² nu ²	g-e ² nu ² (?)	ch-e ⁴ nu ⁴	n-e ² nu ²	'spy'
5	KU	YI	CHI	NI	(8)
	ku ² -ne ²	yi ² -ne ²	chi ⁴ -ne ⁴	ni ² -ne ²	'ripen'
6	KU	YI	CHI	NU	(8)
	ku ⁴ -nd ^{?i} 1 ⁴	yi ⁴ -nd ^{?i} 1 ⁴	chi ⁴ -nd ^{?i} 1 ⁴	nu ² -nd ^{?i} 1 ⁴	'cover'
7	KA	YI	CHI	NDI	(26)
a	ka ³ -k ² a ² ka ²	yi ³ -k ² a ² ka ²	chi ⁴ -k ² a ⁴ ka ³	ndi ² -k ² a ² ka ²	'drive'
b	k- ² a ³ ka ³	?i ³ ka ³	ch- ² i ⁴ ka ³	nd- ² i ² ka ³	'cut'
c	k-a ⁴ tu ⁴	g-a ³ tu ³	ch-a ⁴ tu ³	nd-a ³ tu ³	'fracture'
8	KA	YI	CHI	N	(25)
a	ka ⁴ -nd ^{?i} 4 ⁴ chi ⁴	yi ⁴ -nd ^{?i} 4 ⁴ chi ⁴	chi ⁴ -nd ^{?i} 4 ⁴ chi ⁴	n-n ^{?i} 2 ⁴ chi ¹	'see'
b	k- ² a ³ nga ¹⁴	?i ³ nga ¹⁴	ch- ² i ⁴ nga ¹⁴	n- ² i ² nga ¹⁴	'spill'
c	k-a ² nu ⁴	g-e ² nu ⁴ (?)	ch-e ² nu ⁴	n-e ² nu ⁴	'run'

2 In the original source, this form is given as *gue²nu²*, following the conventions of Spanish orthography, where *gu-* before /i/ or /e/ represents [g].

Table 7: (continued)

	POT	PROG	CPL	PRF	
9	KA	YI	CHI	NI	(5)
	ka ² -ŋd ² i ⁴ ya ⁴	yi ² -ŋd ² i ⁴ ya ⁴	chi ² -ŋd ² i ⁴ ya ⁴	ni ² -ŋd ² i ⁴ ya ⁴	‘frighten’
10	KA	YI	CHI	∅	(3)
	ka ⁴ -nne ⁴	yi ⁴ -nne ⁴	chi ⁴ -nne ⁴	nne ²⁴	‘speak’
11	KU	YI	KA	NDI	(7)
a	k ^w -a ¹ t ² un ⁴	g -a ¹ t ² un ⁴	k -a ¹ t ² un ⁴	nd -a ¹ t ² un ⁴	‘swallow’
b	k ^w - ² a ⁴	² a ⁴	k - ² a ⁴	nd - ² a ⁴	‘call’
12	KU	YI	KA	N	(19)
a	ku ⁴ -chi ⁴ ka ³	yi ⁴ -chi ⁴ ka ³	ka ⁴ -chi ⁴ ka ³	n -chi ⁴ ka ¹	‘stroll’
b	k ^w - ² an ²	² an ²	k - ² an ⁴	n - ² an ²	‘accept’
c	k ^w -an ⁴	g -an ⁴	k -an ⁴	n -an ⁴	‘scratch’
13	KA	YI	{POT}	N+{POT}	(15)
a	k -u ³ vi ³	g -u ³ vi ³	k -u ⁴ vi ³	n -ku ²³ vi ³	‘appear’
b	k - ² a ² n ^{nu} ⁴³	² a ² n ^{nu} ⁴³	k - ² a ² n ^{nu} ⁴³	n -k ² a ² n ^{nu} ⁴³	‘fold’
14	{CPL}	YI	CHI	N+{CPL}	(9)
	ch -o ² nd ^ā ¹⁴	g -o ² nd ^ā ¹⁴	ch -o ⁴ nd ^ā ¹⁴	n -cho ² nd ^ā ¹⁴	‘exceed’

It is worth highlighting here that the different prefix series mainly emerge from the combination of a few alternant markers as exponents of a given value. For example, the prefixes \emptyset -, *ka*-, and *ku*- in the POT, with their respective realizations, combine with other alternant markers realizing other values, such as \emptyset -, *chi*-, *ka*- and *ku*- in the CPL. There is also a degree of syncretism in prefixes series, such as \emptyset - for POT and CPL.

Note that, in prefix series 13 and 14, the notations {POT}, {CPL} and {PRF} indicate that the inflected forms given within the braces serve as the inflectional stem of the given form, i.e. as Priscianic stems. Thus, the stem of the verb ‘appear’ is vowel-initial *-uvi*, and the PROG therefore takes the allomorph *g*-, resulting in *guvi*. The POT form takes the reduced form of the prefix *ka*-, resulting in *k*-*uvi*. The form of the CPL, however, is (segmentally) identical to the POT, and the PRF form is, likewise, formed from the POT, but in the latter case adding the PRF prefix *n*-.

5 Introduction to inflectional tone – where should we look for it?

Having explained the basics of the segmental exponents of inflection, we are now in a position to turn to the role that tone plays in Cuicatec verbal inflection. While identifying inflectional prefixes is a straightforward process of finding those parts of verb forms that were not present in all paradigm forms, analysing the tonal patterns of Cuicatec verbs is far from simple, for a number of reasons, which we will now go through one by one.

Firstly, every syllable in a Cuicatec word carries a tone, or a series of tones. Given the huge variance in tonal patterns across Cuicatec verb forms, we believe it is highly unlikely that every tone plays a role in conveying inflectional information; therefore, it is necessary to distinguish between inherent lexical tones and those tones which are exponents of inflection.

In attempting to address this issue, let us consider where, if anywhere, inflectional tone might present itself on a verb form. Since Cuicatec verb stems are overwhelmingly bisyllabic there are three obvious positions to consider: (i) the inflectional prefix, (ii) the first syllable of the stem and (iii) the second syllable of the stem.

It is clearly plausible to think that aspectual prefixes might bear inflectional tone since they are, by their very nature, involved in the inflection of the verb. However, this raises the question of what happens when the prefix is either non-syllabic or absent. The first syllable of the stem is also a potential candidate as a host for inflectional tone, but then how should we treat forms that do have a tone-bearing prefix? Of the three positions in a verb form which we are considering as potential hosts for inflectional tone, we consider the second syllable of the stem to be the least likely. The reason for this is because the second syllable of the stem is associated with person marking, as we have already encountered in Section 2. In this paper we therefore consider both the tones of any prefix, together with the first tone of the stem.

Secondly, given the fact that aspectual prefixes may, on the one hand, be syllabic and thus tone-bearing, while on the other hand they may be non-syllabic or absent and thus unable to bear tone, it is important to establish which tone in a given verb form corresponds to the same tone in the corresponding form of another verb. To illustrate the importance of establishing the relevant tone, consider the examples presented below in Table 8 (prefixes are indicated in bold typeface).

Table 8: Illustration of the tone patterns which arise from both prefixes and stems

	POT	PROG	CPL	PRF	
(1)	ku ³ di ²	yi ³ di ²	chi ⁴ di ³	ndi ² di ³	'fondle'
PREFIX	3	3	4	2	
STEM	2	2	3	3	
(2)	ku ⁴ di ³ nu ³	yi ⁴ di ³ nu ³	chi ⁴ di ³ nu ³	ndi ²³ nu ³	'be encouraged'
PREFIX	4	4	4	–	
STEM	3	3	3	23	
(3)	nte ²	yi ³ nte ²	nte ⁴	nte ²	'restore'
PREFIX	–	3	–	–	
STEM	2	2	4	2	
(4)	k- ² a ² ka ²	?i ² ka ²	ch- ^{?i} 4ka ³	nd- ^{?i} 2ka ²	'clap'
PREFIX	–	–	–	–	
STEM	2	2	4	2	

In example (1), we can either consider the tones of the prefix, giving a pattern 3~3~4~2, or those of the stem, giving us a tone pattern across the four forms of 2~2~3~3. In example (2), although all four forms have a prefix, three of these are syllabic and carry a tone, whereas the fourth is non-syllabic, so we can either consider the tones of the first syllable of each stem, giving 3~3~3~23, or alternatively simply consider the first tone of each inflected form, which corresponds to the prefix in three of these forms and to the first syllable of the stem in the PRF form, which would give the pattern 4~4~4~23. Example (3) shows a similar issue, but here only the PROG form has a prefix, and the two alternative patterns are 2~2~4~2 or 2~3~4~2. In example (4), the prefix is non-syllabic in all forms in which it is present leaving only the stem to carry tone and giving a pattern of 2~2~4~2.

In order to address the second issue, where there is no prefix present, or the prefix is non-syllabic, we represent this with '0', allowing us to align the prefix "slot" and the first syllable of the stem "slot" of all verbs with each other. This is illustrated in Table 9.

Table 9: Illustration of how tone patterns are encoded

	POT	PROG	CPL	PRF	
Verb form	ku ³ di ²	yi ³ di ²	chi ⁴ di ³	ndi ² di ³	‘fondle’
Prefix (3a)	ku-	yi-	chi-	ndi-	
Tone pattern	3~2	3~2	4~3	2~3	
Verb form	ku ⁴ di ³ nu ³	yi ⁴ di ³ nu ³	chi ⁴ di ³ nu ³	ndi ²³ nu ³	‘be encouraged’
Prefix (4a)	ku-	yi-	chi-	n-	
Tone pattern	4~3	4~3	4~3	0~23	
Verb form	n ^{te} ²	yi ³ n ^{te} ²	n ^{te} ⁴	n ^{te} ²	‘restore’
Prefix (2)	∅-	yi-	∅-	∅-	
Tone pattern	0~2	3~2	0~4	0~2	
Verb form	k- ² a ² ca ²	ʔi ² ka ²	ch- ² i ⁴ ka ³	nd- ² i ² ka ²	‘clap’
Prefix (7b)	k-	∅-	ch-	nd-	
Tone pattern	0~2	0~2	0~4	0~2	

Thirdly, it is possible that there are word-internal prosodic changes that render otherwise transparent patterns opaque. Tones may spread from one syllable to another within the same word, in much the same way as vowel-harmony, or a given word may block the application of an inflectional tone. Unfortunately, it is not possible to answer all these questions with the limited data we have access to, but it is important to bear these possibilities in mind while analyzing tonal inflection.

6 Tonal patterns

As no previous analysis of Cuicatec verbal inflection exists, it is unclear if inflectional prefixes and tone interact with each other. We thus look at the behavior of tone within the context of the aspectual prefixes. In doing so, we attempt to ascertain, firstly, if tone is conveying any inflectional information and, secondly, if there is any correlation between inflectional prefixes and tonal patterns.

The second issue raised in Section 5 concerned the fact that aspectual prefixes may be realized in one of three phonological ‘shapes’ – i.e. as syllabic, non-syllabic or zero prefixes. We also showed in Section 4 that a single (macro) prefix series can consist of micro-series which exhibit different prefix shapes, depen-

dent on the phonological shape of the verb stem. Thus, in order to analyze the behavior of tones within the context of the aspectual prefixes, we begin by assigning each micro-series to one of four groups based on the phonological shapes of their prefixes. The characteristics of the four groups are outlined below in Table 10. Tables showing the prefix series which belong in each group will also be given in the relevant sections.

Table 10: Prefix shapes of verbs belonging to Groups 1–4

	POT	PROG	CPL	PRF	#	%
Group 1	–	CV-	–	C- / –	319	59.7
Group 2	C-	C-	C-	C-	143	26.8
Group 3	CV-	CV-	CV-	CV-	42	7.9
Group 4	CV-	CV-	CV-	C-	30	5.6
					534	100.0

We take each group of verbs in turn, starting with Group 1. The reason for choosing to begin with verbs belonging to Group 1, as opposed to Group 3, for example, which might appear a more obvious starting point given the presence of syllabic prefixes in all forms, is due to the fact that Group 1 contains the largest number of verbs (319 or ~60%) and therefore any patterns we observe in this set of data are more likely to be statistically significant.

Groups 2–4 are instantiated by significantly fewer verbs and these are thus dealt with following our analysis of Group 1 verbs. Group 2 consists of verbs with a non-syllabic prefix in all forms (including verb forms with the \emptyset - allomorph in the PROG); Group 3 consists of verbs with a syllabic prefix in all forms; and Group 4 consists of verbs with a syllabic prefix in the POT, PROG and CPL, and a non-syllabic prefix in the PRF.

6.1 Tonal patterns of Group 1

As explained in the preceding section, Group 1 verbs display the prefix shapes presented in Table 11. Only two micro-prefix series (Series 1 and 2) follow this pattern, although they account for a large proportion of Cuicatec verbs (319 verbs or ~60%). In this section, and those which follow, a schematic representation of the micro-prefix series which belong to each group will be presented (example verbs for each micro-series are presented in Section 4).

Table 11: Schematic representation of micro-prefix series belonging to Group 1

	POT	PROG	CPL	PRF	#
Group 1	–	CV-	–	C- / –	319
Series 1	∅–	yi–	∅–	n–	
Series 2	∅–	yi–	∅–	∅–	

The 319 verbs which belong to Group 1 display a large number of tonal patterns, which are presented in Table 12. These are arranged according to a number of observed patterns. Since only the PROG form has a prefix, in Group 1 we compare the tone of the first syllable of the stem across all four forms. In the following paragraphs we outline the patterns which we observe.

The first pattern which emerges from the data is that in all but fourteen verbs, out of a possible 319, the tone of the first syllable of the POT stem is identical to the corresponding tone of the PROG. Looking across all four paradigm forms we can also see that a small number of verbs are invariable with regards to the first tone of the stem (labeled as ‘Invariable’ in Table 12), but since this only applies to 29 verbs it is by no means the norm.

Turning our attention to the CPL, we see that 249 forms out of a possible 319 bear either tone 4, or begin with a tone 4. In the PRF, 299 forms either bear tone 2 or begin with tone 2. Verbs which follow both these patterns are listed in Table 12 as belonging to one of the A sets.

Looking more closely at those forms that *begin* with a tone 4 or tone 2, we can see that in most cases the following tone(s) mirror(s) the tones seen in the POT/PROG. From these observations we deduce that the CPL is primarily associated with tone 4 and the PRF is primarily associated with tone 2. We also deduce that this tone can either replace the tone of the POT/PROG (which we refer to as ‘replacive tone’, e.g. 0~2 in POT → 0~4 in CPL) or be added to the tone of the POT/PROG (which we refer to as ADDITIVE TONE, e.g. 0~3 in POT → 0~43 in CPL)³.

³ In those cases where tone 4 (CPL) or tone 2 (PRF) is not followed by an exact copy of the POT/PROG tone, a simple explanation for why this does not happen is usually easy to find. For example, when a tone 4 appears before the POT/PROG tone 43, this is not represented as 443, since the internal moraic structure of syllables is not reflected in tone marking in Cuicatec, thus a bimoraic syllable bearing tone 4, although phonetically it might be better represented with 44, is instead represented only with 4. Therefore, a tone 4 associated with the CPL, when added to tone 43 of the POT/PROG, would be represented as 43. We are, however, unable to be completely certain that verb entries in the dictionary displaying tone 43 in the POT/PROG and 43 in the CPL are, in fact, realized phonetically as 443 in the CPL, but we deduce this from the following two observations. Firstly, the fact that tone 4 and tone 2 are so common in the CPL and PRF forms, respectively,

Also worth noting here is that a single verb may exhibit both replacive and additive tone in its paradigm. In Table 12, verbs which only display replacive tone are listed as set A1, verbs which display both replacive and additive tone are set A2 and those which only display additive tone are listed as set A3.

The fact that the tones of the POT and PROG are the same in almost all verbs, coupled with the fact that these tones are sometimes retained in the CPL and PRF (in cases of additive tone) allows us to further deduce that the POT/PROG tone is an inherent lexical tone of the verb.

The final observations to be made about the data in Table 12 is that 35 verbs (or just over ten percent of Group 1 verbs) adhere to the observations above, except that the tone of the CPL is 3; these are listed in the rows marked B1 and B2 (B1 showing only those verbs displaying replacive tone, B2 showing those verbs which involve additive tone). In a similar fashion, another small subset of verbs display tone 1 in the PRF, but adhere to the above observations in all other respects; these are labeled as set C. Finally, 14 verbs are irregular with respect to at least one of the above observations and are listed as ‘Other’.

Table 12: The tonal patterns of verbs belonging to Group 1

		POT	PROG	CPL	PRF	Count
A	A1	0~2	3~2	0~4	0~2	56
		0~4	4~4	0~4	0~2	54
		0~3	3~3	0~4	0~2	18
		0~2	2~2	0~4	0~2	2
	A2	0~4	4~4	0~4	0~24	52
		0~3	3~3	0~4	0~23	25
		0~32	3~32	0~4	0~232	1
		0~43	4~43	0~4	0~243	1
	A3	0~43	4~43	0~43	0~243	10
		0~3	3~3	0~43	0~23	9
		0~32	3~32	0~43	0~232	2
		0~42	4~42	0~42	0~242	1

leads us to believe that despite being rendered opaque in some cases, there is nevertheless an additional tone 4 or 2 present. Secondly, in those verbs where the addition of a tone 4 in the CPL is only assumed, rather than given, the corresponding PRF forms do display the expected additional tone 2 (see A3 in Table 12).

Table 12: (continued)

	POT	PROG	CPL	PRF	Count
B B1	0~3	3~3	0~3	0~2	5
	0~2	3~2	0~3	0~2	4
B2	0~3	3~3	0~3	0~23	22
	0~3	4~3	0~3	0~23	4
C	0~3	3~3	0~4	0~1	5
	0~32	3~32	0~4	0~1	3
	0~2	3~2	0~4	0~1	2
Invariable	0~2	2~2	0~2	0~2	8
	0~24	2~24	0~24	0~24	8
	0~1	3~1	0~1	0~1	7
	0~2	3~2	0~2	0~2	3
	0~3	4~3	0~3	0~3	1
	0~14	3~14	0~14	0~14	1
	0~24	3~24	0~24	0~24	1
Other	0~2	3~2	0~43	0~2	2
	0~2	3~2	0~4	0~24	1
	0~2	4~3	0~3	0~23	1
	0~2	3~3	0~2	0~2	1
	0~3	3~3	0~4	0~24	1
	0~3	4~3	0~3	0~24	1
	0~23	3~3	0~3	0~23	1
	0~24	3~34	0~34	0~24	1
	0~3	3~3	0~43	0~12	1
	0~2	3~3	0~4	0~2	1
	0~3	3~2	0~4	0~2	1
	0~4	3~3	0~4	0~2	1
	0~2	3~3	0~3	0~2	1
Total					319

Table 12 shows all surface tone patterns of Group 1 verbs. However, taking into account lexical tone, we would no longer want to consider a verb with lexical tone 2 any differently to a verb with a lexical tone 3, for example, if they pattern in the same way with regards their inflection. We are therefore able to simplify Table 12 in Table 13 (disregarding the 14 verbs listed under ‘Other’).

Table 13: The tonal patterns of verbs belonging to Group 1

		POT	PROG	CPL	PRF	Count
A	A1	0~lex	3~lex	0~4	0~2	74
		0~lex	4~lex	0~4	0~2	54
		0~lex	2~lex	0~4	0~2	2
	A2	0~lex	4~lex	0~4	0~2/lex	53
		0~lex	3~lex	0~4	0~2/lex	26
	A3	0~lex	4~lex	0~4/lex	0~2/lex	11
0~lex		3~lex	0~4/lex	0~2/lex	11	
B	B1	0~lex	3~lex	0~3	0~2	9
	B2	0~lex	3~lex	0~3	0~2/lex	22
		0~lex	4~lex	0~3	0~2/lex	4
C		0~lex	3~lex	0~4	0~1	10
Invariable		0~lex	2~lex	0~lex	0~lex	16
		0~lex	3~lex	0~lex	0~lex	12
		0~lex	4~lex	0~lex	0~lex	1
Total						305

By disregarding the tone of the prefix in the PROG this table can be further simplified to provide the summary presented in Table 14. The motivation for a verb displaying either replacive or additive tone requires further investigation and, therefore, both these sub-types have been conflated in Table 14.

Table 14: Summary of patterns observed in Table 13

	POT	PROG	CPL	PRF	Count	Percent
A	LEX	LEX	4	2	231	72.4 %
B	LEX	LEX	3	2	35	11.0 %
C	LEX	LEX	4	1	10	3.1 %
Invariable	LEX	LEX	LEX	LEX	29	9.1 %
Other	X	X	X	X	14	4.4 %
Total					319	100 %

Table 14 shows that the general observations made about the tonal patterns in Group 1 hold true for 72% of verbs (Pattern A). The patterns which differ only in the marking of either the CPL or PRF (Patterns B and C) hold true for a further 14 % of verbs, invariable verbs account for around 9 % of verbs, leaving less than 5 % of verbs which appear to be irregular.

For clarity, we end Section 6.1 with a list of the observations made with respect to Group 1 verbs:

- 1) Tone 4 is an exponent of inflection in the CPL.
- 2) Tone 2 is an exponent of inflection in the PRF.
- 3) The tones in the POT and PROG forms are lexical tones.
- 4) Inflectional tones can either replace, or be added to, lexical tones.

We will now turn our attention to seeing if the same generalizations can be made about verbs which belong to Groups 2–4.

6.2 Tonal patterns of Group 2

In Group 2 we consider verbs that have a non-syllabic prefix in all forms (although in some cases the prefix may be realised as a zero allomorph). There are 143 verbs which follow this structural pattern, belonging to the 17 micro-prefix series presented in Table 15.

Table 15: Schematic representation of micro-prefix series belonging to Group 2

	POT	PROG	CPL	PRF	#
Group 2	C-	C-	C-	C-	143
Series 3b	k ^w -	∅-	ch-	nd-	
Series 3c	k- (+u)	∅-	ch-	nd-	
Series 3d	k ^w -	g-	ch-	nd-	
Series 4b	k ^w -	∅-	ch-	n-	
Series 4c	k- (+u)	∅-	ch-	n-	
Series 4d	k ^w -	g-	ch-	n-	
Series 7b	k- (+a)	∅-	ch-	nd-	
Series 7c	k-	g-	ch-	nd-	
Series 8b	k- (+a)	∅-	ch-	n-	
Series 8c	k- (+a)	g-	ch-	n-	
Series 11a	k ^w -	g-	k-	nd-	
Series 11b	k ^w -	∅-	k-	nd-	
Series 12b	k ^w -	∅-	k-	n-	
Series 12c	k ^w -	g-	k-	n-	
Series 13a	k-	g-	k-	n+k-	
Series 13b	k-	∅-	k-	n+k-	
Series 14	ch-	g-	ch-	n+ch-	

Table 16: The tonal patterns of verbs with a non-syllabic prefix in all forms

		POT	PROG	CPL	PRF	Count
A	A1	0~lex	0~lex	0~4	0~2	42
	A2	0~lex	0~lex	0~4	0~2/lex	9
	A3	0~lex	0~lex	0~4/lex	0~2/lex	3
	A4	0~lex	0~lex	0~4/lex	0~2	4
B	B1	0~lex	0~lex	0~3	0~2	15
	B2	0~lex	0~lex	0~3	0~2/lex	3
Invariable		0~4	0~4	0~4	0~4	57
Other		0~3	0~3	0~4	0~3	2
		0~3	0~3	0~1	0~2	1
		0~2	0~2	0~3	0~23	1
		0~34	0~34	0~34	0~24	1
		0~2	0~3	0~4	0~2	1
		0~3	0~1	0~4	0~2	1
		0~4	0~3	0~4	0~3	1
		0~4	0~3	0~43	0~2	2
Total					143	

Table 16 presents the tone patterns found, organized and labeled in a similar way as in the preceding section. That is, A1 verbs have a replacive tone 4 and 2 in the CPL and PRF, respectively; A2 verbs differ from A1 in that the tone is additive in the PRF; and A3 verbs show additive tone in both the CPL and PRF, B1 and B2 verbs are divided in the same manner except that the tone in the CPL is 3. Again, we have represented those tones we believe to be lexical as LEX, allowing us to reduce the number of apparent patterns. In this table it was necessary to add a new pattern, not seen in Group 1 verbs, namely Pattern A4. These verbs display an additive tone in the CPL, but a replacive tone in the PRF. A summary of the patterns in Table 16 is presented below in Table 17.

In general, Group 2 verbs do not adhere quite as well to our observations as we might have predicted, with only 40 % of verbs following Pattern A. However, the most striking difference observed here is the significant number of verbs which are invariable with respect to tone alternations, amounting to almost 40 % of this sample.

Table 17: Summary of patterns observed in Table 16

	POT	PROG	CPL	PRF	Count	Percent
A	LEX	LEX	4	2	58	40.5 %
B	LEX	LEX	3	2	18	12.6 %
C	LEX	LEX	4	1	0	0.0 %
Invariable	LEX	LEX	LEX	LEX	57	39.9 %
Other	X	X	X	X	10	7.0 %
Total					143	100 %

A simple explanation for this large number of invariable verbs, which would not be a valid argument for Group 1 verbs, is the fact that all paradigm forms of Group 2 verbs require an aspectual prefix. Since there is already a four-way distinction at the segmental level, we hypothesize that the large number of invariable verbs with respect to tone is due to tone no longer being required to play a role in differentiating between different inflectional forms.

Thus, by taking Pattern A and invariable verbs together we are able to account for 80 % of our sample of Group 2 verbs. Let us now consider Groups 3 and 4.

6.3 Tonal patterns of Group 3

In Group 3 we consider verbs that have a syllabic prefix in all forms. There are only 42 verbs which follow this structural pattern, belonging to the five prefix series presented in Table 18.

Table 18: Schematic representation of micro-prefix series belonging to Group 3

	POT	PROG	CPL	PRF	#
Group 3	CV-	CV-	CV-	CV-	42
Series 3a	ku-	yi-	chi-	ndi-	
Series 5a	ku-	yi-	chi-	ni-	
Series 6	ku-	yi-	chi-	nu-	
Series 7a	ka-	yi-	chi-	ndi-	
Series 9	ka-	yi-	chi-	ni-	

The first thing we notice when looking at the tone patterns of these 42 verbs (Table 19) is that, in all but six of them, the tone of the first syllable of the stem is invariable across all four forms, which provides even stronger evidence than before for this being the lexical tone of the verb. Nevertheless, tone 4 and 2 are still associated with the CPL and PRF, respectively, but these tones now appear on the prefixes. Note the adjusted notation used in presenting the patterns below in Table 19, where tone 4 (CPL) and tone 2 (PRF) appear in the prefix slot.

Table 19: The tonal patterns of verbs with a syllabic prefix in all forms

	POT	PROG	CPL	PRF	Count
A	3~lex	3~lex	4~lex	2~lex	11
	4~lex	4~lex	4~lex	2~lex	8
B	3~lex	3~lex	3~lex	2~lex	3
Invariable	2~lex	2~lex	2~lex	2~lex	3
	4~lex	4~lex	4~lex	4~lex	2
Other	2~2	2~2	4~4	2~2	5
	3~2	3~2	4~4	2~2	2
	2~4	3~4	2~4	2~4	1
	3~2	3~2	4~3	2~3	1
	3~2	3~2	4~43	2~2	1
	3~2	3~2	4~43	3~2	1
	4~2	4~2	4~2	4~4	1
	4~4	3~3	4~3	2~2	3
Total					42

Previously, we claimed that tone 4 and tone 2 are inflectionally associated with the CPL and PRF, respectively. Having seen that these tones are associated with prefixes in those forms which have syllabic prefixes, we further claim that they are *lexically* associated with the respective aspectual prefixes. When the prefix is syllabic, the tone would thus naturally appear on the prefix. However, when the prefix is reduced to a non-syllabic form, or not present, the *lexical* tone of the underlying or historic prefix shifts over to the first syllable of the stem.

This further claim is consistent with other behavior we have seen in Cuicatec verbal inflection and, indeed, provides a neat explanation for at least two phenomena. Firstly, the presence of additive tone and replacive tone could be indicative of a system in flux, whereby additive tone is the first step in a process of prefix loss, leading eventually to replacive tone.

Secondly, there are no instances of additive tone seen on the prefixes of the verbs represented in Table 19, suggesting that the tone of the prefix is its inherent lexical tone and additive tone only occurs when the stem hosts the residual tone of a reduced or lost aspectual prefix.

Taking these new observations into account, we now present the distribution of tone patterns of Group 3 verbs in Table 20. Note the amended notation under the CPL and PRF columns for patterns A, B and C to account for the fact the stem tone is a lexical tone in these verbs.

Table 20: Summary of patterns observed in Table 19

	POT	PROG	CPL	PRF	Count	Percent
A	LEX	LEX	4-LEX	2-LEX	19	45.3 %
B	LEX	LEX	3-LEX	2-LEX	3	7.1 %
C	LEX	LEX	4-LEX	1-LEX	0	0.0 %
Invariable	LEX	LEX	LEX	LEX	5	11.9 %
Other	X	X	X	X	15	35.7 %
Total					42	100 %

The new definition of pattern A in these verbs accounts for 45 % of verbs. Although this is almost half of all verbs in Group 3, it is nonetheless significantly less than the percentage of verbs which pattern A accounted for in Group 1. The main reason for this lower percentage is due to the large number of verbs which do not adhere to any pattern and might instead be seen as irregular. But again, it is worth stressing that in paradigms that are maximally distinctive at the segmental level, the reliance on tone for marking paradigmatic distinctions becomes unnecessary, which may explain the greater number of verbs in this group which display some other minor pattern.

6.4 Tonal patterns of Group 4

In Group 4 we consider verbs that have a syllabic prefix in all forms except the PRF. There are 30 verbs which follow this structural pattern, belonging to the four prefix series presented in Table 21. The tonal patterns observed in these verbs are given in Table 22.

Table 21: Schematic representation of micro-prefix series belonging to Group 4

	POT	PROG	CPL	PRF	#
Group 4	CV-	CV-	CV-	C-	30
Series 4a	ku-	yi-	chi-	n-	
Series 8a	ka-	yi-	chi-	n-	
Series 10	ka-	yi-	chi-	∅-	
Series 12a	ku-	yi-	ka-	n-	

Table 22: The tonal patterns of verbs with a syllabic prefix in the POT, PROG and CPL

		POT	PROG	CPL	PRF	Count
A	A1	4~lex	4~lex	4~lex	0~2	2
	A2	3~lex	3~lex	4~lex	0~2/lex	4
		4~lex	4~lex	4~lex	0~2/lex	7
C	C1	4~lex	4~lex	4~lex	0~1	3
	C2	4~lex	4~lex	4~lex	0~1/lex	2
		3~lex	3~lex	4~lex	0~1/lex	2
Invariable		4~lex	4~lex	4~lex	0~lex	3
Other		4~2	4~2	4~2	0~4	1
		4~3	4~3	4~3	0~4	1
		4~43	4~43	4~43	0~41	2
		4~4	4~4	4~14	0~4	1
		2~2	2~2	2~24	0~24	1
		4~24	4~14	4~14	0~14	1
Total						30

Despite being such a small group of verbs, the tone patterns provide real support for the hypotheses already discussed. In this group of verbs, tone 4 in the CPL is again associated with the aspectual prefix, in all but one verb. The PRF, however, either has a non-syllabic prefix or no prefix in these verbs, and it is therefore no surprise that these verbs adopt the strategy already seen in Groups 1 and 2 in which tone 2 appears on the first syllable of the stem. Again, this may either involve replacing the lexical tone (A1) or being added to it (A2), corroborating our hypothesis yet further. We also see instances of verbs taking tone 1 in the PRF (C1,

C2), which we have also seen previously. Table 23 summarizes the distribution of patterns in Group 4 verbs.

Table 23: Summary of patterns observed in Table 22

	POT	PROG	CPL	PRF	Count	Percent
A	LEX	LEX	4-LEX	2	13	43.4 %
B	LEX	LEX	3-LEX	2	0	0.0 %
C	LEX	LEX	4-LEX	1	7	23.3 %
Invariable	LEX	LEX	LEX	LEX	3	10.0 %
Other	X	X	X	X	7	23.3 %
Total					30	100 %

Here we see that pattern A accounts for 43 % of Group 4 verbs, a similar figure to that seen with Group 3 verbs. In this case, however, pattern C accounts for a further 23 % of verbs, bringing the total of verbs accounted for by patterns A–C to 66 %. Again, there is also a relatively high proportion of irregular verbs.

7 Conclusion

As it will have become apparent, Cuicatec verbal inflection is complex. This is due, first and foremost, to the significant number of prefix series and tonal patterns that exist. These two systems cross-cut each other and, in doing so, the number of unique inflectional classes is multiplied. However, it is also due to the fact that the two systems of prefixes and tone are so intertwined with each other that one cannot be divorced from the other, but instead both must be looked at together if one is to disentangle, at least to a degree, the tonal inflection of Cuicatec.

Indeed, a number of noteworthy observations made in this paper may never have come to light had we looked at the tonal patterns of Cuicatec verbs in isolation, without paying heed to the segmental realization of the verb forms. For example, our observations about tone 4 being an inflectional tone used to mark the CPL, and tone 2 to mark the PRF, became clear when we looked at verbs which do not have a syllabic prefix in these forms, while the notion that these tones are lexically associated with a syllabic prefix only became apparent when looking at Group 3 verbs in isolation.

In summary, then, we present the following as our hypothesis regarding the synchronic role of tone in Cuicatec verbal inflection:

- Tone 4 is lexically associated with the CPL prefix (or tone 3 in a subset of verbs).
- Tone 2 is lexically associated with the PRF prefix (or tone 1 in a subset of verbs).
- When a non-syllabic prefix is present in CPL/PRF, these tones are hosted by the first syllable of the verbal stem. As a result, the prefix and tone are co-exponents of inflection.
- If no prefix is present in CPL/PRF, the tone associated with these forms is still hosted by the first syllable of the verbal stem, but in these cases tone is the sole exponent of inflection.

Taking this hypothesis, together with the possible explanation for the co-existence of replacive and additive tone which we alluded to earlier, we arrive at the following hypothesis regarding the possible diachronic development of verbal inflection in Cuicatec:

- Stage 1. Verbal aspect is marked by means of syllabic prefixes, which carry lexical tone.
- Stage 2. Prefixes become phonologically reduced to non-syllabic prefixes, and the lexical tone of the aspectual prefix shifts onto the stem (additive tone).
- Stage 3. Additive tone is later reanalyzed as replacive tone.
- Stage 4. The tone now appearing on the stem is reanalysed as an exponent of inflection.
- Stage 5. Aspectual prefixes are lost due to the presence of dual exponents of inflection.
- Stage 6. Verbal aspect is marked solely by tone.

This hypothesis illustrates how a language may go from marking aspect solely by means of prefixes to marking it solely by tone. As we have seen, Cuicatec uses both prefixes and tone in its verbal inflection and, thus, we would suggest that Cuicatec is currently in the process of moving towards Stage 6. Furthermore, it appears that certain verbs are further along in this sequence of events than others, which in turn provides an explanation for the high degree of internal complexity which we observe in Cuicatec.

On the contrary, we should be clear that we have no issues with accepting that the human mind is capable of handling incredibly complex inflectional systems, such as those seen in the Oto-Manguean languages, and it is of course possible for a language to have multiple exponents of inflection acting as systems which are orthogonal to each other, like in Cuicatec. Nevertheless, is it encouraging to see that, even within such complexity, there are often patterns to be found

when one searches long enough and hard enough. These patterns, in turn, might lead one to a possible explanation for the origin of the complexity and, if they do, all the better.

Acknowledgments

This research was funded by the ESRC/AHRC grant RES-062-23-3126 ‘Endangered Complexity: Inflectional classes in Oto-Manguean languages’. We presented different versions of this paper in three different events: the 2013 Paris-3 workshop on “Tons et paradigmes flexionnels”; the 2013 workshop on “Comparative morphology and morphological theory” at the Annual Meeting of the LAGB in London; and at the 2014 “Conference on Oto-Manguean and neighbouring languages” in Oaxaca, Mexico. We thank the audiences for their feedback and we are grateful to Jean Léo Léonard for his comments.

References

- Anderson, E. Richard and Hilario Concepción Roque. 1983. *Diccionario Cuicateco: Español-Cuicateco, Cuicateco-Español*. Mexico City: Instituto Lingüístico de Verano.
- Bradley, C. Henry and Barbara E. Hollenbach (eds.). 1988. *Studies in the syntax of Mixtecan languages*, volume 1. Dallas, Texas: SIL International.
- Bradley, C. Henry and Barbara E. Hollenbach (eds.). 1990. *Studies in the syntax of Mixtecan languages*, volume 2. Dallas, Texas: SIL International.
- Bradley, C. Henry and Barbara E. Hollenbach (eds.). 1991. *Studies in the syntax of Mixtecan languages*, volume 3. Dallas, Texas: SIL International.
- Bradley, C. Henry and Barbara E. Hollenbach (eds.). 1992. *Studies in the syntax of Mixtecan languages*, volume 4. Dallas, Texas: SIL International.
- Daly, John P. and Larry M. Hyman. 2007. On the representation of tone in Peñoles Mixtec. *International Journal of American Linguistics* 73.2: 165–207.
- DiCano, Christian T. 2012a. Tone-glottal coarticulation in Itunyoso Trique. *Journal of Phonetics*, 40: 162–176.
- DiCano, Christian T. 2012b. Cross-linguistic perception of Itunyoso Trique Tone. *Journal of Phonetics*, 40: 672–688.
- DiCano, Christian T., Jonathan D. Amith and Rey Castillo García. 2014. The phonetics of moraic alignment in Yoloxóchitl Mixtec. *Proceedings from the 4th Tonal Aspects of Languages Conference*, Nijmegen, the Netherlands.
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.

- Macaulay, Monica A. 1996. *A Grammar of Chalcatongo Mixtec*. Berkeley: University of California Press.
- McKendry, Inga. 2013. Tonal association, prominence and prosodic structure in South-eastern Nochixtlán Mixtec. Unpublished PhD Thesis at the University of Edinburgh.
- Pike, Kenneth L. 1948. *Tone languages*. Ann Arbor: University of Michigan Press.
- Suárez, Jorge A. 1983. *The Mesoamerican Indian Languages*. Cambridge: Cambridge University Press.

Enrique L. Palancar, Jonathan D. Amith and Rey Castillo García
12 Verbal inflection in Yoloxóchitl Mixtec

1 Introduction

Mixtec is a language family that together with Cuicatec and Triqui forms the Mixtecan branch of Oto-Manguean, a large and very diverse phylum of languages spoken in Mexico. In Mixtec languages tone carries a significant functional load both in inflection and in derivation. For example, verbs in Mixtec languages have at least three main inflected forms: two aspects: incompletive (also called ‘habitual’, ‘continuative’, or even ‘present’) and completive (also called ‘past’) and one mood: irrealis (also called ‘potential’, or even ‘future’). It is particularly common for tone alone to distinguish the irrealis and the incompletive, though tone alone may also mark the completive. Some Mixtec languages use tone productively to mark negation, as in Metlatonoc (Overholt, 1961:620) and Xochapa (Stark et al., 2006:120). As for derivation, though no longer a productive process, tone may also be manifested in transitivity alternations as in Xochapa (Stark et al., 2006:118) and San Juan Colorado (Stark et al., 1986:168).¹ Similarly, tone is also involved in denominal adjective derivation in Chalcatongo Mixtec in what has been called the ‘adjectival high’ by Hinton et al. (1991) and which Macaulay (Stark et al., 1996:64–65) considers “somewhat productive”.² Yoloxóchitl Mixtec manifests all the preceding uses of tone: a low tone alone is an alternate way to mark the completive, a rising tone (/14/) is used to mark negation, a contrast between /1/ and /3/ tones distinguishes pairs of intransitive and transitive verbs, and a high tone marks not only denominal adjectives but also attributive nouns and stative or resultative verbs as well as the incompletive.³

Despite what we consider interesting tonal and segmental alternations in marking verbal aspect and mood, there have been few studies dedicated to this characteristic of Mixtec and even fewer based on a large lexical database and corpus. In this article, we broach this topic in relation to verbal inflection in

1 Note that for Yosondúa Mixtec a segmental morpheme (*x-* or *s-*) marks the transitive of intransitive/transitive pairs (Beatty de Farris et al. 2002:164). A similar segmental difference, along with tonal alternations, marks transitivity in San Juan Colorado Mixtec (Stark, Johnson, and Lorenzo 1986:167–8).

2 Tonal variation accompanied by a final glottal stop is used to derive stative verbs in Ayutla (Hills 1990:198–9).

3 It is not altogether unlikely that there is some historical link between the use of high tone in the production of adjectives and its function in verbal morphology to mark the incompletive aspect as well as to form stative forms (see McKendry 2013:32–34, and Macaulay 1996:65 and *passim*).

Yoloxóchitl Mixtec, a Mixtec language spoken in a small cluster of villages located approximately three-and-a-half hours by car south of Acapulco along the Pacific Coast of the state of Guerrero (see map 1). We base our analysis on a large sample of 554 verbs from primary data collected by Amith and Castillo García.



Map 1: Location of Yoloxóchitl Mixtec in Mexico

Yoloxóchitl Mixtec (henceforth YM) is characterized by its high number of tonal contrasts (nine basic tones and over twenty-three lexical tonal melodies on bimoraic stems [see (1) and fn. 9]) and, not unexpectedly, the almost complete absence of tonal sandhi.⁴ Nevertheless, the tone of some enclitics may be conditioned by either the phonological environment (i.e., a stem-final tone affects the tone of a following enclitic) or by the syntactic environment (i.e., certain enclitics manifest a lower tone in phrase medial as opposed to phrase final position). As in other Mixtec languages, tone in YM carries a high functional load: tonal contrasts not

⁴ One of the rare examples of possible sandhi is the sequence *ku²un¹* ('go' [IRR]) + verb with initial tone /3/, which shows a lowering of the initial tone of the verb: *ku²ũ¹ + ku³chi³* ('bathe') → *ku²ũ¹ ku³chi³* ('go to bathe'). Interestingly, the Mixtec of Cuanacaxtitlan, a closely related neighboring village about 5 km southeast of Yoloxóchitl, manifests sandhi in some cases. For example, Christian DiCano observed that at least in some cases word-final high tones (/4/) spread to low (/1/) initial tones on following words. Thus *ku¹mi⁴* ('four') followed by *ka¹ta¹* ('sugar press [trapiche]') is realized as *ku¹mi⁴ ka¹ta¹* in Yoloxóchitl but *ku¹mi⁴ ka⁴ta¹* in Cuanacaxtitlan. The extent of this tone spreading (i.e., whether all word-final high tones spread to a low initial tone on a following word or whether this phenomenon may be lexically determined) still needs to be investigated.

only mark the aforementioned aspect and mood distinctions but person marking as well. After a stem final mid-high (/3/) or high (/4/), a tone /2/ encodes the first person (i.e., the subject of a predicate or the possessor of a noun). This tone is in complementary distribution with =*yu*¹, used to mark first person after stems that end in tone /2/ or /1/ (see §4.1). YM tone is also involved in a set of other lexico-semantic changes. One, the ‘adjectival high’, is reported elsewhere (in YM, for example, *yu*¹*u*⁴ ‘stone’ > *yu*⁴*u*⁴ ‘solid’, *xaɽ*²³*an*² ‘fat’ > *xaɽ*²⁴*an*² ‘greasy’, and *ñuɽ*²³*u*⁴ ‘earth’ > *ñu*²⁴*u*⁴ ‘earth colored’). The function of other, somewhat irregular, tonal variations is not so easily defined and requires further study although many such variations seem related to the adjectivization function already mentioned. These additional tone-marked distinctions include:

- Marking of a noun as an attributive, a function particularly common in plant and animal names: *tio*¹*ko*⁴ ‘ant’ > *tio*¹⁴*ko*³ in the compound *i*³*ta*² (‘flower’) + *tio*¹⁴*ko*³ ‘*Montanoa grandifolia* DC’, a plant of the Asteraceae family known for its pungent ‘ant-like’ smell, and *tio*⁴*ko*⁴ in the compound *ña*¹*ña*⁴ + *tio*⁴*ko*⁴ (‘ant eater’); *tiɽ*²¹*ɽ*⁴ ‘skunk’ > *tiɽ*⁴*ɽ*⁴ in the compound *yu*³*ba*² (‘edible green’) + *tiɽ*⁴*ɽ*⁴ ‘*Solanum nigrescens* M. Martens & Galeotti’, an edible plant of the Solanaceae family; *yu*³*ku*⁴ ‘hill’ > *yu*¹*ku*⁴ in the compound *ki*³*ni*² (‘pig’) + *yu*¹*ku*⁴ ‘wild boar’; and finally *i*³*su*² ‘deer’ > *i*⁴*su*⁴ in the compound *ko*¹*o*⁴ (‘snake’) + *i*⁴*su*⁴ ‘boa’ (‘deer snake’ is a common Mesoamerican calque).⁵
- Irregular adjectival/nominal marking: *laɽ*⁴*la*¹ ‘worn-out (bags, hammocks)’ > *laɽ*²⁴*la*¹ ‘worn-out clothes’.
- Marking of scale: *bi*¹*xī*³ ‘cold’ ~ *bi*¹*xī*¹ ‘lukewarm’.
- Marking of the characterization of time periods: *sa*¹*bi*⁴ ‘rain’ > *sa*⁴*bi*⁴ ‘rainy season’.
- Different modification functions: *ti*¹*ka*¹*yu*¹ ‘jet’ as a modifier used only with *nda*⁴*a*⁴ ‘black’ to indicate ‘jet black’ and in *ti*⁴*ka*¹*yu*¹ ‘carbonized’ (as in a heavily burnt tortilla).

Despite the rich and unexplored occurrence of tonal alternations in derivational morphology, in this paper we limit our study of tone to the marking of mood and aspect on verbs, a marking that manifests a fairly high degree of morphological regularity. That is, specific tones are consistently associated to specific aspect/mood values. Section 2 presents a brief overview of YM. The following section begins with a summary presentation of YM segmental and autosegmental phonology. It continues with a discussion of the role of tone in YM inflectional and derivational morphology and patterns of tonal allophony that are dependent on

5 For calques, see Campbell et al. (1986) and Smith-Stark (1982, 1994)

phonological and morphosyntactic factors. Section 4 presents patterns of tonal variation related to verbal inflection in the irrealis, incomplete, completive and negative irrealis. We examine both tonal marking and segmental alternations. This study ends, section 5, by comparing YM verbal inflection to that found in closely related Xochapan Mixtec (state of Guerrero) and that of the more distantly related San Juan Colorado Mixtec (state of Oaxaca). As a result of this comparison we suggest the possibility of a regional pattern of verbal inflection through tone that typifies Guerrero Mixtec.

2 Yoloxóchitl Mixtec

Mixtec is here considered to be a language *family*, part of a larger unit, Oto-Manguean, that Suárez [1983:26] considers to be a ‘hyper-family’ or a ‘stock’. Mixtecan languages (spoken in Oaxaca [156 municipalities], Guerrero [13 municipalities], and Puebla [10 municipalities]) are highly varied internally, the result of approximately 2,000 years of diversification. Estimates of the number of Mixtec languages vary (*Ethnologue* lists 53, Smith Stark [1995] mentions 45; Suárez (1983) estimates about 29 distinct languages; Bradley & Hollenbach [1988:1] suggest “perhaps twenty mutually unintelligible languages”). Likewise, the criteria utilized for such determinations also vary: mutual intelligibility is favored by SIL and *Ethnologue*; others (e.g., Josserand, 1983) compare lexicon and morphology in establishing isoglosses.

Following Josserand (1983), Mixtec languages are now commonly organized into five branches, represented in Figure 1. Recognition of this internal diversity is not recent. The first grammarian of Mixtec, Fray Francisco de Alvarado, already remarked on the internal diversity of the family in his grammar of 1593.

Both YM and Xochapan Mixtec (which, along with San Juan Colorado Mixtec, we discuss in the conclusion) are treated by Josserand (1983) as belonging to the Guerrero subgroup, which comprises six main varieties; Xochapan Mixtec is classified under the Alcozauca variety. Castillo García (2007), a native speaker of YM, reports reasonably good mutual intelligibility across the Guerrero subgroup, but the degree of intelligibility drops considerably when one compares YM with neighboring Mixtec varieties of the Southern Baja group such as Ayutla Mixtec. San Juan Colorado Mixtec is more distant and would probably show a low level of mutual intelligibility with YM although Castillo García does not provide any qualitative judgements on the degree of intelligibility.

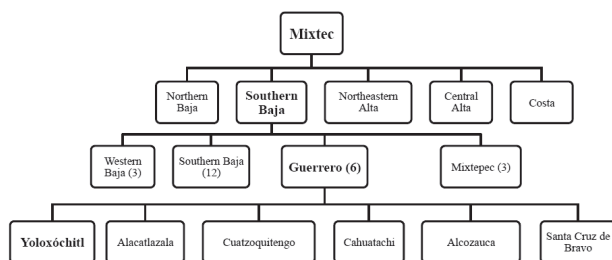


Figure 1: Mixtec family based on Josserand (1983), from DiCano et al. (2014)

Yoloxóchitl Mixtec is spoken in four villages along the Pacific Coast of southern Guerrero, Mexico. The greatest linguistic vitality is found in the villages of Yoloxóchitl (pop. approx. 3,000) and Arroyo Cumiapa (pop. approx. 1,500), about 6.5 miles to the northeast of Yoloxóchitl. In both of these villages almost all speakers are highly fluent. Yoloxóchitl Mixtec is also spoken, though rapidly disappearing, in Cuanacaxtitlan (pop. approx. 4,000) and Buenavista and its surrounding farms (pop. approx. 5,000). In both these latter areas language loss, shift to Spanish and significant relexification is virtually complete among the younger generation. It is fair to say that in both Cuanacaxtitlan and Buenavista the ‘tipping point’ has been reached and the language has passed from one side to the other of what Fishman (1991) calls “the continental divide”, that is, the point at which intergenerational transmission of language ceases and maintenance becomes an often fruitless endeavour.

3 Basics of the segmental and tonal phonology of Yoloxóchitl Mixtec

YM has 17 consonant segments and five cardinal vowels (DiCano et al. 2014). We consider both nasalization and laryngealization to be autosegmental features. With few exceptions (apparently the result of recent compounding) nasalization is restricted to the stem-final mora; laryngealization occurs only on the penultimate mora. Thus when a laryngealized stem combines with a following bimoraic stem, laryngealization is lost in the now non-penultimate position: $xaʔ^l a^4$ ‘foot’ + $kuʔ^l u^l$ ‘area (be it urban or distant from a village) overgrown with weeds’ becomes $xa^{14} kuʔ^l u^l$ ‘distant (from an inhabited area) woods’. There are other co-occurrence restrictions: a laryngealized vowel can only precede a stem-final consonant if it is

voiced; a nasalized vowel can only follow a stem-final consonant if it is voiceless. Given these restrictions there are no CVCV words in YM with both a laryngealized and nasal vowel although there are CVV and CVʔV stems with both features: *iʔ³in³* ‘scratch’ (transitive verb).⁶

The five cardinal vowels also manifest interesting distributional patterns. With virtually no exception the two vowels of CVV stems are identical, a pattern that suggests either vowel lengthening to fulfil phonotactic constraints on minimal word formation or a rigorous process of vowel harmonization. Regardless of the stem syllable structure, however, stem nasal vowels are limited (with one exception) to /ã/, /ĩ/, and /ũ/.⁷ The occurrence of nasal /ẽ/ and /õ/ in enclitics is probably the result of phonologically motivated vowel quality shift from underlying /ã/ and /ũ/, respectively.

YM has nine basic tones that can occur on a single mora: four level tones (written as superscript numbers from /1/ for the low tone through /4/ for the high tone); three rising tones (/13/, /14/, and /24/) and two falling tones (/32/ and /42/).⁸ Again there are significant distributional gaps. Tone /2/ is never found on the initial mora; /32/ and /42/ are only word final and occur in the absence of any contrast with /31/ and /41/, both of which are absent in YM tonal phonology. Even though it does bear significant functional weight as a 1st-person enclitic after stems that end in final /4/ or /3/, it is highly likely that the low middle /2/ is innovative. Finally, of the 89 occurrences of final-mora rising tones found in the lexicon (/13/, /14/, and /24/) all but two occur after a previous high (/4/) tone.⁹

Monomoraic words are rare. They are limited to some 40 in total and are almost exclusively function words such as modals, clausal markers and adverbs. The vast majority of content words in YM are bimoraic, either mono- or disyllabic, though longer trimoraic words are not uncommon. There are 23 tonal melodies on bimoraic lexical stems. As expected, some tonal melodies are more frequent than others. Indeed, two of the twenty-three melodies on bimoraic stems (3–13 and 3–14; see fn. 9) are limited to one word each. The remaining melodies are exemplified in the nouns and adjectives in (1). However, if morphological tone

⁶ Although we consider laryngealization an autosegmental feature, we still represent this phonation type with <ʔ>. Likewise, following convention we write nasal vowels with a following <n>, although an orthography using forms with a tilde might be preferable.

⁷ The only stem occurrence of /ẽ/ is in the noun *le³en⁴* ‘genital liquid’ (male or female).

⁸ There is one function word and two adverbs, all monomoraic, with /143/. Two plant names have similarly complex tones: *pi²la⁴xi¹³²* ‘*Jaltomata darycana*’ and *ku⁴tu¹pi³⁴²* ‘*Plumbago scandens*’. In this latter case the name is onomatopoeic for the sound the flowers make when blown through by playing children.

⁹ The melody /3–13/ is only attested in the passive-oriented verb *kia²³bi¹³* ‘be sold’, while /3–14/ occurs only in the quantifier *i³nda¹⁴* ‘one’.

were to be taken into account (e.g., verbal inflection for aspect and mood, negation, and 1st-person marking), the number of realized patterns on bimoraic stems would increase considerably. For example, including inflectional morphology, the sequence /na^Tma^T/ alone manifests 21 tonal contrasts. These are different from the 21 lexical melodies as the verb paradigm includes tonally marked morphemes (e.g., negation and first person).

(1)	/1-1/	bi ¹ ka ¹	‘comb’	/4-3/	i ⁴ in ³	‘mute’
	/1-3/	ta ¹ a ³	‘man’	/4-4/	sa ⁴ bi ⁴	‘rainy season’
	/1-4/	xi ¹ i ⁴	‘grandfather’	/4-13/	che ⁴ e ¹³	‘big’
	/1-32/	xa ¹ ko ³²	‘opossum’	/4-14/	na ⁴ ni ¹⁴	‘long (PL)’
	/1-42/	ta ¹ kwi ⁴²	‘water’	/4-24/	ya ⁴ a ²⁴	‘tongue’
	/3-2/	ñu ³ u ²	‘village’	/13-2/	tĩ ¹³ bi ²	‘lightning bug’
	/3-3/	ta ³ ni ³	‘animal breast’	/14-1/	na ¹⁴ a ¹	‘demoniac’
	/3-4/	bi ³ ko ⁴	‘feast’	/14-2/	ma ¹⁴ ñu ²	‘central place’
	/3-42/	ñu ³ u ⁴²	‘night’	/14-3/	nu ¹⁴ u ³	‘face’
	/4-1/	ya ⁴ a ¹	‘darkish, brown’	/14-4/	ye ¹⁴ e ⁴	‘door’
	/4-2/	xa ⁴ a ²	‘lime-soaked maize (nixtamal)’			

Some melodies are common across all syllable types (/1-1/, /1-3/, /1-4/, /3-2/, /3-3/, /3-4/, /4-2/ and /4-4/) and at least six patterns are much more common in disyllabic than in monosyllabic bimoraic words (/1-32/, /4-1/, /4-3/, /13-2/, /14-2/, /14-3/). Finally, in addition to /3-13/ and /3-14/, which occur in one word each, seven patterns are extremely rare in the lexicon (/1-42/, /3-13/, /3-14/, /3-42/, /4-13/, /4-14/, /14-1/). Moreover, the distribution of tones is asymmetrical. While the first mora allows for only five possibilities (/1/, /3/, /4/, /13/, /14/), the second mora, which is lengthened, allows for nine (/1/, /2/, /3/, /4/, /13/, /14/, /24/, /32/, /42/) although two, /13/ and /14/, are relatively rare). Note that /31/ is never found in lexical or inflected words, either in a bimoraic melody /3-1/ or on a single mora /31/.

As the object of our study is the role of tone in the making of the verbal inflection of YM, in the next section we concentrate on verbs.

3.1 Sample and the tonal structure of verbs

Our analysis of the verbal inflection of YM is based on a sample of 554 verbs taken from a large lexical database presently comprising 2,192 lexical entries. This database was compiled by Jonathan D. Amith and Rey Castillo García as a principal outcome for the documentation project “Corpus and lexicon development:

Endangered genres of discourse and domains of cultural knowledge in Tu²un³ i⁴sa¹bi¹⁴ ('the Mixtec language') of Yoloxóchitl, Guerrero" (see also Amith & Castillo García, n.d. 1 and n.d. 2).¹⁰

The sample of 554 YM verbs used in this study is balanced in regard to moraic structure and transitivity. The sample contains an almost equal number of bimoraic and trimoraic verbs in almost equal proportion regarding transitivity (2), ('intr' stands for 'intransitive'; 'tr' for 'transitive' and 'ditransitive').

(2)	intr	tr/dtr	Total
Bimoraic	183	99	282
Trimoraic	182	90	272
Total	365	189	554

Many trimoraic words in YM are formed by compounding or by derivation (the latter may involve productive mechanisms or archaic ones, the latter resulting in what are now frozen forms). This has consequences for the distribution of the tonal patterns in such trimoraic words as they are associated with the patterns found on the bimoraic words on which they are based. For example, a denominal verb formed by adding a particular prefix to a nominal stem will have a tonal pattern properly of nouns over its final two morae. A detailed analysis of the relation between tonal melodies in bimoraic and trimoraic words, however, would require a level of effort that is beyond the scope of the present article. The focus below is on bimoraic verbs.

3.1.1 Bimoraic verbs

The lexicon of YM has no monomoraic verbs, so that like nouns and adjectives a verb with a bimoraic stem instantiates the basic lexical template for a verbal

¹⁰ A note on the structure of the dictionary is in order. Primary verbal entries are found under the irrealis form, but the total number of entries also includes approximately 120 'place-holders' or cross-references. For example, when a verb has an irrealis form which stands in a suppletive relation to the incomplete, the incomplete form is given an entry but this entry is without semantic content. The entry is simply a pointer to the main verb entry, in the irrealis mood. Additionally, dictionary entries have also been established for iterative verbal formations using the prefix *nda*³ when the semantics of the iterative are not transparent.

lexeme.¹¹ There are 282 bimoraic verbs in our sample. Some examples appear in (3), inflected in the irrealis.

- (3) a. ka³a⁴ (intr) ‘slip’
 tu³tu⁴ (intr) ‘whistle’
 ka³xi⁴ (tr) ‘eat’ ‘bite’ or ‘bite into’ (foods such as fruit, popcorn, *pozole* and sometimes meat, chicken or cheese)¹²
 sa³ta⁴ (tr) ‘buy’
 b. ke¹yu⁴ (intr) ‘swim’
 c. chi³kun² (tr) ‘hang’

The verbs in (3) have different tonal melodies: /3-4/ in (3a), /1-4/ in (3b) and /3-2/ in (3c), respectively. With few exceptions, the tonal melody of the irrealis represents lexical tone, i.e., the tonal melody representative of the basic lexical form.

Throughout the YM verbal lexicon, contrastive tonal melodies also correlate with lexical contrast. For example, the pairs in (4) differ only in tone, their segments being identical. The contrasts may involve differences in the first mora as in (4a), differences in the second mora, as in (4b), or differences in both morae, as in (4c).

- (4) a. ka¹ku³ /1-3/ (intr) ‘escape’ ~ ka³ku³ /3-3/ (intr) ‘be born’
 ke¹ta³ /1-3/ (intr) ‘get in’ ~ ke³ta³ /3-3/ (intr) ‘get out’
 (SG sub.) (SG sub.)
 ki¹ni⁴ /1-4/ (intr) ‘be balanced’ ~ ki³ni⁴ /3-4/ (tr) ‘shoot’
 b. ka³ba³ /3-3/ (intr) ‘lie down’ ~ ka³ba⁴ /3-4/ (tr) ‘turn around’
 ka³si² /3-2/ (tr) ‘take a shortcut’ ~ ka³si⁴ /3-4/ (intr) ‘get goose bumps’
 ka¹an¹ /1-1/ (intr) ‘get drilled’ ~ ka¹an³ /1-3/ (intr) ‘get used’
 c. ka^ʔa¹ /1-1/ (intr) ‘drown’ ~ ka^ʔa³ /3-3/ (intr) ‘emit a sound’

¹¹ Our sample includes bimoraic and trimoraic verbs, but excludes several quadrimoraic verbs present in the dictionary (e.g. *ta³xi³kwa^ʔa⁴* ‘keep’, *ka³sa³chii⁴un⁴* ‘work’, among others). Such quadrimoraic verbs, nevertheless, follow the same principles as bimoraic and trimoraic verbs.

¹² *Ka³xi⁴* contrasts with *ku³xi³*, the latter used for eating meals such as mole, tortillas and most other “meal-type” foods. The eating of meat can be expressed by either verb in patterns, if any, that are not yet fully understood.

ka ¹ sun ¹ /1-1/ (intr)	‘creak’	~ ka ³ sun ² /3-2/ (intr)	‘get fried’
ti ³ bi ² /3-2/ (intr)	‘blow out air’ or ‘play a wind instrument’	~ ti ¹⁴ bi ³ /14-3/ (intr)	‘rot’

Minimal triplets also exist, and they emerge as a combination of contrasts such as those in (4), which combine in unpredictable ways. For example, the triplets in (5) each show different patterns of contrast.

- (5) a. kwi¹i⁴ /1-4/ (intr) ‘fight’
 kwi³i⁴ /3-4/ (tr) ‘peel’
 kwi³i³ /3-3/ (intr) ‘be able’
 b. kwi¹so¹ /1-1/ (intr) ‘boil’
 kwi¹so³ /1-3/ (intr) ‘be pricked’
 kwi³so³ /3-3/ (tr) ‘load’
 c. ta¹ni¹ /1-1/ (intr) ‘be spoilt’
 ta¹⁴ni³ /14-3/ (intr) ‘bruise’
 ta³ni³ /3-3/ (intr) ‘be carried away by water’
 d. tu³u³ /3-3/ (tr) ‘sting’
 tu¹u³ /1-3/ (intr) ‘appear’
 tu³u⁴ /3-4/ (intr) ‘roll’

The irrealis forms of the 282 bimoraic verbs in our sample display a total of 12 different tonal melodies, given in Table 1. Lexical verb stems, therefore, manifest only 12 of the 23 documented tonal possibilities for bimoraic stems. This absence is reflective of the fact, discussed in continuation, that high tone in the first mora of verbs is ‘reserved’ to mark the incompletive aspect.¹³ Among the absent patterns, /13–2/ is not uncommon in the overall lexicon. The tonal pattern /3–13/, however, is found only in one verb, *kia²³bi¹³* (‘be sold’), but in no other bimoraic lexical stem. In bimoraic words, /14–1/ is found in one verb, *chu¹⁴ma¹* ‘be censured (with incense)’, two nouns and one adjective; two nouns and one adjective; /14–4/ is quite common although found in only one verb, *xio¹⁴o⁴* (‘become ill after having craved a food’).

In Table 1, ‘0’ indicates tonal melodies that are attested elsewhere in the lexicon but are not found in bimoraic verbs; ‘–’ indicates tonal melodies that are

¹³ Indeed, seven of the eleven absent melodies have an initial /4/ tone.

not attested anywhere in the lexicon. As noted above, there are no bimoraic verbs in the irrealis with a lexical high tone /4/ on the first mora (μ_1)¹⁴ as this tone is the inflectional exponent of the incomplete (see §5.2). A further analysis, Table 2, reveals that by default bimoraic verbs have tones /1/ or /3/ (only a few have /14/) on μ_1 . There is a greater range of tone values for the second mora (μ_2), although 45 percent of the total number of verbs (126 of 282) have tone /3/ on μ_2 . Many verbs also have harmonized melodies: 46 (16.3%) have /1-1/ and 85 (30.0%) have /3-3/. That is, close to 50 percent of bimoraic irrealis verbs have harmonized tone.

Table 1: Tonal melodies in lexical stems of bimoraic verbs (irrealis forms only)

		μ_2						
		/1/	/2/	/3/	/4/	/13/	/32/	
μ_1	/1/	46	–	34	26	–	4	110
	/3/	0	25	85	45	1	0	156
	/4/	0	0	0	0	0	0	0
	/13/	0	0	0	0	0	0	0
	/14/	1	7	7	1	0	0	16
		47	32	126	72	1	4	282

The distribution of tonal melodies in lexical entries displays an interesting correlation with transitivity. Consider for this purpose the data in Table 2, where we have given the transitivity status of bimoraic verbs.

Table 2: Lexical tonal melody and transitivity in bimoraic verbs

		μ_2									
		/1/	/2/	/3/	/4/	/13/	/14/	/32/	intr	tr	
μ_1	/1/	intr	38	0	30	24	0	0	4	96	
		tr	8	0	4	2	0	0	–		14
	/3/	intr	0	11	54	13	1	0	0	79	
		tr	0	14	31	32	0	0	0		77
	/14/	intr	1	3	3	1	0	0	0	8	
		tr	0	4	4	0	0	0	0		8
									183	99	

¹⁴ One stative manifests the melody /4-14/, which is limited to four words, all disyllabic.

We see from Table 2 that almost 90 percent (96 out of 110) of bimoraic verbs with tone /1/ on μ 1 are intransitive. Verbs with /1/ on the first mora constitute half of the intransitive bimoraic verbs (96 out of 183) but only 15 percent of the transitives (14 out of 99). This reveals a strong correlation between phonology and transitivity: a verb bearing lexical melodies such as /1-1/, /1-3/, /1-4/ or /1-32/ is expected to be intransitive. In contrast, a tone /3/ or /14/ on μ 1 is uninformative about transitivity.

The correlation of tonal melody and transitivity is partly exploited in the lexicon through the existence of a set of valence pairs that contrast only in tone. In such patterns, the intransitive verb often depicts an inchoative state of affairs, a sort of passive meaning. That is, the semantics of the intransitives suggest an implicit and unexpressed agent. This indicates that, at least historically, the direction of derivation might have been intransitivization. This is consistent with the fact that tone /1/ is closely associated with intransitivity (+ 90%) while tone /3/ is uninformative in this regard.¹⁵

There are various patterns involved in such intransitive/transitive pairs, given in (6), all in the irrealis form. In all of them, the intransitive verb has tone /1/ on μ 1 and the transitive verb tone /3/. In patterns (6a) and (6b), the tone on μ 2 remains constant, which would seem to indicate that the transitivity alternation, at least historically, was motivated by a tonal change on the first mora alone, a pattern to be expected given the position of morphological tone in general. In (6c) and (6d) the tone on the second mora of the transitive shifts as well, but in (6c) this is clearly due to a phonotactic constraint on /3-1/ melodies, which occur nowhere in YM. In (6d) the upward shift to tone /4/ on the final mora is irregular and unexplained. It might reflect some underlying tonal difference between the two sets: (6c) regular from a phonotactic constraint and (6d) irregular transitive patterns from /1-1/ intransitives.¹⁶

15 The agent can be syntactically expressed but only in a juxtaposed clause with the meaning ‘X did it’. Such constructions are used when the patient is topic. Note the following:

- (i) a. ni¹-ku¹ni⁴ yaʔ3a4 [i13xa3 Juan]
 CPL-be.ground chilli CPL.do John
 ‘The chilli (old information) was ground by John.’
 b. ni¹-koʔ¹ni⁴ ji¹ndi¹ki⁴ [i¹³xa³ Juan]
 CPL-be.milked cow CPL.do John
 ‘The cow (old information) was milked by John.’

16 Other alternations between /1/ ~ /3/, are possibly related to the alternation that marks transitivity: *ke¹ta³* ‘enter’ (SG subj.) and *ke³ta³* ‘exit’ (SG subj.) and, despite the segmental difference, *nde¹e³* ‘enter’ (PL subj.) and *ke³e³* ‘exit’ (PL subj.).

- (6) a. ku^1un^4 /1-4/ (intr) ‘be ground’ ~ ku^3un^4 /3-4/ (tr) ‘grind chilli for sauce’
 $koʔ^1ni^4$ /1-4/ (intr) ‘be milked’ ~ $koʔ^3ni^4$ /3-4/ (tr) ‘milk’
 kwi^1i^4 /1-4/ (intr) ‘peel off’ ~ kwi^3i^4 /3-4/ (tr) ‘peel’
 nda^1ta^4 /1-4/ (intr) ‘get split in two’ ~ nda^3ta^4 /3-4/ (tr) ‘split in two’
 $taʔ^1bi^4$ /1-4/ (intr) ‘get broken’ ~ $taʔ^3bi^4$ /3-4/ (tr) ‘break’
 ta^1xi^4 /1-4/ (intr) ‘be fired from work’ ~ ta^3xi^4 /3-4/ (tr) ‘fire from work’
 tu^1xi^4 /1-4/ (intr) ‘be crushed’ ~ tu^3xi^4 /3-4/ (tr) ‘crush’
 b. na^1ma^3 /1-3/ (intr) ‘get changed’ ~ na^3ma^3 /3-3/ (tr) ‘change’
 c. ka^1an^1 /1-1/ (intr) ‘get drilled’ ~ ka^3an^2 /3-2/ (tr) ‘drill’
 tu^1un^1 /1-1/ (intr) ‘get turned on’¹⁷ ~ tu^3un^2 /3-2/ (tr) ‘turn on’
 d. ka^1ba^1 /1-1/ (intr) ‘turn around’ ~ ka^3ba^4 /3-4/ (tr) ‘turn around’
 $kaʔ^1yu^1$ /1-1/ (intr) ‘be written’ ~ $kaʔ^3yu^4$ /3-4/ (tr) ‘write’

3.1.2 Trimoraic verbs

Apart from bimoraic verbs, a significant number of YM verbs are trimoraic. There are 272 such verbs in our sample, displaying a total of 22 different tonal melodies, illustrated in Table 3.

As with bimoraic verbs, trimoraic verbs may also have tones /1/, /3/ or /14/ on μ_1 , but tones /1/ and /14/ are very rare (found only in 6 verbs). Thus trimoraic melodies almost always start with tone /3/. This would not be unexpected if, historically, trimoraic verbs were formed from a derivational morpheme prefixed to a bimoraic verbal stem. Most derivational verbal prefixes do indeed carry a mid tone /3/ (the completive marker *ni^l-* being an inflectional morpheme). Thus the tonal pattern on μ_1 of trimoraic verbs is correlated to the tones of derivational morphemes. This consistency of tone marking on the first mora of trimoraic verbs means that the functional lexicosemantic contrast lies on μ_2 and μ_3 . In many ways the melodies of the final two morae of trimoraic verbs are similar to the melodies of bimoraic verbs, except for those cases with a high tone /4/ or a low-mid tone /2/ on the middle mora. Instances of tone /4/ occur most often in deadjectival verbs. A mid-mora /2/ occurs only after initial-mora /3/. Compare Table 4 with Table 1. Relevant contrasts have been highlighted.

¹⁷ Said of an electric device.

Table 3: Tone melodies in trimoraic verbs

Melody	#	intr	tr	Example		
/1-1-3/	1	1	0	ku ¹ nda ² a ³	(intr)	'move away'
/1-1-4/	1	1	0	ku ¹ nu ² u ⁴	(intr)	'get back home'
/1-4-4/	1	1	0	ki ¹ tu ² un ⁴	(intr)	'be pulled out (part of a whole)'
/1-4-2/	1	1	0	tu ¹ xu ² u ²	(intr)	'get harmed'
/3-1-1/	20	16	4	nda ³ kwi ¹ kun ¹	(intr)	'sprout, spring'
/3-1-3/	21	19	2	ka ³ ki ² i ³	(intr)	'trip'
/3-1-32/	1	1	1	ku ³ bi ¹ xi ³²	(intr)	'grow white hair'
/3-1-4/	12	11	1	ku ³ ma ¹ ni ⁴	(intr)	'be lacking'
/3-2-2/	8	3	5	ndo ³ ko ² o ²	(intr)	'get up of bed'
/3-3-2/	34	23	11	nda ³ ye ² e ²	(intr)	'shine'
/3-3-3/	57	35	22	ku ³ ndi ³ xi ³	(intr)	'get dressed'
/3-3-4/	43	24	19	nda ³ ke ³ e ⁴	(intr)	'become stretched'
/3-3-42/	2	2	0	ku ³ nū ³ u ⁴²	(intr)	'get dark (at night)'
/3-4-1/	5	4	1	ndu ³ ka ⁴ chi ¹	(intr)	'become equal'
/3-4-13/	1	1	0	xi ³ ka ⁴ ba ¹³	(intr)	'get turned around'
/3-4-2/	7	4	3	ko ³ se ² e ²	(intr)	'hide oneself'
/3-4-24/	11	10	1	xu ³ ku ⁴ tu ²⁴	(intr)	'roll'
/3-4-3/	18	10	8	ku ³ bi ⁴ ka ³	(intr)	'get rich'
/3-4-4/	22	12	10	ta ³ xa ² a ⁴	(intr)	'dance'
/3-14-2/	2	1	1	ku ³ to ¹⁴ o ²	(intr)	'fancy or crave'
/3-14-3/	2	1	1	ku ³ nu ¹⁴ u ³	(intr)	'look around carefully'
/14-1-1/	2	1	1	xo ¹⁴ kwi ¹ in ¹	(intr)	'turn around to see'
Total	272	182	90			

Table 4: Tonal melodies of last two morae in lexical stems of trimoraic verbs

		μ3									
		/1/	/2/	/3/	/4/	/13/	/14/	/24/	/32/	/42/	Total
μ2	/1/	22	0	22	13	0	0	0	1	0	58
	/2/	0	8	0	0	0	0	0	0	0	8
	/3/	0	34	57	43	0	0	0	0	2	136
	/4/	5	8	18	23	1	0	11	0	0	66
	/13/	0	0	0	0	0	0	0	0	0	–
	/14/	0	2	2	0	0	0	0	0	0	4
	/24/	0	0	0	0	0	0	0	0	0	–
	/32/	0	0	0	0	0	0	0	0	0	–
	/42/	0	0	0	0	0	0	0	0	0	–
	Total	27	52	99	79	1	–	11	1	2	272

intransitive verb has the stem formative ku^3/ko^3+ while the transitive varies: the examples in (9) illustrate ka^3+ , chi^3+ and ta^3+ as the transitive formatives.

- (9) a. $ku^3+ta^3ni^3$ /3-3-3/ (intr) ‘hang down’ ~ $ka^3+ta^3ni^3$ /3-3-3/ (tr) ‘hang down’
 $ku^3+ndi^3chi^2$ /3-3-2/ (intr) ‘get up’ (SG sub.) ~ $ka^3+ndi^3chi^2$ /3-3-2/ (tr) ‘place up’ (SG obj.)’
 b. $ku^3+ndu^2u^4$ /3-4-4/ (intr) ‘sit down’ (SG sub.) ~ $chi^3+ndu^2u^4$ /3-4-4/ (tr) ‘put down’ (SG obj.)’
 c. $ko^3+se^2e^2$ /3-4-2/ (intr) ‘hide oneself’ ~ $ta^3+se^2e^2$ /3-4-2/ (tr) ‘hide oneself’

Other pairs such as those in (10) involve trimoraic verbs derived from suppletive bimoraic stems in which the transitivity alternation is marked on the segmental contrast of both the middle syllable (the initial consonant of the base bimoraic stem) and on the stem formative (10a). Another pattern is seen in (11), in which in addition to the suppletive stem, the transitivity alternation is linked to a prefix. This prefix may be associated with the intransitive (11a) or transitive (11b), a situation probably related to the detransitivizing (11a) or the transitivizing (11b) semantics of the formative.

- (10) a. $ndu^3+ndi^3kun^2$ /3-3-2/ (intr) ‘regererate’ ~ $nda^3+chi^3kun^2$ /3-3-2/ (tr) ‘regererate’
 b. $nda^3+ta^2nu^1$ /3-1-1/ (intr) ‘get bent’ ~ $nda^3+ka^2nu^1$ /3-1-1/ (tr) ‘bend’
 (11) a. $ku^3+ndi^3kun^2$ /3-3-2/ (intr) ‘get hung up’ ~ chi^3kun^2 /3-2/ (tr) ‘hang up’
 b. nu^1na^4 /1-4/ (intr) ‘become open’ ~ $ndu^3+ku^3na^4$ /3-3-4/ (tr) ‘open’

Although the base verb of (11b) is suppletive at the segmental level (/n/~k/), the tonal alternation is reminiscent of that found in (6) in which /1/ is associated with the intransitive and /3/ with the transitive. Thus while the valence of the second verb in (11b) might be accounted for by the tone /3/, it is also possible that ndu^3 -performs some sort of transitivizing function, such as that typical of a causative. Such seemingly causative morphemes are illustrated by the verbs in (12). In no case, however, are either cha^3+ or xi^3+ productive, as is the commonly used causative marker sa^4- , discussed below.

- (12) a. ka^3ba^3 /3-2/ (intr) ‘lie down’ ~ $cha^3+ka^3ba^3$ /3-3-2/ (tr) ‘lay down’
 kwi^3in^3 /3-3/ (intr) ‘be hung’ ~ $cha^3+kwi^3in^3$ /3-3-3/ (tr) ‘hang up’
 or ‘hanging up’ (PL sub.)
 sub.)
 b. kwa^2a^3 /3-3/ (intr) ‘be made’ ~ $xi^3+kwa^2a^4$ /3-4-4/ (tr) ‘be made’
 or ‘be built’

Stem formatives such as ndu^3+ , cha^3+ or xi^3+ are not productive and are restricted to a small number of lexical forms. In some cases the historical derivation is apparent: $chi^3+kur^3ba^2$ ‘measure’ (tr) is undoubtedly related to kur^3ba^2 ‘measure’ (noun). In other cases the trimoraic verbs are lexically basic; that is, the final two morae do not correspond to any discernible lexical item presently found in YM. Thus the nouns $ku^1ndu^2u^4$ ‘bush’ and $ti^1ndu^2u^4$ ‘lit firewood’ seem to suggest a stem, $*ndu^2u^4$ or $*ndu^2a^4$, that is non-existent in YM at the present time. These cases may be contrasted to clearly productive derivational processes in which both elements are identifiable: $ku^3+bi^1xi^{32}$ ‘get white hair’ from bi^1xi^{32} ‘white hair’ and the inchoative marker ku^3- .

In contrast to the causative-like semantics of some stem formatives, such as those just discussed, YM has a genuine productive causative prefix, sa^4- , that can be prefixed to both bimoraic and trimoraic verbs (see 13). In some cases the semantic relation between the causative form and a base verb is not clear (13c) as, in this case, the only possible base verb, kwa^2a^3 , is used only in a ritual context with a meaning to the effect of ‘be benefited by’ not clearly related to the meaning, ‘study’, of the causative. In most cases, however, the semantics of the causative derivation is clear.

- (13) a. ka^2a^1 (intr) ‘drown’ > $sa^4-ka^2a^1$ (tr) ‘drown [sb.]’
 ka^2an^1 (intr) ‘talk’ > $sa^4-ka^2an^1$ (tr) ‘make talk’
 ki^2bi^3 (intr) ‘get in’ > $sa^4-ka^1ku^3$ (tr) ‘force in’
 na^3nu^3 (intr) ‘get fat’ > $sa^4-na^3nu^3$ (tr) ‘fatten’
 chi^3chin^4 (intr) ‘suckle’ > $sa^4-chi^3chin^4$ (tr) ‘breast feed’
 b. $nda^3+i^3chi^2$ (intr) ‘dry up’ > $sa^4-nda^3i^3chi^2$ (tr) ‘dry’
 $nda^3+ka^3a^4$ (intr) ‘trip’ > $sa^4-nda^3-ka^3a^4$ (tr) ‘trip [sb.] up’
 $nda^3+ka^2an^4$ (intr) ‘remember’ > $sa^4-nda^3-ka^2an^4$ (tr) ‘remind’
 $tu^1+xu^2u^2$ (intr) ‘get hurt’ > $sa^4-tu^1xu^2u^2$ (tr) ‘harm’
 $ku^1+nda^2a^3$ (intr) ‘move up’ > $sa^4-ku^1nda^2a^3$ (tr) ‘remove’
 c. $sa^4+kwa^2a^3$ (intr) ‘study’

In our analysis of tonal patterns in verbal aspect marking we have excluded any verbs bearing the prefix *sa*⁴-. This responds to the high frequency and productivity of this causative marker. We have also excluded in our sample about a dozen quadrimoraic verbs, such as those in (14).

(14) cha ³ +ka ³ +ndu ² u ⁴	(tr)	‘lay down’
ka ³ +sa ³ +chiu ⁴ un ⁴	(intr)	‘work’
ka ³ +si ³ +kwe ² e ²	(intr)	‘get upset’
ku ³ +ta ³ +nde ³ e ³	(intr)	‘bend down’
ku ³ +ti ³ +sa ² ma ³	(intr)	‘coagulate’
nda ³ +ka ¹ +tu ² un ⁴	(tr)	‘ask’
nda ³ +ki ³ +nde ³ e ⁴	(intr)	‘rest’
nda ³ +xi ³ +ko ² ni ³	(tr)	‘turn around’
ta ³ +xi ³ +kwa ² a ⁴	(tr)	‘keep’

4 Tone and verbal inflection in Yoloxóchtitl Mixtec

The previous sections presented both the wealth and diversity of tonal contrasts in YM as well as its significant functional role. This role includes verbal inflection, the topic of this present section. Tone is not only used for the marking of first-person singular subject but, most importantly, for marking aspect, mood and polarity values, an encoding of grammatical function that is remarkably systematic. We start with a brief description on how tone marks person and then move to the encoding of aspect, mood and polarity.

4.1 Tone and the marking of person

In YM, the subject of verbs and possessor of nouns is encoded by means of enclitics. For the first person, there are four allomorphs, three segmental (=yu¹, =e¹ and =i¹) and one tonal (=²). The distribution of these allomorphs is phonologically conditioned. Enclitics =e¹ and =i¹ are the most restricted; both optionally occur instead of =yu¹ after stems with a tone /1/ on the final mora, the former after stems ending in /a/ or /o/, the latter after stems ending in /u/. For all other cases, enclitic =yu¹ occurs after stems with a final /1/ or /2/ and =² after stems with a final /3/ or /4/.

Tone /2/ is affixed to a stem ending in /3/ or /4/ to mark first person, e.g. ki¹ta³ ‘enter’ > ki¹ta³=² ‘I enter’; ka¹xan⁴ ‘sneeze’ > ka¹xan⁴=² ‘I sneeze’; or ka³ta⁴ ‘feel

itchy' > $ka^3ta^{4=2}$ 'I feel itchy'.¹⁸ Depending on the tonal melody of the stem, the stem-final tone may be elided before the enclitic, creating a surface form contrasting in person, e.g. ku^3xa^3 'mature' > $ku^3xa^{(3)=2}$ > ku^3xa^2 'I mature' or $ndo\gamma^4o^4$ 'suffer' > $ndo\gamma^4o^{(4)=2}$ > $ndo\gamma^4o^2$ 'I suffer'. The general pattern suggests maintenance of stem-final tone when the tonal melody over the final two stem morae is rising (i.e., /1-3/, /1-4/ and /3-4/). Elision of stem-final tone occurs when the tonal melody on the bimoraic stem is level or falling (i.e., /3-3/, /4-3/ and /4-4/).¹⁹ Additionally, syllable structure plays a significant role. The final tone /4/ in a stem with /4-4/ is elided, as expected, if the syllable structure is CVV or CV?V but maintained in disyllabic roots (e.g., $nda^4ta^4=2$ 'I split (something) in two').²⁰

4.2 Tone and the marking of aspect/mood

Verbs in YM inflect for incomplete and complete aspect and for irrealis mood. There is also an inflectional form for the negative of verbs in this mood and, conditionally, in the two aspects. The language has two alternative forms for the complete, which for convenience we call CPL-1 and CPL-2. Two other occasional verb forms, stative and progressive, will not be studied here as they are relatively rare and in many cases irregular in form. Table 5 offers paradigms of four representative verbs.

Tone plays a fundamental role in marking aspect/mood in YM verbs. The complete CPL-1, with the prefix ni^l- , is the only inflected form built by segmental affixation. We take the segments and tonal melody of the CPL-1 stem as representing the basic lemma of the verb. In most cases the tonal melody of the irrealis matches that of CPL-1 although in a few cases (§4.3.3) it is distinct. The following section explores the role of tonal variation in marking aspect and mood.

18 Note that in neighboring Cuanacaxtitlan the use of tone, =2, as a first person enclitic does not occur as an allomorph of = yu^1 adding further evidence that Yoloxóchtitl /2/ is innovative.

19 Note that if the tonal melody is /1-1/ then the enclitic = yu^2 is used to mark the first person.

20 A search through over one hundred hours of transcribed natural speech recordings reveals that other factors may affect the behaviour (maintenance or elision) of stem-final tones before the 1st person enclitic: natural vs. elicited speech, rapidity of utterance, speaker idiosyncracies, and free variation.

Table 5: The verbal paradigm of four exemplary verbs in YM

	‘hang’ (tr)	‘drag’ (tr)	‘break’ (tr)	‘boil’ (intr)
IRR	chi ³ kun ²	ku ³ +ñu ³ u ³	ta ⁷ bi ⁴	kwi ¹ so ¹
NEG.IRR	chi ¹⁴ kun ²	ku ¹⁴ +ñu ³ u ³	ta ⁷ bi ⁴	kwi ¹⁴ so ¹
CPL-1	ni ¹ -chi ³ kun ²	ni ¹ -ju ³ +ñu ³ u ³	ni ¹ -ta ⁷ bi ⁴	ni ¹ -si ¹ so ¹
CPL-2	chi ¹³ kun ²	ju ¹³ +ñu ³ u ³	ta ⁷ bi ⁴	si ¹ so ¹
INCPL	chi ⁴ kun ²	ju ⁴ +ñu ³ u ³	ta ⁷ bi ⁴	si ⁴ so ¹
STAT	ndi ⁴ kun ²	---	ta ⁷ bi ⁴	---
PROG	chi ⁴ +ndi ³ kun ²	ñu ⁴ u ⁴	---	---

4.2.1 Inflectional tone for the CPL-2

The two alternative completive forms (CPL-1 and CPL-2) appear to be in free variation.²¹ While the form for the CPL-1 is built by prefixing *ni¹-* to the base stem, CPL-2 is realized simply by adding a low tone /1/ to the lexical tone on μ 1 of the base. The inflectional tone /1/ of CPL-2 has a lexical origin: the prefix *ni¹-*. When the base has tone /3/ on μ 1, (55% of bimoraic verbs [156 out of 282] and 98% of trimoraic ones [266 out of 272]), the result is an ascending tone /13/ for the CPL-2, as in (15a) and (15b).

- (15)
- | | | | |
|---|--|--------|-----------|
| CPL-1 | CPL-2 | | |
| a. ni ¹ -chi ³ chin ⁴ | chi ¹³ chin ⁴ | (intr) | ‘suckle’ |
| ni ¹ -ka ⁷ an ⁴ | ka ⁷ an ⁴ | (tr) | ‘believe’ |
| b. ni ¹ -chi ³ +nda ⁷ a ⁴ | chi ¹³ +nda ⁷ a ⁴ | (tr) | ‘push’ |
| ni ¹ -ku ³ +i ³ ni ² | ku ¹³ +i ³ ni ² | (tr) | ‘love’ |

The same rule applies to causative verbs with the prefix *sa⁴-*, but for such verbs the CPL-2 that results is homophonous with the negative irrealis, which also bears the inflectional tone /14/ (16).

²¹ This could be taken as a canonical instance of overabundance (Thornton 2011).

(16)	CPL-1	CPL-2	NEG.IRR		
	ni ¹ -sa ⁴ -na ² a ¹	sa ¹⁴ -na ² a ¹	sa ¹⁴ -na ² a ¹	(tr)	‘teach’
	ni-sa ⁴ -ka ³ sun ²	sa ¹⁴ -ka ³ sun ²	sa ¹⁴ -ka ³ sun ²	(tr)	‘fry’

Completive tone /1/ may be present, though not overtly manifested, on stems with lexical tones /1/ or /14/ on μ 1. As a result, the tonal melodies of the CPL-2 and the irrealis of verbs with /1/ or /14/ on μ 1 are homophonous (17). Nevertheless, a significant number of verbs have irrealis forms whose segments are distinct from those of the lexical stem (the stem of the CPL-2) (see §4.3). In cases of lexical tones /1/ or /14/ on μ 1, this segmental variation maintains the formal distinction between the irrealis and CPL-2 despite the fact that the tonal melodies for the two forms are identical.

(17)	CPL-1	CPL-2	IRR		
	ni ¹ -ka ¹ ba ¹	ka ¹ ba ¹	ka ¹ ba ¹	(intr)	‘turn’
	ni ¹ -tu ¹ +xu ² u ²	tu ¹ +xu ² u ²	tu ¹ +xu ² u ²	(intr)	‘get hurt’
	ni ¹ -xi ¹⁴ ta ³	xi ¹⁴ ta ³	xi ¹⁴ ta ³	(tr)	‘pull’
	ni ¹ -si ¹⁴ +kwe ¹ kun ¹	si ¹⁴ +kwe ¹ kun ¹	si ¹⁴ +kwe ¹ kun ¹	(tr)	‘spread a disease to (sb)’

4.2.2 Inflectional tone for the incomplete

In the default situation, the form for the incomplete aspect is built by overwriting the lexical tone on μ 1 with high tone /4/. In (18) instances of CPL-1 are given alongside the incomplete to provide a reference for lexical tone of the stem.

(18)	CPL-1	INCPL		
a.	ni ¹ -chi ³ chin ⁴	chi ⁴ chin ⁴	(intr)	‘suckle’
	ni ¹ -ka ² an ⁴	ka ² an ⁴	(tr)	‘believe’
	ni ¹ -ku ³ +i ³ ni ²	ku ⁴ +i ³ ni ²	(tr)	‘love’
	ni ¹ -ka ³ +xi ⁴ ta ³	ka ⁴ +xi ⁴ ta ³	(tr)	‘grind’
b.	ni ¹ -ka ² an ¹	ka ² an ¹	(intr)	‘talk’
	ni ¹ -nda ² yu ¹	nda ² yu ¹	(intr)	‘confess’
	ni ¹ -tu ¹ +xu ² u ²	tu ⁴ +xu ² u ²	(tr)	‘get hurt’

As causatives already have a tone /4/ on μ_1 , which is lexically associated with the prefix sa^4 -, the incomplete is indistinguishable in surface form from the irrealis (19).

(19) IRR	INCPL		
sa^4 -ndo ³ to ³	sa^4 -ndo ³ to ³	(tr)	‘wake up’
sa^4 -ndu ¹ xin ¹	sa^4 -ndu ¹ xin ¹	(tr)	‘bury’

With level tones in a bimoraic stem, the formation of the incomplete is quite regular, although in some cases there is a pattern split depending on the syllabic structure of the verbal stem. This is demonstrated by the examples in Table 6. Note that there are no YM words with /3-1/ tones.

Table 6: Tone allomorphy for the incomplete

	Syllable structure	LEX	INCPL	CPL-1	INCPL	
a	disyllabic	/1-1/ > /4-1/	ni^1 - ki^1 xin ¹	ki^4 xin ¹		‘fall asleep’
	monosyllabic		ni^1 -tu ¹ un ¹	tu ⁴ un ¹		‘catch fire, light up’
b	disyllabic	/1-3/ > /4-13/	ni^1 -ka ¹ ku ³	ka ⁴ ku ¹³		‘escape’
	monosyllabic		ni^1 -ka ¹ an ³	ka ⁴ an ¹³		‘get accustomed’
c	disyllabic	/1-4/ > /4-14/	ni^1 -ka ¹ xan ⁴	ka ⁴ xan ¹⁴		‘sneeze’
	monosyllabic		ni^1 -ku ¹ un ⁴	ku ⁴ un ¹⁴		‘for chilli to be ground into sauce’
d.1	disyllabic	/3-3/ > /4-3/	ni^1 -ka ³ ba ³	ka ⁴ ba ³		‘lie down to sleep’
			ni^1 -nda ³ ba ³	nda ⁴ ba ³		‘fall’
d.2	monosyllabic	/3-3/ > /4-4/	ni^1 -chi ³ i ³	chi ⁴ i ⁴		‘get wet’
			ni^1 -ka ³ a ³	ka ⁴ a ⁴		‘emit a sound’
			ni^1 -ku ³ u ³	ku ⁴ u ⁴		‘occur, happen’
			ni^1 -chi ³ i ³	chi ⁴ i ⁴		‘harvest’ (tr)
e.1	disyllabic	/3-4/ > /4-4/	ni^1 -ku ³ chi ⁴	ku ⁴ chi ⁴		‘feel sad’ ²²
			ni^1 -ka ³ ba ⁴	ka ⁴ ba ⁴		‘turn’ (tr.)
e.2	monosyllabic	/3-4/ > /4-24/	ni^1 -ka ³ a ⁴	ka ⁴ a ²⁴		‘slip’
			ni^1 -ku ³ un ⁴	ku ⁴ un ²⁴		‘grind [chilli] for sauce’

²² In collocation with the noun i^3ni^2 ‘heart’.

The preceding examples reveal interesting, though regular, patterns. If the lexical tone has /1/ on μ_1 then /4/ is assigned to μ_1 and the lexical tone /1/ on μ_1 is assigned to μ_2 . This can be seen most clearly in (b) and (c) whose outcomes result in a rising tone /13/ or /14/ on μ_2 . When the lexical tone of μ_2 is also /1/, it remains unchanged in the incomplete (a).

If the lexical tone of μ_1 is /3/ then a split pattern develops depending on the syllabic structure of the stem. With disyllabic stems, the incomplete is formed by simply assigning /4/ to μ_1 ; there is no change in the lexical tone of μ_2 (d.1/e.1). If the stem is monosyllabic (CVV or CV?V), then for lexical melodies /3-3/ the incomplete is /4-4/, an outcome that can be accounted for by positing a single lexical tone /3/ that spreads to both morae in the lexical base. The same spreading would occur when the single lexical tone is overwritten by inflectional tone /4/. If the lexical tonal melody is /3-4/ for disyllabic verbs the tone of μ_1 is simply raised to /4/. In monosyllables, however, the mid tone /3/ on μ_1 is pushed to μ_2 and lowered to /2/ in the context of surrounding high tones (a rising /34/ tone is not found in YM) resulting in a /4-24/ tonal melody.

A final interesting incomplete formation involves the few irrealis stems that have a lexical tone /14/ on μ_1 (for a complete list, see Table 7 below). Such stems require the incomplete prefix i^4 - if the verb is bimoraic, as shown in (20a) and Table 7.²³ But with trimoraic verbs, tone /4/ overwrites the lexical tone, as shown in (20b). Like other inflectional tones, tone /4/ for the incomplete is the reflex of a segmental form, in this case the archaic incomplete prefix i^4 -, only observable nowadays in this restricted prosodic context of initial /14/ in verbal lexical stems.

(20)	CPL-1	INCPL		
a.	$ni^1\text{-}ki^{14}tu^3$	$i^4\text{-}ki^{14}tu^3$	(intr)	'dawn'
	$ni^1\text{-}ta^{14}ni^3$	$i^4\text{-}ta^{14}ni^3$	(intr)	'get bruised' (fruit)
	$ni^1\text{-}ti^{14}bi^3$	$i^4\text{-}ti^{14}bi^3$	(intr)	'rotten'
	$ni^1\text{-}chu^{14}tu^2$	$i^4\text{-}chu^{14}tu^2$	(tr)	'kiss'
b.	$ni^1\text{-}xo^{14}\text{+}kwi^1in^1$	$xo^4\text{+}kwi^1in^1$	(intr)	'turn around to look'
	$ni^1\text{-}si^{14}\text{+}kwe^1kun^1$	$si^4\text{+}kwe^1kun^1$	(tr)	'spread a disease to (someone)'

²³ One verb has a suppletive incomplete form that begins with /i/: LEX.STEM $jo^3nde^3e^3$ 'sit down' (PL subj.) > INCPL $i^4nde^4e^4$. Three other bimoraic verbs have an initial /i/ in their lexical stems, none of which have an initial /14/ tone: LEX.STEM $i^3?in^3$ 'scratch' (INCPL $i^2?in^3$); LEX.STEM i^1chi^1 'dry up (something alive, such as a plant)' (INCPL i^4chi^1); LEX.STEM i^3xa^3 'do' (INCPL i^4xa^3).

This incomplete prefix i^4 - in YM provides an example of the retention of a historical form of a bound prefix to mark the continuative, a morpheme discussed by McKendry (2013:80):²⁴

The imperfective prefix provides an example of the process whereby CV segments are lost, but the tones remain as floating tones. In most varieties the difference between the irrealis and the imperfective forms of verbs is the presence of a High tone in the imperfective. Hollenbach (2001) posits that this floating High tone is all that remains of a verbal morpheme which up to colonial times had CV segments.

McKendry (2013:80) goes on to say that the historical prefix form is *yo*, attested in Alvarado (1962 [1593]) as in the following example:

- (21) **yo**-sasi=ndi
 IPFV-eat=I
 ‘I’m eating.’

Finally, there are at least two verbs in our sample that have an irregular tonal outcome in the incomplete (22a). A very small number of irregular verbs (cf. 22b) have the unexpected historical incomplete marker i^4 - despite lacking the initial /14/ lexical tonal melody that otherwise motivates the use of this prefix.

(22)	CPL-1	INCPL	Expected		
a.	$ni^1ka^?^1an^1$	$ka^?^4an^2$	$*ka^?^4an^1$	(intr)	‘talk’
	$ni^1sa^1a^4$	sa^4a^{24}	$*sa^4a^{14}$	(intr)	‘get upset’ ²⁵
b.	$ni^1jo^3nde^3e^3$	$(i^4-)nde^4e^4$	$*jo^4nde^3e^3$	(intr)	‘sit down’ (PL sub.)

4.2.3 Inflectional tone for the negative: Irrealis, incomplete, completive

Negation in YM can be marked inflectionally by means of tone on verbs in the irrealis, completive and, less commonly, the incomplete. In cases in which

²⁴ The same incomplete prefix is also found Juxtlahuaca Mixtec (Ramírez and Beatham, 2012:4), e.g. *i-ndú’ú=ra* ‘he’s sitting down.’ We are indebted to Lucien Carroll for bringing McKendry’s observation and the Juxtlahuaca examples to our attention.

²⁵ There are eight monosyllabic verbs (seven intransitives and one transitive) that have /1-4/ lexical tone. All but sa^1a^4 (‘get upset’) form the incomplete with /4-14/, as expected. However, there are two verbs sa^1a^4 (‘get heated up’ [e.g., a boiled liquid] and ‘get upset’) that are homophonous in the irrealis though distinct in the incomplete: sa^4a^{14} ‘get heated up’ [a liquid] and sa^4a^{24} ‘get upset’. The two verbs would seem to be related as distinct senses of the same lemma. Whether the irregular incomplete tonal melody of one sense (sa^4a^{24} ‘get upset’) can be ascribed to a disambiguating function is, however, unclear.

negation cannot be marked by tone a negative adverb is used. With the incomplete, at times both tonal and syntactic marking of negation can occur with the same verb stem (Examples in (25) below).

(23)	IRR	NEG.IRR		
a.	choʔ ³ ma ⁴	cho ¹⁴ ma ⁴	(tr)	‘squash’
	kaʔ ¹ an ¹	kaʔ ¹⁴ an ¹	(intr)	‘talk’
b.	ka ³ +xi ⁴ ta ³	ka ¹⁴ +xi ⁴ ta ³	(tr)	‘grind’
	tu ¹ +xuʔ ⁴ u ²	tu ¹⁴ +xuʔ ¹⁴ u ²	(intr)	‘get hurt’
c.	sa ⁴ -naʔ ¹ a ¹	sa ¹⁴ -naʔ ¹ a ¹	(tr)	‘teach’
	sa ⁴ -ka ³ sun ²	sa ¹⁴ -ka ³ sun ²	(tr)	‘fry’

Negation of the irrealis and completive through tone is shown in (23) and (24). For the irrealis, negation is realized by overwriting the lexical tone on $\mu 1$, be it /1/ or /3/, by an ascending tone /14/. This occurs regardless of whether the verb has a bimoraic stem as in (23a) or a trimoraic one, as in (23b), or even if the verb is a causative derived verb with the prefix *sa*⁴-, as in (23c).

Note that despite the fact that in causative constructions with *sa*⁴- the irrealis and incomplete are formally indistinct, negation through a rising tone is limited to the irrealis (24) whereas the negative of the incomplete is marked by the adverb *ba*¹⁴³.

(24)	IRR	NEG.IRR	INCPL	NEG.INCPL	
	sa ⁴ -chi ³ i ³	sa ¹⁴ -chi ³ i ³	sa ⁴ -chi ³ i ³	ba ¹⁴³ sa ⁴ -chi ³ i ³	(intr) ‘make wet’

In the causative, then, the irrealis and incomplete utilize distinct strategies to mark the negative: tonal inflection for the irrealis, a syntactic structure for the incomplete. This distinction suggests that despite the exceptions and irregularities noted in Tables 7 and 8, the prototypical markers of negation are tonal for the irrealis and segmental for the incomplete.

Negation of completive aspect only occurs with CPL-1 form, in which the completive aspect is marked by the prefix *ni*¹-. The tonally marked CPL-2 form cannot be negated, either through tone or by a preceding lexical element. For CPL-1, negation is realized in one of two ways. The most common (1385 of a total of 1574 corpus occurrences; 88 percent) is with *ni*¹⁴-, applying the /14/ tonal marker of negation to the prefix. An alternative form is syntactic, using the element *ba*¹⁴³ (189 cases in the corpus; 12 percent). After *ba*¹⁴³ the tone of the completive prefix *ni*¹- is raised to *ni*⁴-.

(25) CPL-1	NEG.CPL-1	NEG.CPL-1	
	WITH TONAL MARKING	WITH SYNTACTIC MARKING AND RAISED TONE ON NI ¹ -	
ni ¹ -na ɾ ¹ ma ¹	ni ¹⁴ -na ɾ ¹ ma ¹	ba ¹⁴³ ni ⁴ -na ɾ ¹ ma ¹	(intr) ‘confess’
ni ¹ -su ³ ma ³	ni ¹⁴ -su ³ ma ³	ba ¹⁴³ ni ⁴ -su ³ ma ³	(intr) ‘go backwards’
ni ¹ -ka ³ chi ²	ni ¹⁴ -ka ³ chi ²	ba ¹⁴³ ni ⁴ -ka ³ chi ²	(tr) ‘say’

Table 7: The negative of irrealis stems with /i4/ on μ1

IRR	NEG.IRR	Alternative IRR	INCPL
ki ¹⁴ tu ³ / ni ¹ ki ⁴ tu ³	kwa ¹⁴ ki ¹⁴ tu ³	----	i ⁴ -nki ¹⁴ tu ³ ‘dawn’
chu ^{ɾ14} ma ¹	kwa ¹⁴ chu ^{ɾ14} ma ¹	----	i ⁴ -chu ^{ɾ14} ma ¹ ‘be censured (with incense)’
ta ¹⁴ ni ³	kwa ¹⁴ ta ¹⁴ ni ³	----	i ⁴ -ta ¹⁴ ni ³ ‘get bruised’ (fruit)
ti ¹⁴ bi ³	kwa ¹⁴ ti ¹⁴ bi ³	----	i ⁴ -ti ¹⁴ bi ³ ‘get broken’ (a machine)
xa ¹⁴ bi ²	kwa ¹⁴ xa ¹⁴ bi ²	----	i ⁴ -xa ¹⁴ bi ² ‘get tired’
xa ¹⁴ ni ²	kwa ¹⁴ xa ¹⁴ ni ²	----	i ⁴ -xa ¹⁴ ni ² ‘dream’
xa ¹⁴ ta ²	kwa ¹⁴ xa ¹⁴ ta ²	----	i ⁴ -xa ¹⁴ ta ² ‘shave’ (sb’s head)
xi ¹⁴ nda ²	kwa ¹⁴ xi ¹⁴ nda ²	----	i ⁴ -xi ¹⁴ nda ² ‘carve’
xi ^{ɾ14} ni ³	kwa ¹⁴ xi ^{ɾ14} ni ³	----	i ⁴ -xi ^{ɾ14} ni ³ ‘rub’
xi ^{ɾ14} n̄u ³	kwa ¹⁴ xi ^{ɾ14} n̄u ³	----	i ⁴ -xi ^{ɾ14} n̄u ³ ‘patch up’
xio ^{ɾ14} o ⁴	kwa ¹⁴ xio ^{ɾ14} o ⁴	----	i ⁴ -xio ^{ɾ14} o ⁴ ‘become ill after having craved a food’
cha ^{ɾ14} bi ³ / cha ^{ɾ4} bi ³	kwa ¹⁴ cha ^{ɾ14} bi ³ kwa ¹⁴ cha ^{ɾ4} bi ³	cha ^{ɾ4} bi ³	i ⁴ -cha ^{ɾ14} bi ³ ‘pay’ i ⁴ -cha ^{ɾ4} bi ³
chu ¹⁴ tu ² / chu ⁴ tu ²	kwa ¹⁴ chu ¹⁴ tu ² kwa ¹⁴ chu ⁴ tu ²	chu ⁴ tu ²	i ⁴ -chu ¹⁴ tu ² ‘kiss’
xi ¹⁴ ko ³	kwa ¹⁴ xi ¹⁴ ko ³ kwa ¹⁴ xi ⁴ ko ³	xi ⁴ ko ³	i ⁴ -xi ¹⁴ ko ³ ‘sell’
xi ¹⁴ ta ³	kwa ¹⁴ xi ¹⁴ ta ³ kwa ¹⁴ xi ⁴ ta ³	xi ⁴ ta ³	i ⁴ -xi ¹⁴ ta ³ ‘pull’ i ⁴ -xi ⁴ ta ³
ya ¹⁴ kun ²	kwa ¹⁴ ya ¹⁴ kun ² kwa ¹⁴ ya ⁴ kun ²	ya ⁴ kun ²	i ⁴ -ya ¹⁴ kun ² ‘massage’ i ⁴ -ya ⁴ kun ²
yu ^{ɾ14} bi ²	kwa ¹⁴ yu ^{ɾ14} bi ² kwa ¹⁴ yu ^{ɾ3} bi ²	yu ^{ɾ3} bi ²	i ⁴ -yu ^{ɾ14} bi ² ‘get scared’

For irrealis stem verbs with lexical tone /14/ on $\mu 1$, however, negation is marked syntactically by means of the negative adverb *kwa*¹⁴, as shown in Table 7. Note in Table 7 that six verbs with initial /14/ have an alternate negative form. The alternative irrealis forms *xi*⁴*ko*², *xi*⁴*ta*³, *ya*⁴*kun*² and *yu*³*bi*² apparently can occur only after the negative marker *kwa*¹⁴ (i.e., the affirmative irrealis must have tone /14/ on $\mu 1$). However, two irregular forms, *cha*²*bi*³ and *chu*⁴*tu*², can be used in the affirmative, that is, even if not preceded by the negative marker *kwa*¹⁴. As pointed out in (20a) above, the verbs in Table 7 are those whose incomplete form is marked by the prefix *i*⁴-. It is to facilitate the visualization of this pattern that we include their incomplete forms there.²⁶

The negative adverb *kwa*¹⁴ in YM is used only before the irrealis. Comparative evidence from other Mixtec languages strongly suggests that this marker is the reflex of a historical irrealis marker that is still observable in other varieties as an optional or contextually conditioned marker, with various degrees of phonological independence: *kwa*¹ in San Juan Colorado (Stark Campbell et al. 1986:163) (see also Table 13 below), *ku* in Jamiltepec (Johnson 1988:102), *kūn* in Ocotepéc (Alexander 1988:251), or the prefixes *kw-* or *kV-* in Coatzacoapan (Small 1990:398). Both the independent word and the fused prefix are undoubtedly related to the segmental alternations discussed in §4.3 and go back to the historical prefixes **ka-* and **ku-* reconstructed by Kaufman (cited in Macaulay 1996:48) as the allomorphs for the irrealis in Proto-Mixtec.

The other principle independent negative marker in YM is *ba*¹⁴³, which is most commonly used to negate a verb in the incomplete, as in (26).²⁷

- (26) *ba*¹⁴³ *ka*²*an*²=*yu*¹
 NEG INCPL.SPEAK=1SG
 'I do not speak.'

A few incomplete verbs, however, can also accept tonal marking (/14/) of negation. The incomplete verbs that have been identified as accepting tonal as well as segmental marking of negation all have an incomplete stem that is segmentally different from that of the irrealis (for more details on segmental alternations in verb stems, see the following section). Table 8 offers data on the frequency of

²⁶ Note the verb *i*⁴-*nki*¹⁴*tu*³, which is based on the contracted trimoraic stem *ni*¹*ki*⁴*tu*³, an alternate of bimoraic stem *ki*¹⁴*tu*³ 'dawn'. This suggests that an initial /14/ tone in bimoraic stems such as the ones in Table 7 might be historically related to a conflation of two independent level tones /1/ and /4/ realized over the first two morae of historically prior trimoraic verbs.

²⁷ Notice that both *ba*¹⁴³ and the negative irrealis marker *kwa*¹⁴ have the same tone /14/ characteristic of other negative forms in YM.

this alternative marking for four verbal lemmas of this type based on Amith and Castillo García's large corpus of YM texts.²⁸ Notice that the first and fourth verbs have alternative stems for the incomplete at a lexical level with no difference in meaning. The use of the negative form of the incomplete varies greatly depending on the lexical item. For example, for the verb *ku³ni³* 'know', tonal marking of the negative incomplete is found 1,394 times, whereas the syntactic alternative by means of *ba¹⁴³* is only found only 22 times. Both options, inflectional vs. syntactic negation, are used with equal frequency for the verb *kwi³in³* 'accept', but for the verbs *kwa²a³* 'permit' or *ko³to³* 'see', the tonal encoding is less favoured.

Table 8: Incomplete verbs with negation marked by tone /14/

IRR	INCPL	Corpus occurrences of affirmative INCPL	NEG.INCPL	Corpus occurrences of	
				total NEG.INCPL	<i>ba¹⁴³</i>
<i>ku³ni³</i> 'know'	<i>ji⁴ni²</i>	3,483	<i>ji¹⁴ni²</i>	1,394	22
	<i>xi⁴ni²</i>	298	<i>xi¹⁴ni²</i>	261	
	<i>ji⁴ni³</i>	119	<i>ji¹⁴ni³</i>	3	
	<i>xi⁴ni³</i>	45	<i>xi¹⁴ni³</i>	1	
<i>kwi³in³</i> 'accept'	<i>xi⁴in⁴</i>	81	<i>xi¹⁴in⁴</i>	59	60
<i>kwa²a³</i> 'permit'	<i>xa²a¹³</i>	175	<i>xa²a¹³</i>	24	86
<i>ko³to³</i> 'see'	<i>xi⁴to³</i>	146	<i>xi¹⁴to³</i>	2	6
	<i>ji⁴to³</i>	419	<i>ji¹⁴to³</i>	1	11

In this section we have examined the use of tone in YM as an inflectional exponent of aspect, mood and polarity. Inflectional tones in this system work as prosodic affixes with a remarkable consistency in the form-meaning mapping of inflection. Before we compare the behaviour of tone in YM to that found in other Mixtec languages, in the next section we introduce other relevant aspects of the inflectional morphology of YM that involve segmental alternations on the stem.

²⁸ These four verbs are not the only ones that accept tonal marking of negation in the incomplete. They are simply those for which such tonal marking is most frequently found in the corpus. Although not documented in the corpus, according to the judgement of Castillo García other verbs also accept tonal marking of negation in the incomplete: *ka³xi⁴* IRR, *xa⁴xi²⁴* INCPL, *xa¹⁴xi²⁴* NEG INCPL 'eat', 'bite into'; and *xi²⁴in¹³* INCPL, *xi²⁴in¹³* NEG INCPL 'leak' (a house roof) (this verb does not have an irrealis form).

4.3 Stem alternation patterns in verbal morphology

In addition to tone, segmental stem alternations also play an important role in YM verbal inflection. Most verbs, here called "invariant verbs", maintain segments constant throughout their paradigms, as for example the verbs *ku³ki⁷⁴bi³* 'be a nuisance' and *xi³kwi⁷⁴na⁴* 'steal' in Table 9. Other verbs, here called "variant verbs", may have two segmentally distinct stems: *ku³ka⁷⁴an³* 'be ashamed', for example, has one stem for the irrealis mood (and for the negative irrealis) and another for the incomplete and complete aspects. The same stem formatives (e.g., **ku-* in Table 9) may be found with either invariant or variant verbs. This type of segmental variation is very common in YM and in Mixtec languages in general. But it is still poorly understood and further study is needed to understand any possible conditioning factors.

Table 9: Invariant and variant verbs in YM

	Invariant verbs	Variant verbs	
	(intr) 'be a nuisance'	(tr) 'steal'	(intr) 'be ashamed'
IRR	<i>ku³+ki⁷⁴bi³</i>	<i>xi³+kwi⁷⁴na⁴</i>	<i>ku³+ka⁷⁴an³</i>
NEG.IRR	<i>ku¹⁴+ki⁷⁴bi³</i>	<i>xi¹⁴+kwi⁷⁴na⁴</i>	<i>ku¹⁴+ka⁷⁴an³</i>
CPL-1	<i>ni¹-ku³+ki⁷⁴bi³</i>	<i>ni¹-xi³+kwi⁷⁴na⁴</i>	<i>ni¹-xi³+ka⁷⁴an³</i>
CPL-2	<i>ku¹³+ki⁷⁴bi³</i>	<i>xi¹³+kwi⁷⁴na⁴</i>	<i>xi¹³+ka⁷⁴an³</i>
INCPL	<i>ku⁴+ki⁷⁴bi³</i>	<i>xi⁴+kwi⁷⁴na⁴</i>	<i>xi⁴+ka⁷⁴an³</i>

Most YM verbs are invariant: 89 percent of our sample verbs are of this type (495 of 554).²⁹ The remaining 109 are variant. Such verbs appear to be remnants of an older system in which the stem formatives involved in the building of the irrealis stem were once irrealis markers of some sort. Macaulay (1996:48) makes this point and cites a manuscript by Terrence Kaufman in which he reconstructs the proto-Mixtec prefixes **xi-* (durative), **ka-* (irrealis) and **ku-* (also irrealis) as evidence for the historical basis of the *xi-/ku-* alternation noted in many Mixtec languages. These same archaic stem formatives are evident in YM. At some point this situation became morphologically opaque, giving rise to the invariant~variant verb patterns now in evidence. For some verbs the irrealis stem generalized to

²⁹ As pointed out above, our sample does not include any causative verb with the prefix *sa⁴-*.

the entire paradigm (e.g., $ku^3ki^?^4bi^3$ ‘be a nuisance’) whereas for other verbs it was the stem bearing Kaufman’s ‘durative’ **xi-* that generalized (e.g., $xi^3kwi^?^4na^4$ ‘steal’).

Note that in the case of *invariant verbs* with lexical tone /1/ or /14/ on μ_1 , the irrealis and CPL-2 forms are indistinct (Table 10, first and third columns, rows 1 and 3). This homophony is avoided, however, with *variant verbs* as in these cases the irrealis stem is not identical to the lexical stem of the incomplete and complete (Table 10, second column):

Table 10: Homophony between irrealis and CPL-2

	Invariant (tr) ‘sew’	Variant (tr) ‘ask for’	Initial /14/ (intr) ‘dream’
IRR	ki^1ku^3	ka^1kan^1	$xa^{14}ni^2$
CPL-2	ki^1ku^3	xi^1kan^1	$xa^{14}ni^2$
INCPL	ki^4ku^{13}	xi^4kan^1	$i^4-xa^{14}ni^2$

Of the 109 variant verbs in our sample, 101 (93%) fall into two inflectional classes attending to the proposed stem formative of the irrealis: *ku+* (82 cases) or *ka+* (19 cases). The remaining eight verbs are irregular to some degree and are treated separately. The distribution of the classes in our sample is given in (27).

(27)		<i>ku+</i>	<i>ka+</i>	TOTAL
Bimoraic	intr	29	3	
	tr	13	10	
Trimoraic	intr	35	6	
	tr	5	0	
TOTAL		82	19	101

The following sections analyze the patterns of variation manifested in the 109 variant verbs of our sample.

4.3.1 The ku+ stem class

Eighty-two verbs in the sample have a stem change to initial *ku+* in the irrealis, e.g. *chi+chi* /1-3/ ‘ripen, mature’ > IRR.STEM *ku+chi*. We refer to this variable segmental portion of the stems as the stem formative (indicated by the plus sign +). The unstressed /u/ in the formative *ku+* surfaces as [o] in conditions that are not entirely clear; there may be some free variation involved. Note also that *ku+* is labialized to [kw]³⁰ when preceding a non-back initial stem vowel (e.g. *ch+achi* /3-3/ ‘be torn apart’ > IRR.STEM *ku+achi* [k^wachi]). Representative examples of the *ku+* stem class are given in (28). Verbs with variable lexical stems are illustrated in (28b). In a few cases the irrealis is built by adding *ku+* to a verb lacking a stem formative in the lexical form (28c). As the tone melody is independent from segmental changes, we have disassociated the segments from the tone structure.

(28) The *ku+* class

	LEX.STEM	IRR.STEM	TONE	
a.	<i>cha+achi</i>	<i>ku+achi</i>	/3-3/	(intr) ‘be torn apart’
	<i>chi+chi</i>	<i>ku+chi</i>	/1-3/	(intr) ‘ripen, mature’
	<i>ja+nuʔu</i>	<i>ku+nuʔu</i>	/1-1-4/	(intr) ‘go back to one’s place’ ³¹
	<i>ji+ni</i>	<i>ku+ni</i>	/3-2/	(tr) ‘see’
	<i>jo+ndaa</i>	<i>ko+ndaa</i>	/3-2-2/	(tr) ‘look after’
	<i>ju+ndaʔa</i>	<i>ku+ndaʔa</i>	/1-1-3/	(intr) ‘remove oneself (from a place)’
	<i>ju+eta</i>	<i>ku+eta</i>	/3-3/	(intr) ‘be measured’
	<i>ki+xin</i>	<i>ku+sun</i>	/1-1/	(intr) ‘sleep’
	<i>si+iso</i>	<i>ku+iso</i>	/1-1/	(intr) ‘come to a boil’
	<i>xa+aʔa</i>	<i>ku+aʔa</i>	/1-3/	(dtr) ‘give’
	<i>xi+iin</i>	<i>ku+iin</i>	/1-1/	(intr) ‘copulate’
	<i>xi+iko</i>	<i>ku+iko</i>	/3-4/	(tr) ‘carry’
	<i>xi+kaʔan</i>	<i>ku+kaʔan</i>	/3-3-3/	(intr) ‘be ashamed’
	<i>xu+xa</i>	<i>ku+xa</i>	/3-3/	(intr) ‘become an adult’
b.	<i>xi+to/ji+to</i>	<i>ko+to</i>	/3-3/	(intr) ‘look’
	<i>ju+ndeta/</i> <i>ndu+ndeta</i>	<i>ku+ndeta</i>	/3-3-3/	(intr) ‘stand up’ (PL sub.)
c.	<i>ndoʔo</i>	<i>ko+ndoʔo</i>	/3-3/	(intr) ‘be the victim of disfortune’
	<i>nani/ju+nani</i>	<i>ku+nani</i>	/(3-)3-2/	(intr) ‘become resolved (a problem)’

³⁰ Labialization and palatalization commonly occur before enclitics: /Cu/, and sometimes /Co/, is labialized before non-back vowels and /Ci/ is palatalized before non-front vowels.

³¹ Alternative stem: *kwa¹nu²u⁴*.

4.3.2 The *ka+* stem class

Nineteen verbs have an initial *ka+* sequence for the irrealis that is not present in the lexical stem (incompletive). For example, *xa+ta* /3-3/ ‘dig out’ is *ka+ta* in the irrealis. Representative examples of *ka+* stem class verbs are given in (29).

(29) The *ka+* class

	LEX.STEM	IRR.STEM			
a.	sa+si	ka+si	/3-2/	(tr)	‘look after cattle’
	xa+ʔnu	ka+ʔnu	/1-1/	(tr)	‘split in two’
	xi+kan	ka+kan	/1-1/	(tr)	‘ask for’
	ja+nduʔu	ka+nduʔu	/3-4-4/	(intr)	‘lie down’
b.	xa+ta/ja+ta	ka+ta	/3-3/	(tr)	‘dig out’
	xa+kwiin/ju+kwiin	ka+kwiin	/3-1-3/	(intr)	‘be left hanging’
	xi+ta/ji+ta	ka+ta	/3-3/	(intr)	‘sing’

As previously noted, we treat the segmental stem of the completive (CPL-1) as the lexical stem. It has a wider distribution in the aspect paradigm, being found in the incompletive as well as the completive. Additionally, its form is more diverse and unpredictable. A number of verbs may have alternative lexical stems, almost invariably manifesting a *x/j*-type of alternation in the historical prefix, such as the one illustrated by the verbs in (29b).

In general, there is little in the lexical stem that indicates with certainty that the verb is a member of the variant class although the phonology of the stem does provide some clues. For example, no invariant verb has a lexical stem with an onset /j/, so verbs such as *ju³+nda³tu³* ‘wait’, *ju³+ndi³xi³* ‘dress’, *ja³+ndu⁴u⁴* ‘lie down’ and *ja³ta²* ‘graze’ can be assumed to be variant, with a different onset segment in the irrealis. Furthermore, in most cases verbs beginning with *ju+* belong to the *ku+* class, and with *ja+* to the *ka+* class. The initial sequences of irrealis stems are less revealing: irrealis stems with /ka/ or /ku/ onsets may or may not be variant: 224 regular verbs in the sample have irrealis stems with /ka/ or /ku/ onsets and only 101 of these (43%) are variant. Of the 123 remaining invariant verbs, 59 have a /ka/ onset and 64 a /ku/ onset.³²

³² Like the verbs of the *ku+* class, most of the invariant verbs with an onset in /ku/ are intransitive (50 out of 59). The transitivity profile of verbs having /ka/ is much more mixed.

4.3.3 Irregular verbs

Finally, eight verbs of the 109 variant verbs are irregular, although most of them could be alternatively analyzed as members of the *ku+* class. The verbs listed in (30) have suppletive irrealis stems that at times also bear tonal melodies that are distinct from those of the lexical stem.

(30)	LEX. STEM	IRR		
a.	i^3+xa^3	a^3sa^3	(tr)	‘do’ ³³
	$ju^3+na^3ni^2$	na^3ni^2	(intr)	‘be able to reach out to grab (something)’
	$xi+\tau^3i^3$	$ko+\tau^3o^3$	(tr)	‘drink’
	$xa+\tau^1an^1$	$ku+\tau^1un^1$	(intr)	‘go’
b.	xi^1xi^1	ko^3ko^3	(intr)	‘get burnt’
	xi^1yo^3	ko^3o^3	(tr)	‘be (existential)’
	$xi+\tau^1i^1$	ku^3u^2	(intr)	‘die’
	xi^1ka^3	ka^3ka^3	(intr)	‘walk’

To this point we have presented and analyzed tonal and segmental variation in the marking of inflection in YM verb. In the following section, we conclude this article by comparing YM verbal inflection with patterns found in two other languages within the Mixtec family for which comparable data exists.

5 Conclusion: Yoloxóchitl verbal inflection in a Mixtec context

This paper has discussed tonal and segmental variation in the inflection of verbs in YM. We have shown that in this language tones work as fully fledged inflectional exponents. That is, YM manifests a system of verbal inflection whereby a given tone X is assigned a certain morphosyntactic value Y. The system has little tonal allomorphy overall, and when we find it, especially in the building of the incomplete, the different tonal structures can be accounted for as conditioned by morphophonological properties of the stems. In other words, one does not need to appeal to information encoded in the lexicon to account for the selection of tonal alternants.

³³ The same verb apparently has an alternative lexical stem *ki+xa* /3-3/ and irrealis stem *ka+sa* /3-3/, in the *ka+* class, although these forms now mean (intr) ‘copulate’ in YM.

We conclude this article by placing the inflectional system of YM within the context of Mixtec languages in general in order to suggest how representative the YM system is considering Mixtec as a language family. The answer is not straightforward. By comparing the YM system to that found in Xochapan Mixtec, we show that both systems share characteristics that we suggest are features of the Guerrero subgroup of Mixtec languages. A very brief excursion into another more distantly related Mixtec language for which sizeable lexical samples exist, San Juan Colorado Mixtec, reveals that outside the Guerrero subgroup of Mixtec languages the inflectional system appears to be very different. At this stage, however, we can only suggest possible patterns of tonal and segmental morphology to explore in analyzing Mixtec dialectology.

Xochapan Mixtec (henceforth XM) is a member of the Guerrero subgroup and thus closely related to YM, also a Guerrero Mixtec language. Xochapan Mixtec is spoken by approximately 8,000 people in Xochapa, a village in the municipality of Alcozauca about 60 kilometers north-northeast of Yoloxóchitl. In general, XM bears a strong resemblance to YM though differing in some interesting ways. The dictionary of XM by Stark, Johnson, and González de Guzmán (2006) includes 111 verbs with bimoraic lexical stems. For each the authors give three inflected forms: the irrealis (called ‘future’), the completive (called ‘past’) and the incompletive (called ‘present’). Xochapa has four level tones. To facilitate comparison we have converted the orthography of the original, which uses accents, to that we employ for YM: 1 (low) to 4 (high), and two contour tones, one rising and one falling.

In XM, the completive may be encoded in either of two ways, found in complementary distribution. Verbs that have a tone /3/ on $\mu 1$ in their lexical representation (i.e., irrealis), receive a tone /2/ in the completive, as in (31a). Verbs with a lexical tone /1/ on $\mu 1$ (mostly, but not exclusively, associated with intransitive verbs as in YM; see examples below), receive the prefix *ndi¹*, cognate of YM *ni¹* used in the CPL-1 (31b). Among the bimoraic examples in (31), there are also instances of variant verbs of the two classes *ku+* and *ka+*, again cognate with YM.

(31)	IRR	CPL		
a.	cha ³ bi ³	cha ² bi ²	(tr)	‘pay’
	ka ³ a ³	ka ² a ²	(intr)	‘bathe (at sweat lodge)’
	ku ³ un ³	ku ² un ²	(intr)	‘produce’
	ku+ ³ tu ⁴	chu+ ² tu ⁴	(intr)	‘get filled’
	ka+ ³ kin ³	xa+ ² kin ²	(tr)	‘place’
	ka+ ³ mi ³	xa+ ² mi ²	(tr)	‘burn’

b. ka ¹ nda ¹	ndi ¹ -ka ¹ nda ¹	(intr)	‘move, tremble’
ku ¹ ni ¹	ndi ¹ -ku ¹ ni ¹	(tr)	‘want’
kaʔ ¹ an ¹	ndi ¹ -kaʔ ¹ an ¹	(intr)	‘talk’
ndaʔ ¹ i ¹	ndi ¹ -ndaʔ ¹ i ¹	(intr)	‘shout’, ‘bark’, ‘scold’

In contrast to XM, in which the two markings available for the completive appear in complementary distribution, in YM there are two alternative completive forms: CPL-1 and CPL-2. As we have pointed out above, for invariant verbs the form of CPL-2 with lexical tones /1/ or /14/ on μ 1 is homophonous with that of the irrealis. A topic for future research, therefore, is whether in YM simple tonal marking of the completive is more common when a contrast with the irrealis is maintained. This contrast would be maintained when (i) the stem-initial tone of the irrealis is /3/, or (ii) there is a segmental distinction between the lexical (completive) and irrealis stem (see §4.3.3).

The XM sample also includes 13 verbs with the prefixes *ku*³- for the irrealis and *xi*¹- for the completive. Some illustrative examples are given in (32). These verbs form their own inflectional class whose membership appears to be lexically determined.

(32) IRR	CPL	INCPL		
ku ³ -nda ³ a ³	xi ¹ -nda ² a ²	nda ⁴ a ³	(tr)	‘look after’
ku ³ -na ³ ni ⁴	xi ¹ -na ² ni ⁴	na ⁴ ni ⁴	(intr)	‘be called’
ku ³ -nda ³ tu ³	xi ¹ -nda ² tu ²	nda ⁴ tu ⁴	(tr)	‘wait for’
ku ³ -ndi ³ chi ³	xi ¹ -ndi ² chi ²	ndi ⁴ chi ³	(intr)	‘stand’ (SG sub.)

For the verbs in (32), notice that when the lexical stem has tone /3/ in the initial mora, the completive has /2/ (e.g. ‘look after’ (IRR) *ku*³-*nda*³*a*³ vs. (CPL) of *xi*¹-*nda*²*a*²). This may result from tone sandhi rules, pointing to a harmonization of tone /3/ to the low tone of the prefix *xi*¹-, but could also be taken as the bridging context from which the possible reanalysis of tone /2/ as a marker of completive in XM resulted in the forms given in (31a). Similarly, the stem formative *ku*+ that we find in the irrealis of verbs of the *ku*+ class is undoubtedly the historical reflex of the stem formative of that class. As we have pointed out, the irrealis stem of many verbs was historically levelled to the entire paradigm, producing a large number of invariant verbs that now have a stem with a velar onset /k/ in all forms.³⁴

For the incomplete, when the tonal melody of the base is /1-1/ as in (33a), inflectional tone /4/ overwrites lexical tone /1/ on μ 1, just as it does with initial

³⁴ In Cuicatec, the historical reflex of an irrealis prefix *ku*³- can still be observed (see Feist & Palancar, this volume).

lexical tone /3/, as shown in (33b–c). However, when the melody of the base is /1–3/ or /1–4/, as in (33d–e), the inflectional tone of the incomplete is added onto, but does not replace, the lexical tone /1/ of μ_1 , resulting into a falling tone on μ_1 that given the lack of pitch specification in the source we write simply as falling (v^F) though most likely the contours are /41–3/ and /41–4/.

(33)	IRR	INCPL	CPL	Transitivity	Gloss
a. /1–1/	ka ¹ nda ¹	ka ⁴ nda ¹	ndi ¹ -ka ¹ nda ¹	(intr)	‘move, tremble’
	ku ¹ ni ¹	ku ⁴ ni ¹	ndi ¹ -ku ¹ ni ¹	(tr)	‘want’
	ka ² an ¹	ka ⁴ an ¹	ndi ¹ -ka ² an ¹	(intr)	‘talk’
	nda ² i ¹	nda ⁴ i ¹	ndi ¹ -nda ² i ¹	(intr)	‘shout, bark, scold’
b. /3–3/	ka ³ ku ³	ka ⁴ ku ⁴	ka ² ku ²	(intr)	‘be born’
	ka ³ chi ³	ka ⁴ chi ³	ka ² chi ²	(tr)	‘say’
	chi ³ i ³	chi ⁴ i ⁴	chi ² i ²	(intr)	‘get wet’
	nda ² a ³	nda ⁴ a ³	nda ² a ²	(intr)	‘get better, recuperate’
c. /3–4/	ko ³ ko ⁴	ko ⁴ ko ⁴	ko ² ko ⁴	(tr)	‘swallow, gulp down’
	sa ³ ta ⁴	sa ⁴ ta ⁴	sa ² ta ⁴	(tr)	‘buy’
	ta ³ an ⁴	ta ⁴ an ⁴	ta ² an ⁴	(tr)	‘put in, insert’ (PL. OBJ)
d. /1–3/	ka ¹ ku ³	ka ^{F(41)} ku ³	ndi ¹ -ka ¹ ku ³	(intr)	‘escape’
	ndi ¹ ko ³	ndi ^{F(41)} ko ³	ndi ¹ -ndi ¹ ko ³	(intr)	‘cool down’
	ndo ¹ o ³	ndo ^{F(41)} o ³	ndi ¹ -ndo ¹ o ³	(intr)	‘stay’
	ka ² un ³	ka ^{F(41)} un ³	ndi ¹ -ka ² un ³	(intr)	‘burn’
e. /1–4/	chi ¹ yo ⁴	chi ^{F(41)} yo ⁴	ndi ¹ -chi ¹ yo ⁴	(intr)	‘boil’
	nu ¹ na ⁴	nu ^{F(41)} na ⁴	ndi ¹ -nu ¹ na ⁴	(intr)	‘open’
	sa ¹ a ⁴	sa ^{F(41)} a ⁴	ndi ¹ -sa ¹ a ⁴	(intr)	‘get warm’

A summary comparison of the realization of the incomplete in Yoloxóchitl Mixtec and Xochapa Mixtec is given in Table 11 with data of 106 verbs in XM.³⁵ The number of occurrences of each tonal melody pattern is given in the final column.

For the incomplete form of verbs with lexical melody /1–1/, both YM and XM keep the lexical tone on μ_2 , but with /3–3/ YM shows a split pattern based on syllable structure (i.e., the inflectional pattern is predictable from stem morphophonology), whereas the split in XM is encoded in the lexicon (i.e., there are

³⁵ Not included in the table of 106 verbs are five somewhat irregular verbs that are not easily categorized. Three add ku^3 - to an incomplete form with a tonal melody of /4–4/; one adds ku^3 - to an incomplete form with a tonal melody of /3–4/. The final irregular verb is ‘go’.

no apparent morphological or phonological conditioning elements). For verbs with melodies /1-3/ and /1-4/, XM adds the high tone of the incomplete to the lexical low tone on μ_1 , resulting in a falling tone, but maintains the lexical tone of μ_2 . In YM, in contrast, the high tone /4/ of the incomplete replaces the lexical tone /1/ on μ_1 which is aligned then to μ_2 , resulting in a rising tone on μ_2 . It is worth noting that YM never manifests falling tone on the first mora; the rightward shift of lexical tone /1/ might, therefore, reflect this phonotactic constraint. Finally, for verbs with melody /3-4/, XM preserves lexical tone on μ_2 in all cases, whereas syllabic structure is again very relevant for YM, where we find a split between /4-4/ on disyllabic stems and /4-24/ (from underlying /4-34/) on monosyllabic stems.

Table 11: Comparison of the incomplete in bimoraic verbs of YM and XM

		YM	XM	Occurrences
/1-1/	disyllabic	CV ¹ CV ¹ > CV ⁴ CV ¹	CV ¹ CV ¹ > CV ⁴ CV ¹	11
	monosyllabic	CV ¹ V ¹ > CV ⁴ V ¹	CV ¹ V ¹ > CV ⁴ V ¹	4
/3-3/	disyllabic	CV ³ CV ³ > CV ⁴ CV ³	CV ³ CV ³ > CV ⁴ CV ³	11
			CV ³ CV ³ > CV ⁴ CV ⁴	25
	monosyllabic	CV ³ V ³ > CV ⁴ V ⁴	CV ³ V ³ > CV ⁴ V ³	3
			CV ³ V ³ > CV ⁴ V ⁴	15
/1-3/	disyllabic	CV ¹ CV ³ > CV ⁴ CV ¹³	CV ¹ CV ³ > CV ^F CV ³	7
	monosyllabic	CV ¹ V ³ > CV ⁴ V ¹³	CV ¹ CV ³ > CV ^F V ³	6
/1-4/	disyllabic	CV ¹ CV ⁴ > CV ⁴ CV ¹⁴	CV ¹ CV ⁴ > CV ^F CV ⁴	2
	monosyllabic	CV ¹ V ⁴ > CV ⁴ V ¹⁴	CV ¹ CV ⁴ > CV ^F V ⁴	1
/3-4/	disyllabic	CV ³ CV ⁴ > CV ⁴ CV ⁴	CV ³ CV ⁴ > CV ⁴ CV ⁴	20
	monosyllabic	CV ³ V ⁴ > CV ⁴ V ²⁴	CV ³ V ⁴ > CV ⁴ V ⁴	1

The tonal inflection of both XM and YM is in general very similar. We take this systemic similarity to point to a common feature of the Guerrero subgroup. But a glimpse into the inflection of San Juan Colorado Mixtec, another Mixtec language for which a published dictionary provides ample comparative data on verb morphology, suggests that Mixtec languages may also show tonal allomorphy not predictable by morphophonology.

San Juan Colorado Mixtec (henceforth SJCM) is spoken in the district of Jamiltepec, state of Oaxaca, by approximately 6,000 people. SJCM pertains to the Costa branch of Mixtec (see Figure 1) and it is thus only distantly related to the Southern Baja branch containing the Guerrero subgroup. To provide comparative data, we rely on a sample of 179 verbs from the dictionary by Stark et al. (1986) that has the following two properties: (a) the verbs have bimoraic stems in all three forms; and (b) the irrealis and the completive have the same tonal melody although the

irrealis is formed by preceding the irrealis stem with the modal marker *kwa*^{1,36}. This means that the incomplete form is the only form marked by tone, and it is precisely in the tonal patterns found in incomplete that the inflection of SJCM differs from both YM and XM.

Based exclusively on the data available in Stark et al. (1986), the incomplete form in SJCM displays a tonal allomorphy whose conditioning appears to be lexical. This means that the morphology of the incomplete is best analyzed by grouping SJCM verbs according to the tonal melodies manifested in the incomplete. Such an organization results in the four inflectional classes described in Table 12.

Table 12: Tonal classes for the incomplete in SJCM.

CLASS	LEX	IRR	CPL	INCPL	
I	17 /1-1/	<i>kwa</i> ¹	<i>ka</i> ¹ <i>an</i> ¹	<i>ka</i> ¹ <i>an</i> ¹ /3-3/	<i>ka</i> ³ <i>an</i> ³ (tr) 'bite'
	31 /1-2/	<i>kwa</i> ¹	<i>ka</i> ² <i>yi</i> ²	<i>ka</i> ² <i>yi</i> ²	<i>ka</i> ² <i>yi</i> ³ (tr) 'paint'
	1 /3-2/	<i>kwa</i> ¹	<i>ko</i> ¹ <i>kon</i> ²	<i>ko</i> ¹ <i>kon</i> ²	<i>ko</i> ³ <i>kon</i> ³ (tr) 'swallow'
96	47 /2-2/	<i>kwa</i> ¹	<i>nu</i> ² <i>ña</i> ²	<i>nu</i> ² <i>ña</i> ²	<i>nu</i> ³ <i>ña</i> ³ (tr) 'open'
II	5 /1-1/	<i>kwa</i> ¹	<i>ndi</i> ¹ <i>yi</i> ¹	<i>ndi</i> ¹ <i>yi</i> ¹ /3-2/	<i>ndi</i> ³ <i>yi</i> ² (intr) 'get scorched'
	4 /1-2/	<i>kwa</i> ¹	<i>tsi</i> ¹ <i>yo</i> ²	<i>tsi</i> ¹ <i>yo</i> ²	<i>tsi</i> ³ <i>yo</i> ² (intr) 'have diarrhea'
29	20 /2-2/	<i>kwa</i> ¹	<i>ki</i> ² <i>ku</i> ²	<i>ki</i> ² <i>ku</i> ²	<i>ki</i> ³ <i>ku</i> ² (tr) 'sew'
III	4 /1-1/	<i>kwa</i> ¹	<i>nda</i> ¹ <i>tsi</i> ¹	<i>nda</i> ¹ <i>tsi</i> ¹ /3-1/	<i>nda</i> ³ <i>tsi</i> ¹ (intr) 'become loose'
	1 /1-2/	<i>kwa</i> ¹	<i>na</i> ¹ <i>ñi</i> ²	<i>na</i> ¹ <i>ñi</i> ²	<i>na</i> ³ <i>ñi</i> ¹ (tr) 'hit'
6	1 /2-1/	<i>kwa</i> ¹	<i>ka</i> ² <i>yi</i> ¹	<i>ka</i> ² <i>yi</i> ¹	<i>ka</i> ³ <i>yi</i> ¹ (intr) 'use make-up'
IV	20 /1-1/	<i>kwa</i> ¹	<i>ka</i> ² <i>an</i> ¹	<i>ka</i> ² <i>an</i> ¹ /2-3/	<i>ka</i> ² <i>an</i> ³ (tr) 'talk'
	3 /1-2/	<i>kwa</i> ¹	<i>ja</i> ¹ <i>ta</i> ²	<i>ja</i> ¹ <i>ta</i> ²	<i>ja</i> ² <i>ta</i> ³ (tr) 'buy'
	21 /2-1/	<i>kwa</i> ¹	<i>nda</i> ² <i>ba</i> ¹	<i>nda</i> ² <i>ba</i> ¹	<i>nda</i> ² <i>ba</i> ³ (tr) 'switch off'
48	4 /2-2/	<i>kwa</i> ¹	<i>na</i> ² <i>ma</i> ²	<i>na</i> ² <i>ma</i> ²	<i>na</i> ² <i>ma</i> ³ (tr) 'confess'

Verbs in classes I to III have a high tone /3/ on $\mu 1$ of the incomplete. Again, this marking strategy is reminiscent of the use of a high tone /4/ in both YM and XM for the same function. What is different about SJCM is that there are many verbs (class IV, approximately 27 percent of the total number of verbs analyzed) that mark the incomplete with a mid tone /2/. Additionally, taking simply the surface forms of verbs in isolation, there is no apparent motivation for the tonal changes manifested on the second mora of the verbs analyzed. Even if the default rule

³⁶ The appearance of this mood marker is undoubtedly the cognate of the negative irrealis marker *kwa*¹⁴ used with YM verbs whose irrealis form has /14/ on $\mu 1$. This point was discussed in section 4.2.3.

were to raise the lexical tone to tone /3/ in all morae, the patterns manifested in classes II and III would still require an explanation.

Tonal allomorphy involving the marking of the incomplete is also found in YM, as we saw in Table 6 above, but in YM the outcome of tonal variance (except for the verb noted in fn. 23) is predictable from a combination of tonal melody and syllabic structure. Nevertheless, from an Oto-Manguan perspective, the tonal classes of SCJM are not typologically odd. Similar lexically conditioned patterns are also found in many languages of this stock including Cuicatec (Feist & Palancar, this volume), Triqui (DiCanio, this volume), Mazatec (Léonard & Fulcrand, this volume), Amuzgo (Kim, this volume) and Chatino (Campbell, this volume). It remains possible that a more refined phonological analysis carried out on SJCM verbs with novel data might reveal tonal structures or the influence of floating tones that have escaped our understanding. This paper is a cordial invitation to carry out such research in this fascinating area of the grammar of Mixtec. Documenting these languages by way of well-informed lexical databases should be given priority before it is too late.

Acknowledgements

Research for this paper has been supported by grants PPG0048 and MDP0201 from the Hans Rausing Endangered Languages Project at the School of Oriental and African Studies, London (Amith); by NSF award #0966462 from the Documenting Endangered Languages program (Amith) and by the ESRC/AHRC grant RES-062–23–3126 ‘Endangered Complexity: Inflectional classes in Oto-Manguan languages’ (Palancar). In Yoloxóchitl, we (Amith and García Castillo) have been fortunate to work with many excellent consultants. In particular we would like to thank Constantino Teodoro Bautista, Esteban Guadalupe Sierra and Esteban Castillo García, all of whom have consistently accompanied us during all phases of fieldwork. In addition to the contributions of the preceding three consultants, the corpus has benefitted immensely from the generosity of the following, who recorded material in a wide range of genres: Alberto Prisciliano, Carlos Rómulo, Constantino Teodoro Celso, Estela Santiago Castilla, Fernando Niño, Guillermina Nazario Sotero, Lamberto García, Marcelina Encarnación, Mario Salazar, Martín Severiano, Maximiliano Francisco, Maximino Meza, Pedro Castillo, Pedro Verónico, Raúl Felipe Maragarito, Santa Cruz Tiburcio, Silvino Ramírez García, Simón Castillo, Soledad García Bautista and Victorino Ramos. Finally, our understanding of Mixtec phonetics, phonology, and morphosyntax has been enhanced through conversations and collaborative work with Christian DiCanio and Ryan

Shosted. We would also like to thank Lucien Carroll for reading an earlier draft of this paper. All errors and deficiencies remain our responsibility.

References

- Alexander, Ruth Mary. 1988. A syntactic sketch of Ocoatepec Mixtec. In C. Henry Bradley and Barbara E. Hollenbach, (eds.), *Studies in the syntax of Mixtecan languages*, volume 1, 151–304. Dallas and Arlington, Texas: Summer Institute of Linguistics and University of Texas at Arlington.
- Alvarado, Fray Francisco de. 1962 [1593]. *Vocabulario en lengua misteca*. Mexico City: Instituto Nacional Indigenista and Instituto Nacional de Antropología e Historia.
- Amith, Jonathan D., and Rey Castillo García. n.d. 1. A dictionary of Yoloxóchitl Mixtec. Ms.
- Amith, Jonathan D., and Rey Castillo García. n.d. 2. A grammar of Yoloxóchitl Mixtec. Ms.
- Beaty de Farris, Kathryn. 2002. *Diccionario básico del mixteco de Yosondúa, Oaxaca*. In collaboration with Pablo García Sánchez, Rubén García Sánchez, Jesús Ojeda Sánchez, Agustín San Pablo García and Apolonio Santiago Jiménez. Mexico City: Instituto Lingüístico de Verano.
- Bradley, C. Henry. 1970. *A Linguistic Sketch of Jicaltepec Mixtec*. Summer Institute of Linguistics Publications in Linguistics and Related Fields, 25. Normal, Okla.: Summer Institute of Linguistics and University of Oklahoma.
- Bradley, C. Henry and Barbara Hollenbach, eds. 1988–92. *Studies in the syntax of Mixtecan languages*. 4 vols. Dallas and Arlington, Texas: Summer Institute of Linguistics and the University of Texas at Arlington. (I: Introduction and syntactic sketches of Jamiltepec, Ocoatepec, Silacayoapan; II: Ayutla, Coatzacoapan; III: Alacatlátlazala, Diuxi-Tilantongo, Concepción Pápalo Cuicatec; IV: Yosondúa, Copala Trique).
- Campbell, Lyle, Terrence Kaufman, and Thomas C. Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62:530–70.
- Castillo García, Rey. 2007. La fonología tonal del mixteco de Yoloxóchitl, Gro. MA thesis Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City, Mexico.
- DiCanio, Christian T., Caicai Zhang, Jonathan D. Amith, Rey Castillo García, and Douglas H. Whalen. Forthcoming. Phonetic structure in Yoloxóchitl Mixtec consonants. *International Journal of American Linguistics*.
- Ferguson de Williams, Judith. 2007. *Gramática popular del mixteco del municipio de Tezoatlán, San Andrés Yutatío, Oaxaca*. Serie gramáticas de lenguas indígenas, 9. Mexico City: Instituto Lingüístico de Verano.
- Fishman, Joshua. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon, England: Multilingual Matters.
- Hills, Robert A. 1990. A syntactic sketch of Ayutla Mixtec. In C. Henry Bradley and Barbara E. Hollenbach, (eds.), *Studies in the syntax of Mixtecan languages*, volume 2, 1–260. Dallas and Arlington, Texas: Summer Institute of Linguistics and University of Texas at Arlington.

- Hinton, Leanne, Gene Buckley, Marv Kramer, and Michael Meacham. 1991. Preliminary analysis of Chalcatongo Mixtec tone. In James E. Redden, (eds.), *Papers from the American Indian Languages Conferences held at the University of California, Santa Cruz, July and August 1991*. Occasional Papers on Linguistics No. 16, 173–82. Carbondale: Southern Illinois University.
- Hollenbach, Barbara E. 2001. The Lord's Prayer from the 1567 Hernández catechism. Downloaded at http://barbaraelenahollenbach.com/Bibliography.htm#Mixtec_Mixtecan.
- Johnson, Audrey F. 1988. A syntactic sketch of Jamiltepec Mixtec. In C. Henry Bradley and Barbara E. Hollenbach, (eds.), *Studies in the syntax of Mixtecan languages*, volume 1, 11–150. Dallas and Arlington, Texas: Summer Institute of Linguistics and University of Texas at Arlington.
- Josserand, Judy K. 1983. Mixtec dialect history. PhD thesis, Tulane University.
- Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec*. University of California Publications in Linguistics 127. Berkeley: University of California Press.
- Macaulay, Monica, and Joseph C. Salmons. 1995. The phonology of glottalization in Mixtec. *International Journal of American Linguistics*, 61:38–61.
- McKendry, Inga. 2013. Tonal association, prominence and prosodic structure in south-eastern Nochixtlán Mixtec. PhD thesis, University of Edinburgh.
- Overholt, Edward. 1961. The tonemic system of Guerrero Mixtec. In Manuel Gamio and Raúl Noriega, (eds.), *A William Cameron Townsend en el vigésimoquinto aniversario del Instituto Lingüístico de Verano*, 597–626. Mexico City: Instituto Lingüístico de Verano.
- Ramírez Flores, Esperanza (narrator), and Candice Brown de Beatham. 2012. *Tu'un na ñuu Kiyaa. Mixteco del oeste de Juxtlahuaca*. Electronic version. 2nd edition. Mexico City: Instituto Lingüístico de Verano.
- Small, Priscilla C. 1990. A syntactic sketch of Coatzacoapan Mixtec. In C. Henry Bradley and Barbara E. Hollenbach, (eds.), *Studies in the syntax of Mixtecan languages*, volume 2, 261–479. Dallas and Arlington, Texas: Summer Institute of Linguistics and University of Texas at Arlington.
- Smith Stark, Thomas. 1995. El estado actual de los estudios de las lenguas mixtecanas y zapotecanas. In Leonardo Manrique, Yolanda Lastra, and Doris Bartholomew, (eds.), *Panorama de los estudios de las lenguas indígenas de México*, 2 vols. Colección Biblioteca Abya-Yala 16, 5–186. Quito, Ecuador: Ediciones Abya-Yala.
- Smith Stark, Thomas. 1994. Mesoamerican calques. In Carolyn J. MacKay and Verónica Vázquez, (eds.), *Investigaciones lingüísticas en Mesoamérica*, 5–186. Mexico City: Universidad Nacional Autónoma de México.
- Smith Stark, Thomas. 1982. Mesoamerican calques. Paper presented at the symposium on Mesoamerican dialectology and language history, 44th International Congress of Americanists, Manchester, England.
- Stark, Sharon (Sara), Audrey (Andrea) Johnson P., and Benita Gonzalez de Guzman. 2006. *Diccionario básico del mixteco de Xochapa, Guerrero*. Segunda edición. Mexico City: Instituto Lingüístico de Verano.
- Stark Campbell, Sara, Andrea Johnson Peterson, Filiberto Lorenzo Cruz (compiladores), and Catalina López de García, Daniel Fidencio García Alavez (colaboradores mixtecos). 1986. *Diccionario mixteco de San Juan Colorado*. Serie de Vocabularios y Diccionarios Indígenas “Mario Silva y Aceves”, no. 29. Mexico City: Instituto Lingüístico de Verano.
- Suárez, Jorge A. 1983. *The Mesoamerican Indian languages*. Cambridge, England: Cambridge University Press.

Thornton, Anna M. 2011. Overabundancy (Multiple forms realizing the same cell): A non-canonical phenomenon in Italian. In Maria Goldbach, Marc-Olivier Hinzelin, Martin Maiden, and John Charles Smith, (eds.), *Morphological Autonomy*, 362–85. Oxford: Oxford University Press.

A mixed system of agreement in the suffix classes of Lealao Chinantec

Enrique L. Palancar

Received: 24 March 2014 / Accepted: 27 November 2014 / Published online: 25 December 2014
© Springer Science+Business Media Dordrecht 2014

Abstract In Lealao Chinantec, an Oto-Manguean language of Mexico, verbs receive a set of suffixes which index person and number of the subject. There are various sets of such affixes producing the sort of allomorphy that is typical of inflectional classes. In this paper, I study the distribution of such suffixes in a large sample of verbs from Rupp and Rupp's (1996) dictionary and show that the allomorphy is largely morphosyntactically motivated for a group of verbs that have two paradigms, animate and inanimate. For many other verbs this morphosyntactic mapping fails and the allomorphy we observe is essentially morphological in nature. This allomorphy reflects the interaction of what I treat as two competing systems of agreement. One such system is a more canonical agreement system based on subject (S/A), the other is based on animacy agreement with the absolutive argument (S/O) (also found in other Chinantecan languages). These two systems and the way they talk to each other has important consequences for the morphology-syntax interface of Lealao Chinantec.

Keywords Chinantec · Oto-Manguean · Inflectional classes · Animacy · Agreement

1 Introduction

The Chinantecan languages (Oto-Manguean) of Mexico have started to awaken interest in word-and-paradigm approaches to morphology. This can be seen in a number of recent publications, such as for example Baerman (2013b), Baerman and Palancar (forthcoming), Baerman (2014a), Brown and Evans (2013), Corbett and Baerman (2006), Finkel and Stump (2007, 2009), Palancar (2014), Stump and Finkel (2013), etc.

E.L. Palancar (✉)
Structure et Dynamique des Langues (UMR8202), CNRS; Surrey Morphology Group, University
of Surrey, Guildford, UK
e-mail: palancar@vjf.cnrs.fr

Such an interest is justified. It is based on a growing acknowledgment—sometimes not without a tinge of bewilderment—that the inflection of Chinantecan languages ranks amongst the most intricate in the world’s languages. This is mainly due to the fact that inflected forms in these languages result from the interaction of multiple systems of classification, which are for the most part mutually independent from each other. Typical such systems in the family are the prosodic classes (made by tone and stress) and the stem alternation patterns, but the language known as Lealao Chinantec is remarkable within this family in also having different sets of person and number suffixes. These suffixes add yet another layer of complexity to the inflectional system because some verbs take a given set, as for example the verb in (1), while other verbs take a different set, as for example the verb in (2).¹ For convenience, the examples illustrate only singular forms in the future tense.²

<p>(1) $\text{ʔi}^4\text{nuu}^4$ (tr) ‘hear something’ FUT</p> <table border="1" style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px;">1SG</td> <td style="padding: 2px;">$\text{ʔi}^2\text{-nuu}^4\text{-y}$</td> </tr> <tr> <td style="padding: 2px;">2SG</td> <td style="padding: 2px;">$\text{ʔi}^2\text{-nuu}^3\text{-y}$</td> </tr> <tr> <td style="padding: 2px;">3SG</td> <td style="padding: 2px;">$\text{ʔi}^4\text{-nuu}^4\text{-}\emptyset$</td> </tr> </table>	1SG	$\text{ʔi}^2\text{-nuu}^4\text{-y}$	2SG	$\text{ʔi}^2\text{-nuu}^3\text{-y}$	3SG	$\text{ʔi}^4\text{-nuu}^4\text{-}\emptyset$	<p>(2) $\text{ʔi}^4\text{xiee}^2\text{y}$ (tr) ‘deduct something’ FUT</p> <table border="1" style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px;">1SG</td> <td style="padding: 2px;">$\text{ʔi}^2\text{-xiaa}^1\text{-a}^2$</td> </tr> <tr> <td style="padding: 2px;">2SG</td> <td style="padding: 2px;">$\text{ʔi}^2\text{-xiaa}^4\text{-u}^3$</td> </tr> <tr> <td style="padding: 2px;">3SG</td> <td style="padding: 2px;">$\text{ʔi}^4\text{-xiee}^2\text{-y}$</td> </tr> </table>	1SG	$\text{ʔi}^2\text{-xiaa}^1\text{-a}^2$	2SG	$\text{ʔi}^2\text{-xiaa}^4\text{-u}^3$	3SG	$\text{ʔi}^4\text{-xiee}^2\text{-y}$
1SG	$\text{ʔi}^2\text{-nuu}^4\text{-y}$												
2SG	$\text{ʔi}^2\text{-nuu}^3\text{-y}$												
3SG	$\text{ʔi}^4\text{-nuu}^4\text{-}\emptyset$												
1SG	$\text{ʔi}^2\text{-xiaa}^1\text{-a}^2$												
2SG	$\text{ʔi}^2\text{-xiaa}^4\text{-u}^3$												
3SG	$\text{ʔi}^4\text{-xiee}^2\text{-y}$												

The Lealao Chinantec verbs in (1) and (2) could be said to belong to two different inflectional classes because they select different suffix sets: the class selecting $-\text{y}/-\text{y}/-\emptyset$ and the class selecting $-\text{a}^2/-\text{u}^3/-\text{y}$.³ Rupp and Rupp (1996) identify four such classes, given in (3), but I argue in Sect. 3.1 that in reality we can talk of six. The classes have a different membership size: some contain hundreds of verbs, some just a handful.

(3)

	Class I	Class II	Class III	Class IV
1SG	-y	-a ⁴	-a ⁴	-y
2SG	-y	-u ³	-y	-u ³
3SG	-∅	-y	-∅	-∅/-ʔ

In this paper, I show that these suffix classes constitute an interesting object of study for morphological theory because they instantiate a type of inflection mixing two systems of verbal agreement, one involving person-number agreement with subject (S/A) and another animacy agreement with the S/O argument. Different verbs make use of such systems differently.

¹Tone is indicated by numbers. There are four level tones (very high 1, high 2, mid 3 and low 4) and two ascending contour tones (mid-high 32 and low-high 42). Acute stress indicates ‘ballistic stress’ treated in Sect. 2.1. A nasal vowel is represented by underlining.

²Abbreviations: AN: animate; CL: class; CPL: completive; DEICT: deictic; EXCL: exclusive; FUT: future; INAN: inanimate; INCL: inclusive; INCLP: incomplete; INFL: inflectional; intr: intransitive; IRR: irrealis; PL: plural; PRF: perfect; PRS: present; PST: past; RELAT: relational; SG: singular; SUB: subject; tr: transitive.

³As we will see in Sect. 3.1, the suffix -y is a palatal consonant that cannot bear tone.

The lexicographic tradition of the SIL treats animacy agreement as a lexical phenomenon that involves two different verbs which are given their own independent entry in the dictionaries: one being the inanimate verb, the other the animate one as in an animacy pair. An example of such a pair is given in (4)–(5). The treatment of these verbs as being two lexical items has been justified by them having different prosodic inflection, but in this paper I argue that the forms involved could be alternatively viewed as instantiating two different paradigms of the same verbal lexeme.

<p>(4) $\text{ʔi}^d\text{láʔ}^2$ (tr) ‘get something down’</p> <table style="margin-left: 40px; border-collapse: collapse;"> <tr> <td></td> <td style="text-align: center;">FUT</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">1SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">2SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^2$</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">3SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^4\text{-lá}\text{ʔ}^2\text{-}\emptyset$</td> </tr> </table>		FUT	1SG	$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$	2SG	$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^2$	3SG	$\text{ʔi}^4\text{-lá}\text{ʔ}^2\text{-}\emptyset$	<p>(5) $\text{ʔi}^d\text{láyʔ}^d$ (tr) ‘get an animate entity down’</p> <table style="margin-left: 40px; border-collapse: collapse;"> <tr> <td></td> <td style="text-align: center;">FUT</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">1SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^2\text{-lá}\text{ʔ}^{d2}\text{-á}^2$</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">2SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^2\text{-lá}\text{ʔ}^1\text{-u}^3$</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 5px;">3SG</td> <td style="border: 1px solid black; padding: 2px;">$\text{ʔi}^4\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$</td> </tr> </table>		FUT	1SG	$\text{ʔi}^2\text{-lá}\text{ʔ}^{d2}\text{-á}^2$	2SG	$\text{ʔi}^2\text{-lá}\text{ʔ}^1\text{-u}^3$	3SG	$\text{ʔi}^4\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$
	FUT																
1SG	$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$																
2SG	$\text{ʔi}^2\text{-lá}\langle\text{y}\rangle\text{ʔ}^2$																
3SG	$\text{ʔi}^4\text{-lá}\text{ʔ}^2\text{-}\emptyset$																
	FUT																
1SG	$\text{ʔi}^2\text{-lá}\text{ʔ}^{d2}\text{-á}^2$																
2SG	$\text{ʔi}^2\text{-lá}\text{ʔ}^1\text{-u}^3$																
3SG	$\text{ʔi}^4\text{-lá}\langle\text{y}\rangle\text{ʔ}^4$																

Besides prosodic formatives, the animacy pair in (4)–(5) also contrasts in the affix set they select. Like (1) and (2), the verb in (4) would belong to Class I given in (3), while (5) would belong to Class II.⁴ I argue in Sect. 4 that the selection of the suffix sets responds to the morphosyntactic machinery of animacy agreement, and for such verbs there is no need to understand the suffix sets as having inflectional class organization. However, I also show in Sect. 5 that such machinery ceases to be operational when one zooms out from this core into the lexicon at large, where the distribution of the suffixes with inanimate verbs such as (3) becomes less predictable and should be better accounted for as a lexical matter, which justifies in turn their organization in inflectional classes. Notice for example that the verb in (2) has an inanimate O just like the verbs in (1) or (4), but receives the same suffixes as the animate verb in (5). In this respect, I claim that the set in (2) and to a large extent in (4) responds to the other agreement system at play and signals person-number agreement of subject, instead of animacy.

The paper has the following structure. In the next section, I introduce basic notions about both verbal inflection and animacy gender in Lealao Chinantec. This leads in Sect. 3 to introducing the classes as the object of study together with the sample I use. As a great deal of the allomorphy we observe involving the selection of subject suffixes has to do with animacy agreement, in Sect. 4 I study in detail the verbs of the sample in the relation they hold to animacy and the suffixes they select. It then becomes clear that for the verbs with two animacy paradigms, the distribution of the suffixes selected talks about the presence of a functional morphosyntactic mapping, which I develop in Sect. 5.1. In Sect. 5.2, I further suggest that when the system is taken as a whole, the default class is used to index subject agreement. Section 6 sums up and concludes.

⁴In the 1st and 2nd person in (4) and in the 3rd person in (5), the suffix -y is realized as an infix. When this happens, the affix is indicated within angle brackets following the Leipzig Glossing Rules convention. The reason for this infixation is given in Sect. 3.1.

2 Overview of the verbal inflection of Lealao Chinantec

Lealao Chinantec is spoken in Mexico by approximately 2,000 people in the village of San Juan Lealao in the state of Oaxaca. Lealao Chinantec—called *Fáʔ⁴-hmii⁴²* by its speakers (lit. language-Chinantec)—is one of a dozen mutually unintelligible languages that form the Chinantecan branch of Oto-Manguean; itself one of the most diverse linguistic phyla in the Americas. Lealao Chinantec is spoken at the southern borders of the Chinantla (the traditional region inhabited by the Chinantecan people), and because of this it is a geographical outlier with respect to other Chinantecan languages.

Lealao Chinantec is a typical Chinantecan language in many respects. In this section, I elaborate on two aspects of the verbal inflection of this language that are important in understanding the suffix classes as the object of the study. One is the role of tone and stress in the general making of verbal inflection, although I also mention in passing stem alternation patterns, and the other is animacy agreement.

2.1 Prosodic inflection and stem alternation patterns

Chinantecan languages are tone languages. Lealao Chinantec has up to six tonal contrasts; five of which are used in inflection. Tones are linked to morae mainly in vocalic nuclei. There are twelve vowels (six oral and six nasal) and five diphthongs. Vowels can be short and long. Syllables are always open. Syllables of the type /CVʔ/ (e.g. stem *laʔ⁴²* in the 1SG.FUT form *ʔi²-laʔ⁴²-á²* in (4)), /CVj/ (e.g. stem plus the suffix *-y /nuuj⁴²/* in the 1SG.FUT form *ʔi²-nuu⁴²-y* in (1)), or even /CVjʔ/ (e.g. stem plus infix *-y /lájʔ⁴/* in the 1SG.FUT form *ʔi²-lá<y>ʔ⁴* in (3)) could be treated as complex nuclei. Besides tone, every syllable receives one of two phonation types, traditionally talked about in terms of stress as ‘ballistic stress’ (represented with an acute accent) and ‘control stress’ (non-represented). “Ballistic syllables are characterized by an initial surge and rapid decay of intensity, with a resulting fortis articulation of the consonantal onset”, while “controlled syllables generally display a more gradual surge and decay of stress, as well as a longer duration of the maximum stress” (Foris 2000:16, but also in Rensch 1978, amongst others).⁵

Prosodic features are exploited by the morphology to realize a great range of inflectional distinctions in verbs (and to a certain extent also in nouns). I refer to the use of such features (tone and stress type together) as ‘prosodic inflection’. As an illustration of how prosodic inflection works, consider the partial paradigm in (6) for the transitive verb *ʔi²tiuu²a²* ‘spill something’ (Rupp and Rupp 1996:458).

⁵The received view on ballistic stress is acoustic in nature. Alternatively, ballistic stress has also been characterized in articulatory terms as a reflection of laryngeal phonation either by an increased sub-glottal pressure in Mugele (1982) or by a laryngeal abduction in Silverman (1994).

(6) $?i^2tiuu^2a^2$ (tr) ‘spill something’

	Three grammatical tenses			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG	tuu^4-y	ma^3-tuu^4-y	$?i^2-tuu^{42}-y$	tuu^4	tuu^4	tuu^{42}
2SG	tuu^3-y	ma^3-tiuu^1-y	$?i^2-tuu^3-y$	tuu^3	$tiuú^1$	tuu^3
3SG	$tiuú^4-\emptyset$	$ma^3-tiuu^3-\emptyset$	$?i^4-tiuu^4-\emptyset$	$tiuú^4$	$tiuú^3$	$tiuú^4$
1PL.INCL	$tiuú^{42}-a^2$	$ma^3-tiuu^2-a^2$	$?i^2-tiuu^2-a^2$	$tiuú^{42}$	$tiuú^2$	$tiuú^2$

The linguistic literature on Chinantecan languages commonly acknowledges that the twelve cells in (6) are sufficient to derive the other cells in the paradigm through complex sets of referral rules.⁶ In (6), we may see that prefixes encode tense values: a bare stem is used for the tense called ‘Present’, $?i^2$ —for the ‘Future’ ($?i^4$ —for 3SG/FUT) and ma^3 —for the ‘Perfect’. There are many such tenses encoded by other similar prefixes. All tenses are built on stems which are often further inflected by means of prosodic features for aspect (incomplete, complete) and mood (irrealis), sometimes also conflating information about the person-number of the subject. There are also verbs with invariant stems throughout the paradigm. Because stems play a crucial role for the verbal inflection of Lealao Chinantec, it is convenient to keep them visually separated from other inflected forms at least for some of the examples. For this reason, the inflected stems appear under the heading ‘INFL STEMS’.

While the inflection of stems is carried out by prosodic features, there is no consistent one-to-one mapping between form and meaning (for details, see Baerman and Palancar, [forthcoming](#)). In other words, there is not a single tone that may be consistently associated with the realization of a specific morphosyntactic value. For example, in the inflected stems of the verb in (6) tone 4 is used for 1SG/INCPL and 1SG/CPL, but also for 3SG/INCPL and 3SG/IRR; tone 42 for 1SG/IRR but also for 1PL/INCPL, etc. Similarly, ballistic stress (marked by the acute accent) is used for the CPL in all singular persons in (5), but also for 3SG/IRR. Other verbs have other prosodic patterns, and there are many of them, making verbs fall into a great number of inflectional classes defined by these prosodic distinctions (i.e. tone and ballistic stress together). Notice for instance the inflected stems of the transitive verb $?i^2la?^1a^2$ (tr) ‘get something down’ in (7)—also in (4) above—which belongs to a different prosodic class than $i^2tiuu^2a^2$ ‘spill something’. I use shading in (7) to highlight the contrasts in prosody with (6).

(7) $?i^2la?^1a^2$ (tr) ‘get something down’

	Three grammatical tenses			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG	$la?y?^4$	$ma^3-la?y?^4$	$?i^2-la?y?^4$	$la?^4$	$la?^4$	$la?^4$
2SG	$la?y?^{32}$	$ma^3-la?y?^1$	$?i^2-la?y?^2$	$la?^{32}$	$la?^1$	$la?^2$
3SG	$la?^{32}-\emptyset$	$ma^3-la?^2-\emptyset$	$?i^4-la?^2-\emptyset$	$la?^{32}$	$la?^2$	$la?^2$
1PL.INCL	$la?^{32}-a^2$	$ma^3-la?^1-a^2$	$?i^2-la?^1-a^2$	$la?^{32}$	$la?^1$	$la?^1$

⁶Contrary to the received view, Foris (2000) stipulates that a greater number of cells involving the third person is indeed needed to reconstruct the full paradigm of a verb.

Besides prosodic distinctions, verbs may further contrast in having different segmental stems in their paradigm which are arrayed in several stem alternation patterns. For this purpose compare the segmental part of the stems in the partial paradigm of the verb $\text{ʔi}^2\text{niuu}^2\text{a}^2$ (tr) ‘hear something’ in (8) with those of $i^2\text{tiuu}^2\text{a}^2$ ‘spill something’ in (6), (Rupp and Rupp 1996:457).

(8) $\text{ʔi}^2\text{niuu}^2\text{a}^2$ (tr) ‘hear something’

	INFL STEMS					
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG	$\text{nuu}^4\text{-y}$	$\text{ma}^3\text{-nuu}^4\text{-y}$	$\text{ʔi}^2\text{-nuu}^{42}\text{-y}$	nuu^4	nuu^4	nuu^{42}
2SG	$\text{nuu}^3\text{-y}$	$\text{ma}^3\text{-niuu}^3\text{-y}$	$\text{ʔi}^2\text{-nuu}^3\text{-y}$	nuu^3	niuu^3	nuu^3
3SG	$\text{nuu}^3\text{-}\emptyset$	$\text{ma}^3\text{-nuu}^4\text{-}\emptyset$	$\text{ʔi}^4\text{-nuu}^4\text{-}\emptyset$	nuu^3	nuu^4	nuu^4
1PL.INCL	$\text{nuu}^{32}\text{-a}^2$	$\text{ma}^3\text{-niuu}^2\text{-a}^2$	$\text{ʔi}^2\text{-niuu}^2\text{-a}^2$	nuu^{32}	niuu^2	niuu^2

Both verbs require two stems with fundamentally the same shape: a non-palatalized one (*tuu/nuu*) vs. a palatalized one (*tiuu/niuu*). However they differ in requiring the palatalized stem in different cells. Verbs like (6) use a palatalized stem for the 2SG/CPL and all aspectual stems involving 3SG and 1PL. Verbs like (8) require the palatalized stem only in three cells: 2SG/CPL, 1PL/CPL and 1PL/IRR. Such patterns are commonly independent of prosodic classes. In other words, knowing the prosodic class of a verb is generally uninformative of whether the verb undergoes stem alternations, and if it does of which type (but see Baerman 2014a, for trends in the correlations).

2.2 Animacy agreement

2.2.1 Nouns and gender based on animacy

Similarly, like a Chinantecan language nouns in Lealao Chinantec fall into one of two gender classes: inanimate vs. animate. There is no morphological indication of gender class on nouns, but gender class assignment is mostly based on semantics: nouns denoting animal life, including humans, are classified as animate whereas nouns referring to other entities are inanimate. This is coherent with what is typologically expected of a gender system of this type (Corbett 1991). There are a handful of exceptions to this assignment: nouns for certain atmospheric phenomena (rainbow, lightning, meteor, etc.) and astronomic entities (sun, moon, star, etc.) are animate.⁷ In the nominal domain within the NP, agreement in gender may target descriptive adjectives, anaphoric deictic determiners, alienable possession markers, quantifiers and numerals. In formal terms, the encoding of inanimate agreement is left unmarked, while the forms involved in the realization of animate agreement take various forms depending on the language. In Comaltepec Chinantec, animate agreement involves nasalization of the nuclear vowel or a postnuclear nasal consonant. This is shown by the contrasts in the two examples in (9) from Anderson (1989:56–7).

⁷In other Chinantecan languages like Comaltepec Chinantec (Anderson 1989:57), inanimate nouns can be upgraded to animate to make them more salient under certain discourse circumstances.

(9) Comaltepec Chinantec

- a. ʔe^{M} hʔ^{L} né^{M} ké^{M} dó^{M}
 that.INAN orange(INAN) yellow.INAN of. 1SG DEICT
 ‘that yellow orange of mine’.
- b. ʔi^{L} hʔ^{L} ná:n^{M} ké^{M} dó^{M}
 that.AN wasp(AN) yellow.AN of. 1SG DEICT
 ‘that yellow wasp of mine’

In Lealao Chinantec, animate marking may also involve nasalization (indicated by underlining), but for the most part the stem in animate targets also features a suffix *-y*. This is shown in (10) from Rupp (1989:63), glosses mine.

(10) Lealao Chinantec

- a. ni^3 mi^3 tíaa^3 ké^4 $\text{ʔi}^3\text{-ni}^3$
 three.INAN basket(INAN) white.INAN of.INAN. 1SG RELAT-that.INAN
 ‘Those three white baskets of mine.’
- b. $\text{aa}^3\text{-y}$ laa^4 $\text{tíe}^4\text{-y}$ kíaa^2 $\text{ʔi}^3\text{-ni}^3\text{-y}$
 three.AN-AN mule(AN) white.AN-AN of.AN. 1SG RELAT-that.AN-AN
 ‘Those three white mules of mine.’

The morphology involved in animacy agreement within the nominal domain involves strategies including prosodic formatives (e.g. demonstratives: ni^3 ‘that.INAN’ vs. $\text{ni}^3\text{-y}$ ‘that.AN’), stem apophony (e.g. possessive relationals: ké^4 ‘of.mine.INAN’ vs. kíaa^2 ‘of.mine.AN’) and even occasional lexical differences (e.g. numerals: ni^3 ‘three.INAN’ vs. aa^3 ‘three.AN’),⁸ but most notably there is also the suffix *-y* in the animate form of the adjective ‘white’ and in the relational element used for possession. This suffix is undoubtedly cognate of the suffix *-y* that encodes 3SG in forms such as $\text{ʔi}^4\text{láy}^4$ ‘s/he’ll get an animate entity down’ in (5) above. This has important consequences for the understanding of the distribution of this suffix and others like it that appear in the classes of Lealao Chinantec.

2.2.2 Animacy agreement in verbs

In the Chinantecan literature verbs are also talked about as agreeing in gender with one of their nominal arguments (for example, Anderson 1989; Rupp 1989; Merrifield 1968, etc.). The agreement pattern is ergative, as first pointed out explicitly by Foris (2000): intransitive verbs are said to agree in gender with their subject while transitive verbs agree with their object.⁹ In my opinion, the term ‘gender agreement’ raises expectations about the Chinantec system to be like the more canonical gender systems known in Indo-European languages where assignment is not semantically transparent

⁸The examples in (9) and (10) further illustrate that Chinantecan languages also differ as to targets for animate gender. This can be seen in the difference between the elements involving possession in the examples.

⁹In this, like in other typological configurations like the use of a morphological inverse, Chinantecan is typologically very similar to Algonquian. Whether animacy agreement and inversion are to be treated as implicational features is an open question.

for a large portion of the nominal lexicon. Because of this, I prefer using the term ‘animacy agreement’.

Animacy agreement is at its clearest in verbs that inflect both as animate and inanimate depending on the animacy of one of its core arguments. For example, the intransitive verb in (10). When a verb behaves like this, I refer to it as representing an animacy pair.

- (11) a. $\text{ʔi}^4\text{-k}\acute{\text{i}}\text{ʔ}^4$
 FUT-fall.IRR.3[3SG]
 ‘It’ll fall over.’
 b. $\text{ʔi}^4\text{-k}\acute{\text{i}}\langle\text{y}\rangle\text{ʔ}^4$
 FUT-fall.IRR.3<3SG>
 ‘He’ll fall over.’

As we will see in more detail in Sect. 4, the animate and the inanimate paradigms of verbs more often than not involve differences in prosodic inflection, such as for example the case of the verb in (12).

- (12) a. $\text{chia}\acute{\text{a}}^3$
 [PRS]cook_in_pot.INAN.INCPL.3[3SG]
 ‘It’s cooking (i.e. the vegetable).’
 b. $\text{chie}\acute{\text{e}}^{32}\text{-y}$
 [PRS]cook_in_pot.AN.INCPL.3-3SG
 ‘It’s cooking (i.e. the chicken).’

When discussing examples (4)–(5) above, I have already pointed out that animacy agreement is important to understand the selection of the suffix classes, and I claim in Sect. 5 it plays a crucial role in keeping animate and inanimate paradigms distinct for verbs of the type in (11) and (12). In this direction, Rupp (1989:20) points out that animacy agreement in intransitive verbs often involves “a syllable-final *i*”. In reality, he is referring to cases where the affix -y is used to index a 3SG animate subject as in (11b) and (12b), whereas using a bare stem (i.e., with no exponent) in (11a) and (11b) encodes 3SG inanimate subject. Such an analysis appears to be correct because inflected forms such as (13) are ungrammatical because they use agreement morphology which points to the wrong animacy value of the intended subject.

- (13) a. $*\text{chia}\acute{\text{a}}^3\text{-y}$
 [PRS]cook_in_pot.INAN.INCPL.3-3SG
 Intended reading: ‘It’s cooking (i.e. the vegetable).’
 b. $*\text{chie}\acute{\text{e}}^{32}$
 [PRS]cook_in_pot.AN.INCPL.3[3SG]
 Intended reading: ‘It’s cooking (i.e. the chicken).’

However, I will also show in Sect. 4.2 that things are far from being that clear when other verbs are involved, and for this reason I have refrained so far from using labels such as 3SG.INAN or 3SG.AN in the glosses so far.

The forms in (11) and (12) are treated in Rupp and Rupp’s (1996) dictionary as instantiating two different verbal entries; that is, as being two different, but semantically and formally related, lexemes. Under this view, the differences in form we

observe are accounted for as stemming from derivation (i.e., they represent lexical pairs, each with their own inflectional properties). Alternatively, in this paper I prefer to see the contrasts in (11) and (12) as representing inflected forms of two distinct paradigms of the same verbal lexeme. Under such a view, the differences in form stem from inflection (i.e., it is one lexeme whose paradigm includes cells contrasting in an animacy value).

As we will see in Sect. 4, only a portion of the Lealao Chinantec lexicon behaves like the pairs in (11) and (12). For expository purposes, I find it useful to talk about the animate and the inanimate paradigms in terms of ‘animacy pairs’ because the concept of ‘a pair’ evokes this inherent duality. Seeing the phenomenon as agreement also has advantages for the analysis I pursue in this paper, as I claim in Sect. 4.2 that the suffix sets in (3) play an important role in keeping the morphosyntax of animacy agreement working which is particularly important for verbs with two animacy paradigms.

In this section, I presented a general overview of the inflection of verbs in Lealao Chinantec involving prosodic features and stem alteration patterns. I have also introduced the notion of animacy agreement. With this background, it seems like a good point to introduce the suffix classes in further detail.

3 Introducing the suffix classes of Lealao Chinantec

3.1 A quick look at the classes

We have seen in Sect. 1 that the inflection of a verb in Lealao Chinantec includes a series of suffixes that Rupp (1989) analyzes as encoding the person-number of the subject. Rupp and Rupp (1996) propose that verbs fall into the four inflectional classes given above in (3) attending to the suffix set they select. The classes in question are given in full in (14) with examples. The class Rupp and Rupp (1996) treats as Class IV includes verbs that take \emptyset for the third person and others that take $-ʔ$. This formal difference suggests that this Class IV can be further subdivided into two different classes, which for the moment I treat as ‘Class IVa’ and ‘Class IVb’. All verbs in (14) appear in the future tense.¹⁰

(14)	$ʔi^2\tilde{n}uu^2a^2$ (tr) ‘tie something’	$ʔi^2heé^2a^2$ (tr) ‘run animate over’	$ʔi^2hmeéʔ^2a^2$ (tr) ‘treat animate’	$ʔi^4hií^4a^2$ (intr) ‘for animate to be ashamed’	$ʔi^4gwii^2a^2$ (intr) ‘for animate to be cold’
	Class I	Class II	Class III	Class IVa	Class IVb
1SG	$ʔi^2\tilde{n}uu-y^4$ ²	$ʔi^2-heé^2-a^4$	$ʔi^2-hmeéʔ^4-a^4$	$ʔi^4-hii-y^4$	$ʔi^4-gwii-y^4$
2SG	$ʔi^2\tilde{n}uu-y^3$	$ʔi^2-heé^2-u^3$	$ʔi^2-hmeé(y)ʔ^1$	$ʔi^4-hii^4-u^3$	$ʔi^4-gwii^2-u^3$
3SG	$ʔi^4\tilde{n}uu^4-\emptyset$	$ʔi^4-heé-y^2$	$ʔi^4-hmeéʔ^4-\emptyset$	$ʔi^4-hii^4-\emptyset$	$ʔi^4-gwii-ʔ^4$
1PL.INCL	$ʔi^2\tilde{n}uu^2-a^2$	$ʔi^2-heé^2-a^2$	$ʔi^2-hmeéʔ^2-a^2$	$ʔi^4-hii^4a^2$	$ʔi^4-gwii^32-a^2$
1PL.EXCL	$ʔi^2\tilde{n}uu^2-aʔ^1$	$ʔi^2-heé^2-aʔ^1$	$ʔi^2-hmeéʔ^2-aʔ^1$	$ʔi^4-hii^4aʔ^1$	$ʔi^4-gwii^32-aʔ^1$
2PL	$ʔi^2\tilde{n}uu^3-aʔ^3$	$ʔi^2-heé^2-aʔ^3$	$ʔi^2-hmeéʔ^1-aʔ^3$	$ʔi^4-hii^4-aʔ^3$	$ʔi^4-gwii^2-aʔ^3$
3PL	$ʔi^4-daʔ^2\tilde{n}uu^4-\emptyset$	$ʔi^4-daʔ^2-heé-y^2$	$ʔi^4-daʔ^2-hmeéʔ^4-\emptyset$	$ʔi^4-daʔ^2-hii^4-\emptyset$	$ʔi^4-daʔ^2-gwii-ʔ^4$

¹⁰Verbs fall into other inflectional classes attending to the marker they select for the future tense. A group of verbs (only considering dynamic verbs) take the prefix $ʔi^2-$ for all persons except the third person, marked by $ʔi^4-$. Another large group of verbs (including stative verbs but not only) take $ʔi^4-$ for all persons. A third class takes dsa^3- for all persons except for the third person, which is marked by dsa^4- .

Before analyzing the classes in (14) further, a couple of comments about the forms are in order:

- First, the suffix *-y* in forms such as $\text{ʔi}^2\tilde{n}uuy^{42}$ ‘I’ll tie something’ or $\text{ʔi}^4heéy^2$ ‘he will run somebody over’ is a palatal sonorant /j/ that does not receive tone.¹¹ It also occurs in a phonological position that Rupp (1990) treats as the ‘post-nuclear margin’.¹² When the stem has a glottalic nucleus /Vʔ/, the combination /Vʔj/ (e.g. $*hmee\text{ʔy}^1$ $*/m̥e:\text{ʔj}^1/$) is not phonotactically possible, and there is metathesis of the palatal element resulting in infixation (e.g. $hmee<y>\text{ʔ}^1$ $/m̥e:j\text{ʔ}^1/$).¹³
- The zero suffix in forms such as $\text{ʔi}^4\tilde{n}uu^4-\emptyset$ ‘he will tie something’ represents a stem (i.e. $\tilde{n}uu^4$) used by its own (i.e., bare) to encode a third person.
- Plural marking is invariant across the different classes: $-a^2$ for 1PL.INCL; $-a^1$ for 1PL.EXCL and $-a^3$ for 2PL. The form for 3PL is based on the stem used for 3SG and is derived from it by means of the number prefix $da\text{ʔ}^2-$ (itself the morphologization of the plural determiner $diá\text{ʔ}^4$ found in NPs).

All this leaves us with the singular forms in (15) (a slightly modified version of (3) above), as the only ones which are relevant for the analysis.

(15)

	Class I	Class II	Class III	Class IV.a	Class IV.b
1SG	-y	-á ⁴	-á ⁴	-y	-y
2SG	-y	-u ³	-y	-u ³	-u ³
3SG	-∅	-y	-∅	-∅	-ʔ

The classes in (15) result from the combination of the five affixes in (16) (counting $-\emptyset$ for convenience as an affix). Notice that in the mapping of form to meaning, only *-y* is used to encode all singular persons.¹⁴

(16)

	1SG	2SG	3SG
-y	✓	✓	✓
-á ⁴	✓	-	-
-u ³	-	✓	-
-ʔ	-	-	✓
-∅	-	-	✓

From the potential combinations of the encoding strategies in (16), we could in principle produce the 12 sets in (17), organized in four groups attending to the strategy

¹¹I follow in (14) Rupp and Rupp’s notion of tone as a superscript that appears at the end of the syllable. This notation can give the impression that tone 42 is associated with the suffix *-y* in forms such as $\text{ʔi}^2\tilde{n}uuy^{42}$ ‘I’ll tie something’, when in reality it is linked to the vocalic nucleus, i.e. $/ʔi^2ju:42j/$. I write tone associated with a vocalic nucleus when examples are segmented, i.e. $\text{ʔi}^2-\tilde{n}uu^{42}-y$.

¹²Palatal codas are very common in lexical words (e.g. $kway^1ná^2$ ‘result (n)’; $chíeéy^1$ ‘with’, etc.).

¹³The elements *-y* and *-ʔ* could be alternatively treated as stem extensions in the fashion of stem-based analyses such as for example the papers in Bonami (2012). However, for this particular case, such an analysis would perhaps not work on a distributional basis because the occurrence of these exponents is mutually exclusive with other more uncontroversial suffixes such as $-á^4$, $-u^3$, etc.

¹⁴The affix for 1SG has two shapes: $-á^4$ and $-á^2$. Although the rules behind their distribution appear to be mostly morphophonological, it is not entirely clear in Rupp and Rupp (1996) what the rules are for verbs. All the suffixes are used to mark possessor in nouns (p. 400). It appears that $-á^4$ is the default realization, whereas $-á^2$ is used when the stem has a tone /1/ or tone /42/.

Table 1 The suffix classes of Lealao Chinantec with an indication of population size

My notation:	1			2			3			4			5			6		
R&R classes:	I			II						III			IV.a			IV.b		
1SG	-y	-y	-y	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-y	-y	-y	-y	-y	-y
2SG	-y	-y	-y	-u ³	-u ³	-u ³	-y	-y	-y	-y	-y	-y	-u ³	-u ³	-u ³	-u ³	-u ³	-u ³
3SG	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ
Total	848	210	-	77	532	-	9	-	-	17	-	-	17	-	-	3	-	-
%	100%	24.7%	-	9.3%	62.7%	-	1%	-	-	2%	-	-	2%	-	-	0.3%	-	-

used to encode 1SG and 2SG (i.e., -y/-y; -á⁴/-u³; -á⁴/-y; and -y/-u³) along with three sets each attending to the marking of 3SG (i.e., -∅/-y/-ʔ). In the next section, I show that not all of these logical possibilities are indeed attested in the language.

(17)

1SG	-y	-y	-y	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-á ⁴	-y	-y	-y
2SG	-y	-y	-y	-u ³	-u ³	-u ³	-y	-y	-y	-u ³	-u ³	-u ³
3SG	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ	-∅	-y	-ʔ

3.2 The classes in more detail

My analysis of the suffix classes of Lealao Chinantec is based on the study of a sample of verbs from the dictionary by Rupp and Rupp (1996). This is a large dictionary with over 1500 verbal entries.¹⁵ For the purposes of the present study, I have considered only 1166 such entries. In reality, if we take each animacy pair in the dictionary to instantiate one and only one lexeme (i.e., not two), the sample consists of a total of 991 verbal lexemes. Not included in the sample are over a hundred compound verbs and a large number of stative and passive forms about which little information is given in the source.

Rupp and Rupp (1996) provide information about class membership for each of the 1166 entries, but for certain ones, they encode deviating behavior for 3SG. For example, for the entry of the verb *ʔi²chiaá⁴a²* (tr) ‘mix something’, while the dictionary informs the user that the verb is a Class II verb, there is further warning that it also inflects as a Class I for 3SG. This means that for a verb such as *ʔi²chiaá⁴a²*, 3SG has no affixal exponent, whereas other verbs of Class II would select -y. I take this to mean that the verb *ʔi²chiaá⁴a²* belongs to a different inflectional class from Class I or Class II. To avoid the treatment of this new class as a subclass of one of the other two, I have opted to rename Rupp and Rupp’s classes altogether this time using Arabic numbers. The new notation is given in Table 1. In the table, I also give the population size of each class from a subtotal of 848 entries out of the total of 1166 which can inflect for all persons (i.e., excluding verbs with inanimate or animal subjects). The reason why the paradigms in Table 1 are arranged in four blocks of three paradigms each is to keep it consistent with the array presented in (17).

Table 1 shows that only certain of the logical combinations in (17) are exploited in the inflection of Lealao Chinantec verbs. 97 % of the 848 verbs that inflect for all person values belong to one of three classes: 1, 2 or 3. This means that for 1SG/2SG

¹⁵These data will be freely accessible online from the spring of 2015 as part of the Surrey Morphology Group databases at the University of Surrey (<http://www.surrey.ac.uk/englishandlanguages/research/smg>).

the greater bulk of the verbs either select *-y* for both values or $-a^4$ and $-u^3$, respectively. The remaining three classes (Classes 4–6) are examples of heteroclitics (Stump 2006) because they are smaller subclasses of lexemes that recombine exponents used in other larger classes. One may also observe that 3SG is largely encoded without an exponent, that is, by using a bare stem, except for verbs of Class 3 and 6 which select *-y* and *-ʔ*, respectively. This means that if *-y* is used for the 3SG of a given verb, it can be predicted that the verb will always also select $-a^4$ and $-u^3$ for 1SG and 2SG (i.e. no verb allows *-y* to realize all person values). Additionally, more than half of the verbs in the sample belong to Class 3, which suggests that this class is a default from the point of view of population size.

Since Aronoff's (1994) seminal book on the autonomy of morphology, inflectional classes like the ones we observe in Table 1 are now seen as purely morphological constructs in word-and-paradigm approaches to morphology. Because of their morphological nature, inflectional classes are an interesting object of study for morphological theory, and they continue to inspire a body of work from different perspectives which is aimed at understanding their typology and internal structure, to mention just a few Ackerman et al. (2009), Baerman (2012, 2014b), Stump and Finkel (2013), Brown and Hippisley (2012, Chap. 3), Müller (2007) among others. Following this trend of thought and basing himself on the principles of canonical typology laid out in Corbett (2007), Corbett (2009) argues that a canonical inflectional class is not motivated by any other levels of grammar. This lack of motivation makes inflectional classes useless in functional terms, and in being so, they introduce a layer of morphological complexity into the inflectional system which speakers need to contend with (Baerman et al. 2010; Baerman 2013a). However, in order to be able to assert that classes such as the ones in Lealao Chinantec constitute a case of morphological allomorphy, one needs to explore possible motivations and evaluate the degree of their impact on the system as a whole. One obvious place to start is the phonology.

In this direction, I have not been able to find pairs of verbs belonging to the major classes with a homophonous stem. Even if I had (or if such cases existed), the homophony should in any case be handled with caution because the citation forms in the dictionary are themselves inflected forms (i.e. mostly 1st person plural future, but also 3rd person singular for inanimate verbs). As verbal inflection is made up of prosodic classes and stem alternations, the prosodic and segmental features in the entries themselves often reflect inflectional distinctions of their respective paradigms, and only after careful scrutiny one may end up having comparable phonological shapes to study prosodic features at a lexical level.

Nevertheless, one may still scan a representative list of partial homophones such as the list of entries in (18) to test if the phonology gives us clues as to the selection of at least some of the allomorphy in Table 1. For consistency, all examples in (18) represent 1st person plural future forms.

- (18) 1PL.FUT
- | | | | | |
|----|---|-----------------|--------------------------|---------|
| a. | γi^2 -chiaa ⁴² -a ² | 'We'll keep it' | (tr., inanimate object) | Class 1 |
| | γi^2 -chiaá ⁴ -a ² | 'We'll mix it' | (tr., inanimate object) | Class 2 |
| b. | γi^2 -ʔeeʔ ⁴² -a ² | 'We'll burp' | (intr., animate subject) | Class 1 |
| | γi^2 -ʔeéʔ ⁴ -a ² | 'We'll nail it' | (tr., inanimate object) | Class 2 |

c.	$\text{ʔi}^2\text{-niu}^2\text{-a}^2$	‘We’ll listen to it’	(tr., inanimate object)	Class 1
	$\text{ʔi}^2\text{-niu}^4\text{-a}^2$	‘We’ll mud it up’	(tr., inanimate object)	Class 2
d.	$\text{ʔi}^2\text{-taa}^{42}\text{-a}^2$	‘We’ll chop it’	(tr., inanimate object)	Class 1
	$\text{ʔi}^2\text{-taa}^3\text{-a}^2$	‘We’ll believe it’	(tr., inanimate object)	Class 2
e.	$\text{ʔi}^2\text{-tiaa}^2\text{-a}^2$	‘We’ll help him’	(tr., animate object)	Class 1
	$\text{ʔi}^2\text{-tiaa}^{42}\text{-a}^2$	‘We’ll mend it’	(tr., inanimate object)	Class 2
f.	$\text{ʔi}^2\text{-si}^4\text{-a}^2$	‘We’ll dip it in sauce’	(tr., inanimate object)	Class 1
	$\text{ʔi}^2\text{-si}^7\text{-a}^2$	‘We’ll gnaw it’	(tr., inanimate object)	Class 2
g.	$\text{ʔi}^2\text{-xaa}^7\text{-a}^2$	‘We’ll grab it for him’	(dtr., inanimate object)	Class 1
	$\text{ʔi}^2\text{-xaa}^{32}\text{-a}^2$	‘We’ll be rich’	(intr., animate subject)	Class 2

Table 2 Correlations involving stress type and forms for {1PL, FUT}

	Ballistic stress in {1PL, FUT}	Controlled stress in {1PL, FUT}	Total
Class 1	3 1.4%	207 98.6%	210 100%
Class 2	37 50%	37 50%	74 100%
Class 3	156 29.5%	372 70.5%	528 100%

In all the contrastive pairs in (18)—except those in (e) and (g)—the stem of a Class 2 verb carries ballistic stress (indicated by the acute accent), while all Class 1 stems carry control stress. How shall we interpret this? We can take stress distinction as a possible motivating factor for the allomorphy. Restricting ourselves for consistency to the stems in 812 citation forms instantiating 1st person plural future forms in the three major classes, the correlations involving ballistic/control stress we obtain are given in Table 2.

The figures in Table 2 suggest that if a verb belongs to Class 1, it has controlled stress at least for the form of the 1st person plural future. Alternatively, one can say the same differently. If the form for 1st person plural future has ballistic stress, the verb is NOT of Class 1 (except for three verbs). Beyond this point, phonology does not help us further.

I have pointed out in Sect. 2.2.2 that the morphology involved in the classes has to do with the dimension of animacy. In this light, the obvious next step to take is visit the sample of verbs according to their inflectional behavior with respect to this dimension. This is what I do in the next section where I show that there are indeed correlations, to such an extent that for verbs with two animacy paradigms there is a fully functional morphosyntactic mapping involving a mixed agreement system of person-number of subject S/A with animacy of S/O.

4 The verbs in the sample and their relation to animacy

The verbs in the sample could be said to fall into two main groups. In one group, we find verbs that can be inflected animate or inanimate attending to agreement in

Table 3 Verbs in the sample attending to their ability to enter in animacy pairs

	non-paired 816		paired 175 lexemes (=350 entries)	Total		Lex. Entries
	inanimate	animate	inanimate/animate	Lex.	Entries	
intr/S	226	212	53 106 entries	491	544	Lex. Entries
tr/O	257	121	122 244 entries	500	622	Lex. Entries
Total	483	333	175 350 entries	991	1,166	Lex. Entries

animacy with the S/O argument. I refer to such verbs as ‘paired-verbs’. The other group is formed by verbs that can only be animate or inanimate, which I call ‘non-paired verbs’. Table 3 gives an indication of the population size of the two groups in the sample attending to verbal entries.¹⁶

Let us see these verbs in more detail investigating their animacy value in relation to their prosodic inflection and the suffix class they select. I introduce first non-paired verbs where I show that the system points out remnants of a motivated core involving animacy agreement, which is still functional with paired verbs.

4.1 Non-paired verbs

4.1.1 Non-paired verbs, prosodic inflection and animacy

In general, when one studies these verbs it is difficult, if not utterly impossible, to determine whether a given one is inanimate or animate if one were to just look at the prosodic features of its inflection. This is because the verbs in Lealao Chinantec can fall into a very large number of prosodic classes, with many such classes having only one member. To show the (absence of) correlations between prosodic inflection and animacy, I compared the prosodic marking involved in all non-paired. Following Baerman and Palancar (forthcoming) that propose that in Chinantec verbal inflection it is convenient to deal with 3rd person forms independently from the inflection of other persons, I have treated the different classes involved in the marking of a 3rd person apart from the classes used to encode other person values (i.e., 1SG, 1PL, 2). The actual forms used for these many prosodic classes are given in the Appendix. The verbs in classes labeled /1/, /’2/, etc. all belong to one large class which does not have prosodic inflection. The notation for these verbs indicate the actual phonology of the stem—the apostrophe indicates ballistic stress. I have included them here to show that lexical phonology of a given stem is not really informative about animacy value. I use shading in the table to highlight prosodic classes aligned with a specific animacy value.¹⁷

¹⁶There are a few entries in the dictionary I have opted to classify as non-paired verbs because the stem of the animate entry has a prefix *bi*³- and appears derived, e.g. *?i²-bi³-na⁴a²* ‘for someone to be absent’ vs. *?i⁴na⁴* ‘something lacking’; or *?i²-bi³-ca⁴a¹a²* ‘play with someone’ vs. *?i⁴caa⁴a²* ‘play with something’; *?i²-bi³-?in⁴a²* ‘forget someone’ vs. *?i⁴-?in⁴* ‘forget something’; or *?i²-bi³-fi⁴a²* ‘increase the amount of an animate entity’ vs. *?i²-fi⁴a²* ‘pile up something’. This prefix also appears in the stems of some paired verbs.

¹⁷The angle bracket > indicates an ‘act on’-type of transitive relation between an A argument and an O argument, it should for example be read as ‘1SG/2/1PL act on INAN O’.

Table 4 Prosodic classes in non-paired verbs contrasting inanimate vs. animate

Cl.	3		Cl.	1SG/2/1PL		Cl.	1SG/2/1PL	
	INAN	AN		>INAN	>AN		>INAN	>AN
<i>a</i>		2	1	20	24	37	1	
<i>b</i>	2		2	9	7	38	1	1
<i>c</i>	1	2	3	3	1	39		2
<i>d</i>	2	2	4		1	40	2	
<i>e</i>	1		5	40	17	41		2
<i>f</i>		1	6	12	33	42	1	
<i>g</i>	25	6	7	22	4	43		1
<i>h</i>	1	1	8	17	16	44		2
<i>i</i>		1	9	16	2	45	2	
<i>j</i>		1	10	17	1	46	1	
<i>k</i>	2	1	11	3	8	47		1
<i>l</i>	115	53	12	16	16	48		1
<i>m</i>	1		13	1	3	49	1	
<i>n</i>	4	1	14	9	9	50		1
<i>o</i>	9	1	15	1	5	51		1
<i>p</i>		1	16	3	2	52	1	
<i>q</i>	32	15	17	4	3	53	1	
<i>r</i>	1	1	18	2	2	54	1	
/1/	4	10	19	3	0	55	1	
/1'/	14	17	20	7	2	56	1	
/2/	10	1	21	5	2	57		1
/2'/	23	13	22	1	1	58	1	
/3/	48	41	23		5	59		1
/3'/	70	46	24		1	60		1
/32/	2	12	25		5	61	1	
/4/	37	35	26	3	2	<i>q</i>	1	2
/4'/	72	63	27		3	<i>l</i>	1	14
/42/	7	6	28		3	/1/	1	9
Total	483	333	29	4		/1'/	2	4
			30	2	2	/2/		1
			31	1	2	/2'/		6
			32	1		/3/	3	18
			33	2	1	/32/	1	11
			34		1	/4/	23	36
			35	1		/42/	3	3
			36	2		(without non-3	230	30)
						Total	483	333

Table 4 shows that verbs belong to 19 classes to inflect for a 3SG, but many such classes have only one verb. Excluding those, there are only two (i.e., *a* and *b*) which are specific to inanimates but they have only two members each, so again they are utterly irrelevant. We can conclude that the form for the 3rd person does not help much in predicting animacy. Exactly the same can be broadly said for the forms inflected in other persons of which we have up to 64 classes, most of them instantiated by one or two verbs. Classes 12 (with 16 members) and 14 (with nine) are the most

Table 5 The 816 non-paired verbs of the sample

		Cl. 3	?	Cl. 1	Cl. 2	Cl. 4	Cl. 5	Cl. 6
Animates	Intransitive							
	1SG	- -á ⁴	-	-y	-á ⁴	-á ⁴	-y	-y
	2SG	- -u ³	-	-y	-u ³	-y	-u ³	-u ³
	3SG.AN	-y -y	-Ø	-Ø	-Ø	-Ø	-Ø	-?
	212 100%	17 142 75.5%	11 5%	23 11%	10 4%	-	6 3%	3 1.5%
Animates	Transitive							
	1SG> AN	- -á ⁴		-y	-á ⁴	-á ⁴	-y	-y
	2SG> AN	- -u ³		-y	-u ³	-y	-u ³	-u ³
	3SG> AN	-y -y		-Ø	-Ø	-Ø	-Ø	-?
	121 100%	2 110 92.5%	-	2 1.5%	7 6%	-	-	-
Inanimates	Intransitive							
	1SG	-		-	-	-	-	-
	2SG	-		-	-	-	-	-
	3SG.INAN	-y			-Ø			-?
	226 100%	65 29%		161 71%				-
Inanimates	Transitive		?					
	1SG> INAN	- -á ⁴	-	-y	-á ⁴	-á ⁴	-y	-y
	2SG> INAN	- -u ³	-	-y	-u ³	-y	-u ³	-u ³
	3SG> INAN	-y -y	-Ø	-Ø	-Ø	-Ø	-Ø	-?
	257 100%	1 110 43%	3 1%	107 42%	28 11%	-	8 3%	-

notable exception as they only include animates. The system does not seem to like absolutes, but there are tendencies. For example a verb that inflects like a class *l* for the 3rd person is likely to be inanimate, but if the verb inflects for the other persons in this class, it is likely to be animate. Likewise, it is highly probable that a verb with no prosodic inflection whose stem bears the prosodic feature /'3/ is animate.

If prosody is not a good indicator of animacy value, this opens the possibility that the morphology in the suffix classes is perhaps used for this purpose. I entertain this hypothesis in the following sections, where I claim that this is precisely what happens, at least with animate verbs; inanimate ones appear to have a more canonical inflectional class distribution.

4.1.2 Non-paired verbs and the suffix classes

The distribution of the 816 non-paired attending to class membership is given in Table 5. In order to present the data in a more convenient way, I have extracted Class 3 to the left.

The data in the Table 5 show two different, but related things: (i) animate verbs have a clear preference to belong to Class 3; and (ii) many of the members of Class 1 are inanimate verbs. This could suggest that the suffixes of these classes (-á⁴, -u³, -y and -y, -y, -Ø) are used to indicate animacy information about their respective S/O arguments in such a way that in an ideal state of affairs, Class 3 would say something about S/O being animate and Class 1, used with transitives, that O is

inanimate. However, a closer look at the table further reveals that in reality things are a bit messier: (iii) almost half of the inanimate verbs also belong to Class 3; and (iv) a score of animate verbs are also found in Class 1. This situation suggests that the system we observe in Lealao Chinantec has some motivated core, while still displaying inflectional class organization. In this, the Chinantec system is typologically very similar to other better known inflectional class systems found in Indo-European languages which are partly motivated by gender. I address this commonality in Sect. 5.2.

Before observing paired verbs in the next section and proposing what the morphosyntax of this core is in Sect. 5.1, I think it is first worth studying whether the exceptions in (iii) and (iv) could be accounted for by other means. To address the issue raised in (iii), I study the 49 intransitive animate verbs that use a bare stem to encode a 3rd person—largely a coding strategy for inanimates—under the hypothesis that perhaps these verbs have some semantic property in common that make them stand out from the rest of animate verbs. For this purpose consider Table 6, where I have arrayed such verbs in different semantic fields (like most animate verbs, the verbs in question appear in their citation form in the 1st person plural future).

When one observes the semantics of the verbs in Table 6, it is difficult to come up with a semantic property that unites all these lexemes into one coherent group. It is true that some of those in Class 1 depict activities some of which are typically found in fluid systems with semantic alignment of the agentive/patientive type, e.g. ‘laugh’, ‘burp’, ‘sneeze’, etc. Others are more agentive ‘talk’, ‘whistle’, ‘shout’, ‘play’, etc. But similar verbs are also largely found in Class 3. It is equally difficult to predict how an intransitive animate verb will be inflected if we were to be led only by the semantics, as verbs with very similar meanings are found in different classes (e.g. $\text{ʔi}^2\text{laa}^2\text{a}^2$ ‘lie down’, $\text{ʔi}^2\text{hi}^2\text{a}^2$ ‘weep’ and $\text{ʔi}^4\text{cuin}^2\text{a}^2$ ‘escape’ are of Class 1, but $\text{ʔi}^2\text{bi}^3\text{i}^3\text{ʔ}^3\text{a}^2$ ‘lie crooked’ and $\text{ʔi}^2\text{ʔe}^4\text{en}^4\text{a}^2$ ‘moan’ and $\text{ʔi}^2\text{bi}^3\text{re}^4\text{a}^2$ ‘hide’ are of Class 2, while having very similar semantic). Admittedly, the first column in Table 6 could give the impression that verbs referring to actions performed by animals receive $-\emptyset$ instead of $-y$ for the 3rd person, but there are also verbs depicting animal actions that belong to Class 3, as for example the ones in Table 7.

Finally, if such verbs were to be inflected in 1SG or 2SG (as characters speaking in a fable), it is not entirely sure what the suffix selected would be. The list in Table 6 suggests that it would be by means of $-y$ for either value, imagining the verb could be interpreted as a Class 1 verb. But the existence of the verb $\text{dsa}^3\text{gue}^4\text{a}^2$ ‘crawl’ in Class 2 (which we can compare with the verb dsa^4ngi^4 ‘for a baby to be able to walk’) together with the list in Table 7 further suggests that it could just as well be by means of $-\acute{a}^4$ and $-u^3$. From all this, we can conclude that semantics does not help in determining class membership, at least for the intransitive verbs considered.

Before moving on to paired verbs in the next section, let us address issue (iv) by considering the inanimate verbs that belong to Class 3. For these verbs, we could entertain the hypothesis that their receiving suffixes $-\acute{a}^4$, $-u^3$, $-y$ is because their prosodic inflection is similar to the animate verbs of Class 3, but different from those of Class 1. We have seen in Table 4 above that prosodic inflection is largely uninformative about animacy, but it could happen that verbs that receive the same prosodic

Table 6 The 49 intransitive animate non-paired verbs that do not belong to Class 3

	Class 1 -y, -y, -Ø	Class 2 -a ⁴ , -ir ³ , -Ø	Class 5 -y, -ir ³ , -Ø	
Miscellanea	<p>Class ? -, -, -Ø</p> <p>ɣ⁴ bɪ⁴ ɣaɑ⁴ɣ⁴ bellow ɣ⁴ bɪ⁴ ɣ⁴ ɣ⁴ ɣ⁴ squeak (squirrel) ɣ⁴ hmiɪ³ dig (a den) ɣ⁴ ɣ⁴ ɣ⁴ gobble (a turkey) ɣ⁴ ɣ⁴ ɣ⁴ neigh ɣ⁴ h⁴ roar ɣ⁴ hmaɑ² fly</p>	<p>ɣ² ɣ² ɣ² ɣ² ɣ² a² laugh ɣ² ɣ² ɣ² ɣ² a² burp ɣ⁴ cue⁴ a² sneeze ɣ² ju ɣ⁴ a² cough ɣ² bɪ³ jmiɪ² a² fast ɣ² ɣ² ɣ² ɣ² a² bathe ɣ² ɣ² ɣ² a² whistle ɣ² ɣ² ɣ² a² blow one's nose ɣ² ɣ² ɣ² a² talk ɣ² ɣ² ɣ² a² call ɣ² ɣ² ɣ² a² shout ɣ⁴ caa² a² play ɣ² bɪ³ ɣ² ɣ² a² wait ɣ² hi ɣ² a² weep ɣ² bɪ³ ɣ² ɣ² a² presume ɣ² dxiu² a² complain</p>	<p>owe be rich</p>	<p>ɣ⁴ saɑ ɣ⁴ a² drown ɣ⁴ hmiɑ⁴ a² be able to see</p>
Emotional	<p>ɣ² ɣ² ɣ² ɣ² a² moan ɣ² bɪ³ ɣ² ɣ² a² feel encouraged ɣ⁴ dxii ɣ⁴ a² be jealous</p>	<p>ɣ² ɣ² ɣ² ɣ² a² lie crooked ɣ² bɪ³ ɣ² ɣ² a² be standing (SG) ɣ⁴ crawl</p>	<p>moan feel encouraged be jealous</p>	<p>ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ a² be afraid ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ a² be embarrassed ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ a² get used to ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ a² be disgusted</p>
Posture	<p>ɣ² ɣ² ɣ² ɣ² a² sit, stand (birds, insects)</p>	<p>ɣ² ɣ² ɣ² ɣ² a² stand up ɣ² ɣ² ɣ² ɣ² a² sit</p>	<p>lie down stand up sit</p>	<p>lie crooked be standing (SG) crawl</p>
Motion	<p>ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ walk (babies) ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ fly</p>	<p>ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ go up a level higher than the ground ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ go downhill ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ go uphill ɣ⁴ ɣ⁴ ɣ⁴ ɣ⁴ escape</p>	<p>go up a level higher than the ground go downhill go uphill escape</p>	<p>hide</p>
Location	<p>ɣ⁴ ɣ⁴ be inside</p>			

Table 7 Animate intransitive verbs of Class 3 referring to actions of animals

Class 3 -, -, -y	
ʔi ⁴ bi ⁴ ci ⁴ y ¹	growl (animal)
ʔi ⁴ féy ⁴	flap wings
ʔi ⁴ nieéy ¹	be tied up (an animal)
ʔi ⁴ ná ⁴ chieey ^{ʔ42}	gather up together (animals)
ʔi ⁴ ná ⁴ cuaay ^{ʔ42}	gather up together (animals)
ʔi ⁴ cíy ^{ʔ42}	fall (animal)
ʔi ⁴ yá ⁴ niuy ^{ʔ4}	go around broody (a hen)
ʔi ⁴ nieey ⁴	shed skin (a serpent)
ʔi ⁴ vi ⁴ y ^{ʔ2}	be born (of an egg)
dsá ⁴ xú ⁴ láy ^{ʔ4}	scratch off (animal)
ʔi ⁴ sii ³	be sullen (animal)
ʔi ⁴ chiéy ^{ʔ4}	shake (animal)

inflection would attract similar sets of suffixes. If this were the case, the selection of Class 3 suffixes by inanimate verbs would be accounted for as being conditioned by the type of prosodic inflection the verb receives, and hence it would be linked to the phonology of the inflected stem. For this purpose, I have compared the prosodic inflection of the 110 inanimate transitive verbs of Class 3 on the one hand with the 105 inanimate transitive verbs of Class 1 and on the other with the 109 animate transitive verbs of Class 3. The results are given in Table 8.

Table 8 shows that there are formal similarities between the animate and the inanimate verbs of Class 3. This is puzzling because prosodic inflection appears to align with animacy on the one hand, while on the other it aligns with class. In this respect, there is for example a tendency for the inanimate verbs of Class 1 to take a different inflection than the animate verbs of Class 3, and hence prosody aligns with animacy value. But the verbs of Class 3 that are inanimate take a similar inflection than animate verbs, and hence prosody aligns with class. One way to understand this situation and the reason why we have transitive inanimate verbs in Class 3 is to regard this class as a default (in the sense of the ‘elsewhere’ case) that bears its own distinctive prosodic properties associated with it. In this light, Class 1 would have to be seen as a subclass within this major class (see Sect. 5.2 for more arguments in favor of Class 3 as a default).

4.2 Paired verbs

4.2.1 Paired verbs, prosodic inflection and animacy

Paired-verbs are verbs that inflect as both animate and inanimate depending on the animacy of their S/O argument. In Rupp and Rupp’s (1996) dictionary, I have been able to identify 175 such verbs: 53 intransitive and 122 transitive. For the intransitive verbs, only the forms for the 3rd person are comparable, because in natural usage the inanimate paradigm of the verb cannot have forms for other persons. To inflect for

Table 8 Transitive verbs of Class 1 and 3 compared for prosodic features

Cl.	3			Cl.	1SG/2/1PL			Cl.	1SG/2/1PL		
	INAN		AN		INAN		AN		INAN		AN
	Cl. 1	Cl.3			Cl. 1	Cl.3			Cl. 1	Cl.3	
a			2	1		14	22	31		1	
b	2			2		4	7	32	1		
d		2	1	3		3	1	35		1	
g	23			5		39	16	36	2		
k	2			6		11	13	37	1		
l	53	36	29	7	22			40	2		
n			1	8	6	11	1	41			2
o	9			9	16			42	1		
q	4	1	8	10	17			45	2		
r		1		11		3	6	46	1		
/1/		1	2	13			3	49	1		
/1'/	1	3	2	14				53	1		
/2/	2			15		1	5	54	1		
/2'/	1	4	6	16	3			55	1		
/3/	3	15	18	17	4			56	1		
/3'/	3	16	23	18		2	2	58	1		
/32/		2	1	19	3			q		1	
/4/		6	9	20	7			l		1	1
/4'/	4	20	6	21	5		2	/1/		1	2
/42/		3	2	24			1	/1'/		2	1
Total	107	110	110	25			5	/3/		2	1
				26	3			/32/		1	1
				28			2	/4/	1	7	4
				29	4			/42/		3	2
				30		2	1				
								Total	107	110	110

animacy agreement, a verb often belongs to different prosodic classes simultaneously. When it does, there are two main patterns.

In the first pattern, the animate paradigm has different tonal inflection than the inanimate one. There are 67 paired verbs that behave like this. The example in (19) illustrates an intransitive verb with its two paradigms, and that in (20) a transitive one.

(19) $\text{ʔ}^4\text{tia}\text{ʔ}^3$ (intr) ‘fall from height’

a. INAN ‘for something to fall from height’

	PRS	PRF	FUT
3SG.INAN	$\text{tia}\text{ʔ}^3$	$\text{ma}^3\text{-tia}\text{ʔ}^3$	$\text{ʔ}^4\text{-tia}\text{ʔ}^3$

INFL STEMS

INCPL	CPL	IRR
$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^3$

b. AN ‘for animate entity to fall from height’

	PRS	PRF	FUT
1SG	$\text{tia}\text{ʔ}^4\text{-}\acute{\text{a}}^4$	$\text{ma}^3\text{-tia}\acute{\text{a}}\text{ʔ}^3\text{-}\acute{\text{a}}^4$	$\text{ʔ}^4\text{-tia}\acute{\text{a}}\text{ʔ}^4\text{-}\acute{\text{a}}^4$
2SG	$\text{tia}\text{ʔ}^4\text{-u}^3$	$\text{ma}^3\text{-tia}\acute{\text{a}}\text{ʔ}^3\text{-u}^3$	$\text{ʔ}^4\text{-tia}\acute{\text{a}}\text{ʔ}^4\text{-u}^3$
3SG.AN	$\text{tia}\langle\text{y}\rangle\text{ʔ}^4$	$\text{ma}^3\text{-tia}\langle\text{y}\rangle\text{ʔ}^3$	$\text{ʔ}^4\text{-tia}\langle\text{y}\rangle\text{ʔ}^4$
1PL.INCL	$\text{tia}\text{ʔ}^4\text{-a}^2$	$\text{ma}^3\text{-tia}\acute{\text{a}}\text{ʔ}^3\text{-a}^2$	$\text{ʔ}^4\text{-tia}\acute{\text{a}}\text{ʔ}^4\text{-a}^2$

INFL STEMS

INCPL	CPL	IRR
$\text{tia}\text{ʔ}^4$	$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^4$
$\text{tia}\text{ʔ}^4$	$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^4$
$\text{tia}\text{ʔ}^4$	$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^4$
$\text{tia}\text{ʔ}^4$	$\text{tia}\text{ʔ}^3$	$\text{tia}\text{ʔ}^4$

(20) $\text{ʔ}^i\text{lia}^{\text{ʔ}^2}$ (tr) ‘push’

a. INAN ‘push something’

	INFL STEMS			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG>INAN	lia<y> ʔ^4	ma ³ -liá<y> ʔ^4	$\text{ʔ}^i\text{ʔ}^2$ -lia<y> $\text{ʔ}^{\text{ʔ}^2}$	lia ʔ^4	liá ʔ^4	lia $\text{ʔ}^{\text{ʔ}^2}$
2SG>INAN	lia<y> ʔ^3	ma ³ -liá<y> ʔ^1	$\text{ʔ}^i\text{ʔ}^2$ -lia<y> ʔ^3	lia ʔ^3	liá ʔ^1	lia ʔ^3
3SG>INAN	lia ʔ^4	ma ³ -liá ʔ^3	$\text{ʔ}^i\text{ʔ}^4$ -liá ʔ^4	lia ʔ^4	liá ʔ^3	liá ʔ^4
1PL>INAN	lia $\text{ʔ}^{\text{ʔ}^2}$	ma ³ -lia ʔ^2	$\text{ʔ}^i\text{ʔ}^4$ -lia ʔ^2	lia $\text{ʔ}^{\text{ʔ}^2}$	liá ʔ^2	lia ʔ^2

b. AN ‘push animate’

	INFL STEMS			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG>AN	lia $\text{ʔ}^{\text{ʔ}^2}$ -á ʔ^4	ma ³ -liá $\text{ʔ}^{\text{ʔ}^2}$ -á ʔ^4	$\text{ʔ}^i\text{ʔ}^2$ -lia ʔ^1 -á ʔ^4	lia $\text{ʔ}^{\text{ʔ}^2}$	liá ʔ^2	lia ʔ^1
2SG>AN	lia ʔ^4 -u ³	ma ³ -liá $\text{ʔ}^{\text{ʔ}^2}$ -u ³	$\text{ʔ}^i\text{ʔ}^2$ -lia ʔ^4 -u ³	lia ʔ^4	liá $\text{ʔ}^{\text{ʔ}^2}$	lia ʔ^4
3SG>AN	lia<y> $\text{ʔ}^{\text{ʔ}^2}$	ma ³ -liá<y> ʔ^2	$\text{ʔ}^i\text{ʔ}^4$ -liá<y> ʔ^2	lia $\text{ʔ}^{\text{ʔ}^2}$	liá ʔ^2	liá $\text{ʔ}^{\text{ʔ}^2}$
1PL>AN	lia $\text{ʔ}^{\text{ʔ}^2}$ -a ²	ma ³ -liá $\text{ʔ}^{\text{ʔ}^2}$ -a ²	$\text{ʔ}^i\text{ʔ}^2$ -lia $\text{ʔ}^{\text{ʔ}^2}$ -a ²	lia $\text{ʔ}^{\text{ʔ}^2}$	liá $\text{ʔ}^{\text{ʔ}^2}$	liá $\text{ʔ}^{\text{ʔ}^2}$

The inanimate paradigm of the verb in (19) belongs to the invariant class and has tone /3/ for all cells. In contrast, its animate paradigm belongs to Class g for the marking of 3rd person but to Class 20 for the marking of other persons, we can represent this as {g/20}. The verb in (20) is {g/7} for the inanimate and {l/ll} for the animate. The multiple permutations give rise to a system with many inflectional possibilities. For the values of such classes see the Appendix. The 67 paired verbs in the sample of this first pattern display 41 such combinations, which are given in Table 9. The combinations pertaining to verbs (19) and (20) are highlighted in shading.

The second pattern is instantiated by 42 paired verbs. In this pattern, the forms for the 3rd person inanimate and animate are prosodically identical, while the forms for the other persons remain contrastive. An example is given in (21).

(21) $\text{ʔ}^i\text{c}^{\text{ʔ}^1}\text{a}^2$ (tr) ‘knock over’

a. INAN ‘knock something over’

	INFL STEMS			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG>INAN	cí<y> ʔ^4	ma ³ -cí<y> ʔ^4	$\text{ʔ}^i\text{ʔ}^4$ -cí<y> ʔ^4	cí ʔ^4	cí ʔ^4	cí ʔ^4
2SG>INAN	cí<y> $\text{ʔ}^{\text{ʔ}^2}$	ma ³ -cí<y> ʔ^2	$\text{ʔ}^i\text{ʔ}^4$ -cí<y> ʔ^1	cí $\text{ʔ}^{\text{ʔ}^2}$	cí ʔ^2	cí ʔ^1
3SG>INAN	cí $\text{ʔ}^{\text{ʔ}^2}$	ma ³ -cí ʔ^2	$\text{ʔ}^i\text{ʔ}^4$ -cí ʔ^2	cí $\text{ʔ}^{\text{ʔ}^2}$	cí ʔ^2	cí ʔ^2
1PL>INAN	cí $\text{ʔ}^{\text{ʔ}^2}$ -a ²	ma ³ -cí ʔ^1 -a ²	$\text{ʔ}^i\text{ʔ}^4$ -cí ʔ^1 -a ²	cí $\text{ʔ}^{\text{ʔ}^2}$	cí ʔ^1	cí ʔ^1

b. AN ‘knock animate over’

	INFL STEMS			INFL STEMS		
	PRS	PRF	FUT	INCPL	CPL	IRR
1SG>AN	cí $\text{ʔ}^{\text{ʔ}^2}$ -á ʔ^4	ma ³ -cí $\text{ʔ}^{\text{ʔ}^2}$ -á ʔ^4	$\text{ʔ}^i\text{ʔ}^4$ -cí $\text{ʔ}^{\text{ʔ}^2}$ -á ʔ^4	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$
2SG>AN	cí $\text{ʔ}^{\text{ʔ}^2}$ -u ³	ma ³ -cí $\text{ʔ}^{\text{ʔ}^2}$ -u ³	$\text{ʔ}^i\text{ʔ}^4$ -cí $\text{ʔ}^{\text{ʔ}^2}$ -u ³	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$
3SG>AN	cí $\text{ʔ}^{\text{ʔ}^2}$	ma ³ -cí ʔ^2	$\text{ʔ}^i\text{ʔ}^4$ -cí ʔ^2	cí $\text{ʔ}^{\text{ʔ}^2}$	cí ʔ^2	cí ʔ^2
1PL>AN	cí $\text{ʔ}^{\text{ʔ}^2}$ -a ²	ma ³ -cí $\text{ʔ}^{\text{ʔ}^2}$ -a ²	$\text{ʔ}^i\text{ʔ}^4$ -cí $\text{ʔ}^{\text{ʔ}^2}$ -a ²	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$	cí $\text{ʔ}^{\text{ʔ}^2}$

The verb in (21) belongs to Class {l, 10} when inanimate and {l, 5} when animate. The 42 verbs display the 17 different inflectional possibilities given in Table 10. One of such classes has up to 16 members.

The remaining 66 paired verbs in the sample do not change in prosodic inflection when they are inflected as animate or inanimate. Relevant examples are given in (22) and (23).

Table 9 Prosodic inflection in the 67 paired verbs of the first pattern

#	INAN		AN		#	INAN		AN	
	3	1SG/2/1PL	3	1SG/2/1PL		3	1SG/2/1PL	3	1SG/2/1PL
1	<i>b</i>	42	/32/	68	1	<i>q</i>	—	/4/	8
1	<i>g</i>	7	/3/	5	3	<i>q</i>	—	/4/	/4/
1	<i>g</i>	7	/4/	6	1	<i>q</i>	—	<i>g</i>	7
2	<i>g</i>	7	<i>l</i>	5	1	/2/	—	/1/	8
4	<i>g</i>	7	<i>l</i>	11	2	/2/	—	/4/	6
3	<i>g</i>	7	<i>l</i>	14	1	/2/	—	/4/	/4/
1	<i>g</i>	7	<i>l</i>	24	1	/2/	—	<i>l</i>	<i>l</i>
3	<i>g</i>	7	<i>q</i>	11	4	/3/	—	/4/	/4/
1	<i>g</i>	37	<i>l</i>	5	1	/3/	—	<i>g</i>	20
2	<i>k</i>	16	/32/	22	3	/3/	22	/4/	6
1	<i>k</i>	16	/32/	68	1	/3/	27	/4/	/4/
1	<i>k</i>	17	/32/	22	1	/3/	63	<i>l</i>	63
1	<i>l</i>	10	/3/	18	1	/3/	—	/4/	6
1	<i>l</i>	10	/4/	8	1	/32/	31	/3/	18
3	<i>l</i>	19	/3/	5	1	/4/	6	/3/	6
2	<i>l</i>	32	<i>q</i>	25	1	/4/	—	/1/	8
2	<i>o</i>	16	<i>q</i>	13	5	/4/	/4/	/3/	5
1	<i>o</i>	20	<i>q</i>	13	1	/4/	13	/3/	5
1	<i>q</i>	—	/2/	/2/	1	/4/	22	<i>l</i>	21
1	<i>q</i>	—	/4/	66	2	/4/	26	/3/	22
					1	/4/	65	/2/	13
Total					67				

Table 10 Prosodic inflection in the 42 paired verbs of the second pattern

#	INAN		AN	
	3	1SG/2/1PL	3	1SG/2/1PL
2	/2/	8	/2/	2
1	/2/	67	/2/	2
1	<i>l</i>	19	<i>l</i>	5
1	<i>l</i>	21	<i>l</i>	5
2	<i>l</i>	10	<i>l</i>	5
1	/3/	27	/3/	6
2	/3/	22	/3/	6
16	<i>l</i>	9	<i>l</i>	11
1	<i>l</i>	64	<i>l</i>	11
1	<i>l</i>	10	<i>q</i>	13
1	<i>q</i>	10	<i>q</i>	13
7	<i>l</i>	10	<i>l</i>	15
1	/3/	16	/3/	18
2	<i>l</i>	32	<i>l</i>	24
1	/3/	61	/3/	24
1	<i>r</i>	34	<i>r</i>	35
1	/32/	26	/32/	62
Total 45				

(22)

ʔi⁴gwiiʔ⁴ (intr) ‘get cold’ {/4/, 6}

a. INAN ‘for something to get cold’

				INFL STEMS		
	PRS	PRF	FUT	INCL	CPL	IRR
3SG>INAN	gwiiʔ ⁴	ma ³ -gwiiʔ ⁴	ʔi ⁴ -gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴

b. AN ‘for animate to get cold’

				INFL STEMS		
	PRS	PRF	FUT	INCL	CPL	IRR
1SG	gwiiʔ ⁴ -á ²	ma ³ -gwiiʔ ⁴ -á ²	ʔi ⁴ -gwiiʔ ⁴ -á ²	gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴
2SG	gwiiʔ ⁴ -u ³	ma ³ -gwiiʔ ⁴ -u ³	ʔi ⁴ -gwiiʔ ⁴ -u ³	gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴
3SG>AN	gwiiʔ ⁴ -y	ma ³ -gwiiʔ ⁴ -y	ʔi ⁴ -gwiiʔ ⁴ -y	gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴
1PL>INCL	gwiiʔ ⁴ -a ²	ma ³ -gwiiʔ ⁴ -a ²	ʔi ⁴ -gwiiʔ ⁴ -a ²	gwiiʔ ⁴	gwiiʔ ⁴	gwiiʔ ⁴

(23)

ʔi²gwaʔ⁴²a² (tr) ‘change’ {/l, 5}

a. INAN ‘change something’

				INFL STEMS		
	PRS	PRF	FUT	INCL	CPL	IRR
1SG>INAN	gwa<y>ʔ ⁴²	ma ³ -gwa<y>ʔ ⁴²	ʔi ² -gwa<y>ʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²
2SG>INAN	gwa<y>ʔ ⁴²	ma ³ -gwa<y>ʔ ⁴²	ʔi ² -gwa<y>ʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²
3SG>INAN	gwaʔ ³²	ma ³ -gwaʔ ³²	ʔi ⁴ -gwaʔ ³²	gwaʔ ³²	gwaʔ ³²	gwaʔ ³²
1PL>INAN	gwaʔ ³² -a ²	ma ³ -gwaʔ ³² -a ²	ʔi ² -gwaʔ ³² -a ²	gwaʔ ³²	gwaʔ ³²	gwaʔ ³²

b. AN ‘change entity’

				INFL STEMS		
	PRS	PRF	FUT	INCL	CPL	IRR
1SG>AN	gwa<y>ʔ ⁴² -á ²	ma ³ -gwa<y>ʔ ⁴² -á ²	ʔi ² -gwa<y>ʔ ⁴² -á ²	gwaʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²
2SG>AN	gwa<y>ʔ ⁴² -u ³	ma ³ -gwa<y>ʔ ⁴² -u ³	ʔi ² -gwa<y>ʔ ⁴² -u ³	gwaʔ ⁴²	gwaʔ ⁴²	gwaʔ ⁴²
3SG>AN	gwa<y>ʔ ³²	ma ³ -gwa<y>ʔ ³²	ʔi ⁴ -gwa<y>ʔ ³²	gwaʔ ³²	gwaʔ ³²	gwaʔ ³²
1PL>AN	gwaʔ ³² -a ²	ma ³ -gwaʔ ³² -a ²	ʔi ² -gwaʔ ³² -a ²	gwaʔ ³²	gwaʔ ³²	gwaʔ ³²

In this section, we have seen that at least two thirds of the paired verbs of the sample inflect as animate and inanimate in rather unpredictable ways. Like with non-paired verbs, if prosody does not work as an optimal way to inform about the animacy of a verb, it makes sense—at least from a morphosyntactic point of view—that other parts of the system do the job. In the following section, I show that the suffix sets of the major suffix classes (i.e., Classes 1, 2 and 3) have a clear morphosyntactic function in the realization of animacy agreement with paired-verbs, which are the ones that have more need of it for inflectional purposes.

Finally, alternative views of paired verbs are possible. For example, one could claim that they represent two independent lexemes related semantically through derivation. Under such a view the morphological irregularities they display would be accounted for as being in effect typical of derivation, while verbs in (22) and (23) would have to be seen as ‘ambi-animate’ by undergoing some sort of zero conversion from animate to inanimate and vice versa. But if the verbs were to be seen as related through derivation, the fact that the suffixes display a clearer mapping from function to form precisely with them would still want an explanation. This is why the treatment of paired verbs as inflectional has advantages for the analysis of the suffixes. Nevertheless, appealing to derivation as an argument to account for the mor-

Table 11 The 175 paired verbs of the sample

Animates	Intransitive		Cl. 3	Cl. 1	Cl. 2	Cl. 4	Cl. 5	Cl. 6	
		1SG	-	-á ⁴	-y	-á ⁴	-á ⁴	-y	-y
		2SG	-	-u ³	-y	-u ³	-y	-u ³	-u ³
		3SG.AN	-y	-y	-Ø	-Ø	-Ø	-Ø	-?
	53		5	47	1				
	100%		98%	2%					
Animates	Transitive								
		1SG> AN	-á ⁴	-y	-á ⁴	-á ⁴	-y	-y	
		2SG> AN	-u ³	-y	-u ³	-y	-u ³	-u ³	
		3SG> AN	-y	-Ø	-Ø	-Ø	-Ø	-?	
	122		122						
	100%		100%						
Inanimates	Intransitive		Cl. 3	Cl. 1	Cl. 2	Cl. 4	Cl. 5	Cl. 6	
		1SG	-	-	-	-	-	-	
		2SG	-	-	-	-	-	-	
		3SG.INAN	-y	-Ø			-Ø	-?	
	53		2	51					
	100%		2%	96%					
Inanimates	Transitive								
		1SG> INAN	-á ⁴	-y	-á ⁴	-á ⁴	-y	-y	
		2SG> INAN	-u ³	-y	-u ³	-y	-u ³	-u ³	
		3SG> INAN	-y	-Ø	-Ø	-Ø	-Ø	-?	
	122		1	79	37	2	3		
	100%		1%	65%	30%	1.5%	2.5%		

phonological irregularity we observe in Tables 9 and 10 would be weak in the light of other tables such as Table 4 that show that prosodic inflection in Chinantec is rich in irregularities.

4.2.2 Paired verbs and the suffix classes

In this section, I show the distribution of the suffix classes used in the inflection of 175 paired verbs. For this purpose consider the figures in Table 11 (shading highlights where marking preferences lie).

In Table 11, we can see that Class 3 is aligned with animacy. We have seen a similar tendency in animate non-paired verbs in Table 5, but there the distribution was more spread out in the other classes. The same can be said at least for the use of a bare stem, which indisputably identifies the presence of an inanimate argument. For transitive inanimate verbs, the situation is very different from non-paired verbs. With the latter, the distribution of Classes 1 and 3 were even, but with paired verbs, the preference is for Class 1. There are still a sizable number of verbs in Class 2, which would need explaining, because their presence is comparable to the 11 % we find in equivalent non-paired verbs. The figures in Table 11 reveal clear preference for the selection of a particular set of suffixes according to the animacy of the verb. I take such preferences to reveal the application of the morphosyntactic mapping I develop in the next section.

5 The suffix classes and the morphosyntax of agreement

5.1 Revisiting the morphosyntax of animacy agreement

One could say that the major marking tendencies we observe with paired verbs responds to a morphosyntactic mapping from function to form. The mapping could be said to have the form in (24).

(24)	Animate S 1SG → -á ⁴ 2SG → -u ³ 3SG → -y	Inanimate S 3SG → ∅
	A > Animate O 1SG → -á ⁴ 2SG → -u ³ 3SG → -y	A > Inanimate O 1SG → -y 2SG → -y 3SG → ∅

The mapping in (24) includes information about subject (S/A) and information about animacy of S/O. Because of this, it could ideally be divided up further. One possibility is to posit that the person-number paradigm bears a split between the cells involving a subject of 1SG and 2SG, on the one side, and those involving a 3SG, on the other. Let us deal with 3SG first.

On the one hand, an intransitive inanimate verb can only be inflected in 3SG, and in paired verbs this information is encoded by a bare stem by default. On the other hand, transitive verbs can have a subject in all persons, but their object is inanimate. To encode a transitive situation involving a 3SG acting on an inanimate object (>), a bare stem is also used. One possible way to interpret these similarities in form (i.e. use of a bare stem for both situations) is to take them to reflect similarities in function. Accordingly, the use of a bare stem could be taken to signal the inanimate gender of an absolutive argument S/O, as in (25), illustrated by the forms in (26), please note that the inflected stem also carries information about person-number of the subject.

(25) INAN.S/INAN.O → bare stem (-∅)

(26) a. 3SG.INAN ʔí⁴-héʔ⁴
 FUT-get.close.IRR.3[3SG.INAN]
 ‘It’ll get closer.’
 a. 3SG>INAN ʔí⁴-liáʔ⁴
 FUT-push.IRR.3[3SG>INAN]
 ‘He’ll push it.’

Similarly, -y is used to index either an animate S or an animate O, so it could be taken to signal agreement with the animacy of the absolutive argument S/O, regardless of whether the O is 3SG like in (27a) or any other person, like in (27b) or (27c). So the mapping could be the one in (28).

- (27) a. 3SG>AN $\text{ʔi}^4\text{-liá}<\text{y}>\text{ʔ}^2$
 FUT-push.AN.IRR.3<3SG>AN>
 'He'll push him/her.'
- b. 3SG>AN $\text{ʔi}^4\text{-liá}<\text{y}>\text{ʔ}^2$ hniá²
 FUT-push.AN.IRR.3<3SG>AN> 1SG
 'He'll push me.'
- c. 3SG>AN diáʔ⁴ za³ ni³ ka⁴-daʔ¹-hoó¹-y niu³
 PL person FOC PST-3PL-see.AN.CPL.3-3SG>AN 2SG
 'Those people saw you.' (R: 30)

- (28) AN.S/AN.O → -y

The mappings in (25) and (28) are only used in the context of a 3rd person subject, so when they are used, they indirectly convey information of subject as well.

We still need to account for the configurations involving a subject of 1SG and 2SG. In this regard, the majority (65 %) of the transitive inanimate verbs in the pairs belong to Class 1. This means that they select the suffix -y as exponent for both 1SG and 2SG. It is more difficult to make sense of this in a coherent way. Probably the easiest way out is to interpret that this -y has a different function from the -y in (29) while being coincidentally homophonous.¹⁸ In this light, one could propose the mappings in (30).

- (29) AN.S/AN.O → -y₁
- (30) 1SG>INAN/2SG>INAN → -y₂

The situations in (25), (29) and (30) are specific. For all other possibilities, namely when 1SG or 2SG function as S or when they function as A of verbs with an animate object, a default or elsewhere case applies. As both situations are equally encoded by -á⁴ and -u³, respectively, it seems that what is important now is the marking of subject (S/A), as in (31).

- (31) 1SG → -á⁴
 2SG → -u³

¹⁸A second possibility is to think that -y₂ is a leftover of a once productive direct/inverse system where it indicated direct in configurations such as 1>3.INAN or 2>3.INAN. While other Chinantecan languages are analyzed as having a fully-fledged morphological inverse, for example in Sochiapam Chinantec (Foris 2000), there is little or no evidence for an inverse marking in Lealao Chinantec, and the fact that -y₂ is not used for situations such as 1>3.AN or 2>3.AN makes it unlikely that this solution is appropriate. A third possibility would involve regarding the -y₂ in such forms as being the same as -y₁. For such an option to work, one would need to resort to a convoluted solution by positing two affixal slots in the inflected forms for 1SG and 2SG subject in inanimate verbs. The first slot would be used by a bare stem in a mapping similar to (25) indicating inanimate O (which in context is interpreted as of 3SG), the other slot corresponding to -y indicating animate subject (which in context is interpreted as 1SG or 2SG subject). Such structure is represented in the glossing of the form in (i).

- (i) 1SG>INAN $\text{ʔi}^2\text{-lia}<\text{y}>\text{ʔ}^4$
 FUT-push.INAN.IRR.1SG[INAN (3OBJ)]<AN(1SUB)>
 'I'll push it.'

A summary of the proposal is given in (32):

(32)

(a)
S/A
1SG → -á ⁴
2SG → -u ³

(b)
A > Inanimate
1SG → -y ₂
2SG → -y ₂

(c)
Animate S/O
3SG → -y ₁

(d)
Inanimate S/O
3SG → -∅

When the subject is of 1SG and 2SG the situation in (a) serves as a default in the sense of the elsewhere case, while (b) applies to a smaller subclass within (a). This allows for the marking of (a) to also apply to situations that have an inanimate object as indeed happens with the verbs that are characterized as Class 2 in Table 11. Such verbs amount to 37, but 30 of them are verbs of the type in (22)–(23), which do not use prosodic inflection to encode animacy. Another example of this type of verb is given in (33), with an indication of the suffixes the verb requires.

(33) *i⁴t₁⁴a²* (tr, an) ‘receive’ {g, 69}

	PRS	PRF	FUT	SUFF. CL. 2
1SG	t ₁ ⁴² -á ²	ma ³ -t ₁ ³ -á ⁴	?i ⁴ -t ₁ ⁴ -á ⁴	-á ⁴
2SG	t ₁ ⁴ -u ³	ma ³ -t ₁ ³ -u ³	?i ⁴ -t ₁ ⁴ -u ³	-u ³
3SG>AN	t ₁ ⁴ -y	ma ³ -t ₁ ³ -y	?i ⁴ -t ₁ ⁴ -y	-y
3SG>INAN	t ₁ ⁴	ma ³ -t ₁ ³	?i ⁴ -t ₁ ⁴	-∅
1PL	t ₁ ⁴² -a ²	ma ³ -t ₁ ³ -a ²	?i ⁴ -t ₁ ⁴ -a ²	-a ²

Verbs like (33) abide by the default in (a) and are thus not perceived as members of (b). This has two very interesting consequences for the system: (i) while a great number of inanimate transitive verbs are inflected attending to (b), it makes (b) become an inflectional class whose membership needs to be dictated by the lexicon. Even if we defined the exceptions to (b) as being of the type of verbs in (33), there are still 8 verbs of the other patterns in Tables 9 and 10 that would escape the generalization; (ii) because (a) is a default, it strongly suggests that the marking of agreement with subject (S/A) is more important to the system than marking animacy, at least when a 1st and a 2nd person are involved. We can see the impact of the preference of the agreement system based on subject in verbs of the type in (33) in that a form such as ?i⁴t₁⁴a⁴ {1SG.FUT} means both ‘I’ll receive something’ and ‘I’ll receive someone’, or rather, just ‘I receive (an object)’.

Finally, at least for the context of a 3rd person subject, the mapping still works for paired verbs attending to the animacy of S/O, but when we compare this with non-paired verbs, the form-function correspondence in (c) and (d) leaks. All these facts reveal that the morphosyntax of animacy agreement in Lealao Chinantec is less prominent than the morphosyntax of subject agreement. This is something I discuss further in the next section.

Table 12 The encoding 1SG and 2SG in 848 entries

1SG,2SG		S.INAN	S.AN	A>INAN	A>AN	Total	
a.	$-a^4, -u^3$	–	199	176	234	609	72%
b.	$-y, -y$	–	24	186	0	210	25%
c.	$-a^4, -y$	–	0	2	7	9	1%
d.	$-y, -u^3$	–	9	11	0	20	2%
Total		–	232	375	241	848	100%

Table 13 The encoding 3SG in the 1166 entries of the sample

3SG		S.INAN	S.AN	A>INAN	A>AN	Total	
a.	$-y$	67	211	112	234	624	53.6%
b.	$-\emptyset$	212	51	267	9	539	46.2%
c.	$-?$	0	3	0	0	3	0.2%
Total		279	265	379	243	1166	100%

5.2 Revisiting the morphosyntax of subject agreement

In the previous section, we have seen that the morphosyntactic mapping in (32) which accounts for most instances of animacy agreement in paired verbs already implies a situation where 1SG and 2SG are subject. If we study the encoding of these person-number values in the sample as a whole, we find four encoding strategies: (a) using the suffixes $-a^4$ and $-u^3$ for either value; (b) using the suffix $-y$ for both values; or the heteroclitite options in (c) or (d), which result as a combination of (a) and (b). The relevant figures are given in Table 12, based on all 848 entries in the sample that can inflect for a 1st or 2nd person subject.

The figures in Table 12 reveal that strategy (a) serves as a default, both in its sense of the elsewhere case and the most frequent option, suggesting that the suffixes $-a^4$ and $-u^3$ function as default indexes of subject of 1SG and 2SG without any significant correlation with animacy. Strategy (b), is the second most favored option, it constitutes a lexically defined class of verbs with a great majority of inanimate transitives.

All this points to the existence of an agreement system in Lealao Chinantec that is hybrid between encoding information about subject agreement and about animacy agreement. Information about animacy is scattered. In the mapping in (32) it was mainly at work in the configurations in (b) and (c). But outside paired verbs, the association of $-y$ is no longer restricted to animacy. To show this, consider Table 13, where I give the distribution for the three coding strategies of a 3SG subject in the whole sample: (a) is the suffix $-y$; (b) is by using a stem with no exponent; and (c) is by means of the suffix $-?$.

While the workload of strategy (c) is utterly insignificant for the syntax, strategies (a) and (b) have comparable frequencies. In other words, they instantiate two inflectional classes which have an equal amount of labor regarding the encoding of 3rd person subject. We have seen in the mapping in (32) that at least for paired verbs, the reference to a 3rd person subject is indirect, the emphasis lying on an animacy distinction of the S/O. However, when the whole system is taken into account, the function of $-y$ or $-\emptyset$ can no longer be said to be linked to animacy. This is particularly

Table 14 The classes and animacy configurations

		S.AN	O.INAN	O.AN	Total	S.AN	O.INAN	O.AN	Total	S.AN	O.INAN	O.AN
Class 1	-y, -y, -Ø	24	186	0	210	11.5%	88.5%	–	100%	10.5%	49.6%	–
Class 2	-á ⁴ , -u ³ , -Ø	10	65	2	77	13%	84.5%	2.5%	100%	4.3%	17.4%	0.8%
Class 3	-á ⁴ , -u ³ , -y	189	111	232	532	35.5%	21%	43.5%	100%	81.5%	29.6%	96.2%
Class 4	-á ⁴ , -y, -Ø	0	2	7	9	–	22%	78%	100%	–	5.2%	3%
Class 5	-y, -u ³ , -Ø	6	11	0	17	35%	65%	–	100%	2.5%	2.8%	–
Class 6	-y, -u ³ , -ʔ	3	0	0	3	100%	–	–	100%	1.2%	–	–
Total		232	375	241	848					100%	100%	100%

relevant for -y because the figures in Table 13 indicate that there is some expectation for a bare stem NOT to be used to realize 3SG subject with an animate O. As no similar expectation can be said of -y, it strongly suggests that -y is a default exponent of a 3SG subject, regardless of animacy.

The favored strategy for the encoding of 1SG and 2SG together with the default strategy for the encoding of 3SG shows that Class 3 serves as a default mapping of subject, given in (34), while still being strongly associated to verbs with an animate S/O argument.

(34)

S/A	
1SG	→ -á ⁴
2SG	→ -u ³
3SG	→ -y

Class 3 is also the largest class; the remainder of the verbal lexicon gathers in the rest of the classes in an uneven way. The membership to the suffix classes is still largely a lexical matter, and hence a treatment in inflectional classes is justified, but there are tendencies in such membership involving the animacy features of the S/O argument of the verb. The general distribution of the classes attending to animacy value is given in Table 14 for the 848 verbs in the samples which can inflect in all persons. Notice that Table 14 cannot give information about {S.INAN}. For the distribution of this value the classes taken as whole paradigms are not informative, and only Table 13 should be informative. The zeros of 3SG could pertain to at least four classes, but as pointed out above, the 67 verbs that manifest -y for {S.INAN} can only be understood as pertaining to Class 3.

Being a default, Class 3 is found for all animacy values, but shows preferences for the animate value as most verbs with an {S.AN} argument and practically all with {O.AN} are members of this class. Most of the verbs of Class 1 have {O.INAN}, but only half of the verbs with {O.INAN} are found in Class 1, the rest are distributed in other four classes.

The system we observe in Lealao Chinantec is remarkably similar in typological terms to the inflectional class systems found in better-known Indo-European languages where gender plays a major part in class membership, as discussed for example in Wurzel (1984), Corbett and Fraser (1993), Fraser and Corbett (1994), Aronoff (1994), among others. For example, Russian nouns are said in the literature to fall into four inflectional classes (see Corbett and Fraser 1993; Müller 2004, etc.), and the membership of such classes show strong correlations with

gender.¹⁹ One class contains all masculine nouns, while a second one contains all neuter nouns. Masculine nouns are also found in a third class that instead is mostly populated by feminine nouns, which in turn are virtually the only type of nouns found in the fourth class. The case of Lealao Chinantec is very similar to this situation if one sees animacy as a type of gender. But perhaps, because verbs in Chinantec are often still talked about as being targets of animacy agreement rather than controllers, the typological simile would perhaps have to be made with agreeing adjectives organized in inflectional classes.

6 Concluding remarks

In this paper, I have shown that verbs in Lealao Chinantec receive different sets of suffixes for the person-number values 1SG, 2SG, and 3SG which have a puzzling distribution because whereas they are used to index agreement of subject, just like other suffixes for 1PL and 2PL which have no allomorphy, they also convey information about the animacy of the S/O argument of the verb. I have identified six inflectional classes involving the suffixes—four in Rupp and Rupp's (1996) account—although three of them comprise the greater bulk of verbs in a sample of over 1000 verbal entries.

Studying the distribution of these suffixes, I have concluded that they manifest a mixed system of agreement, combining a more canonical system based on subject agreement (S/A) which is default and another based on the animacy features of the S/O. The morphology involved in the main three classes responds to a morphosyntactic mapping that is operational in verbs with two animacy paradigms. But as there are important mismatches in the meaning-to-form correspondences in the inflection of verbs with only one animacy value, the system ceases to be functional and a treatment of the suffixal allomorphy in inflectional class terms is justified.

As it is widely accepted nowadays in grammaticalization theory that transparent morphosyntactic systems often develop into more morphologically opaque ones (Givón's (1971) aphorism that today's morphology is yesterday's syntax), it seems reasonable to think that the functional morphosyntactic mapping we observe at work in paired verbs was perhaps prior to the system, but evolved into another system where the link between function and form has been broken for a substantial set of the lexicon. In lack of diachronic evidence to support it, this trend of thought should remain at a speculation level. However, the mixed morphological system we observe synchronically can shed new lights on a typology of agreement, as developed in Corbett (2006), Evans (2003), Polinsky (2003), etc.

Lealao Chinantec is one of the many endangered indigenous languages of Mexico, and we can only hope that it survives and keeps evolving through time. It belongs to the Chinantecan branch of Oto-Manguean which comprises languages that display some of the most complex morphological systems known to us. The fact that Chinantecan languages are gaining attention of morphologists comes as good news for the

¹⁹These classes receive different number labels according to analysis and tradition, so I will only refer to them generically here.

discipline, and probably in just the right time, as it is important for linguists to join efforts to improve a theory of morphology by expanding it beyond the typological limits of the inflection of better-known European languages.

Acknowledgements This paper was written under the auspices of the ESRC/AHRC project RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguean languages”. I want to thank Matthew Baerman and Greville Corbett for their invaluable comments on an earlier draft of this paper. I owe my gratitude to the two anonymous reviewers and the main editor of *Morphology* for all their wise comments and suggestions which helped me improve the article considerably. I presented a version of this paper at a talk at the 2013 Meeting of the SSILA at Boston. Another version was presented at the Seminar TypUlm at the École Normale Supérieure of Paris in 2014. I am grateful for all the feedback received. I am also thankful to Penelope Everson for proofreading my English. All errors and deficiencies remain my own responsibility.

Appendix: The prosodic classes of Lealao Chinantec (´ = ballistic stress)

Cl.	3		
	INCPL	CPL	IRR
a	'1	'1	42
b	3	4	'2
c	'3	'4	'3
d	'3	'3	42
e	4	'1	'1
f	4	3	3
g	4	'3	'4
h	4	32	4
i	4	'42	'42
j	32	32	1
k	32	4	'2
l	32	'2	'2
m	32	'4	'4
n	'3	'3	'4
o	3	'3	3
p	4	'4	'4
q	4	'4	4
r	32	32	'3
s	32	2	2
t	1	1	1
/1/	'1	'1	'1
/2/	2	2	2
/2'/	'2	'2	'2
/3/	3	3	3
/3'/	'3	'3	'3
/32/	32	32	32
/4/	4	4	4
/4'/	'4	'4	'4
/42/	42	42	42

Cl.	1sg			2			1pl		
	INCPL	CPL	IRR	INCPL	CPL	IRR	INCPL	CPL	IRR
1	'1	'1	'1	'1	'1	'1	'1	'1	'1
2	1	1	1	1	1	1	1	1	1
3	42	42	42	42	42	42	42	42	42
4	32	'2	1	32	'2	1	32	'2	1
5	42	42	42	42	42	42	42	42	42
6	42	42	42	4	4	4	42	42	42
7	4	'4	42	3	'1	3	42	2	2
8	42	42	42	'1	'1	'1	2	2	2
9	32	'2	1	'3	42	'3	32	42	42
10	'4	'4	'4	32	'2	1	32	1	1
11	'3	'4	'3	'3	'4	'3	'3	'4	'3
12	42	42	42	3	3	3	32	32	32
13	'4	'4	'4	'4	'4	'4	42	42	42
14	'4	'4	'4	'4	42	'4	42	42	42
15	'4	'4	'4	'4	'4	'4	32	42	42
16	4	'4	42	3	3	3	32	2	2
17	32	'2	1	'3	'2	'3	32	42	42
18	42	42	42	4	4	4	32	32	32
19	'4	'4	'4	32	'1	1	32	2	2
20	4	'4	42	3	'3	3	32	2	2
21	'4	'4	'4	32	'1	1	32	2	2
22	4	4	4	3	3	3	32	32	32
23	'4	'4	'4	32	'1	1	4	2	2
24	32	'2	1	'4	'4	'4	32	42	42
25	32	'2	1	4	'4	'4	42	42	42
26	32	32	1	32	32	1	32	32	1
27	4	4	4	'1	'1	'1	2	2	2
28	42	42	42	4	'4	'4	42	'4	'4
29	'4	'4	'4	32	'2	1	32	42	42
30	42	42	1	42	42	42	42	42	42
31	32	'2	1	3	3	3	32	2	2
32	32	'2	1	'3	'2	'3	32	1	1
33	4	4	4	3	3	3	42	42	42
34	4	'4	4	4	'4	4	'4	'4	42
35	42	'4	42	4	'4	4	'4	'4	42
36	4	'4	42	'3	'3	'3	42	42	42
37	4	'4	42	'1	'1	'1	42	2	2
38	'4	'4	'4	'2	'2	'2	'4	'4	'4
39	'4	'4	'4	32	'2	'2	32	32	32
40	'4	'4	'4	32	'2	1	42	1	1
41	42	42	4	42	42	42	42	42	42
42	42	'2	1	'1	'3	'3	2	2	2
43	32	32	42	4	32	4	42	32	42
44	4	'4	42	'3	'3	'3	2	2	2
45	4	'4	42	'3	'3	'3	32	2	2
46	4	'2	1	'3	'1	'3	'3	2	2
47	4	'4	42	'1	'1	'1	42	2	2
48	4	'4	42	'1	'1	'1	2	2	2
49	32	'2	1	'3	'1	'3	'3	2	2
50	32	'2	1	32	'1	1	'4	2	2
51	42	'2	1	42	'2	1	42	42	42
52	42	'2	1	42	4	4	42	42	42
53	4	4	4	'1	'1	'1	'1	'1	'1
54	'4	'4	'4	32	'1	1	32	2	2
55	'4	'4	'4	32	'2	1	32	2	2
56	'4	'4	'4	'1	'1	'1	2	2	2
57	'4	'4	'4	2	2	2	32	32	32
58	42	'4	42	'1	3	3	2	2	2
59	42	42	42	3	3	3	3	3	3
60	42	42	42	'1	'1	'1	2	2	2
61	4	'4	42	3	'2	3	2	2	2
62	32	'2	1	4	4	4	32	42	42
63	42	42	42	32	'1	1	32	42	42
64	32	'2	1	'3	'1	'3	32	1	1
65	'4	'4	'4	3	3	3	32	32	32
66	42	4	4	4	4	4	32	4	4
67	4	4	4	1	1	1	1	1	1
68	'4	'2	1	32	4	4	32	42	42
69	42	'3	4	4	'3	'4	42	'3	'4
l	32	'2	2	32	'2	'2	32	'2	'2
q	4	'4	4	4	'4	'4	4	'4	'4
/1/	1	1	1	1	1	1	1	1	1
/1'/	'1	'1	'1	'1	'1	'1	'1	'1	'1
/2/	2	2	2	2	2	2	2	2	2
/2'/	'2	'2	'2	'2	'2	'2	'2	'2	'2
/3/	3	3	3	3	3	3	3	3	3
/3'/	'3	'3	'3	'3	'3	'3	'3	'3	'3
/32/	32	32	32	32	32	32	32	32	32
/4/	4	4	4	4	4	4	4	4	4
/4'/	'4	'4	'4	'4	'4	'4	'4	'4	'4
/42/	42	42	42	42	42	42	42	42	42

References

- Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: implicative patterns in inflectional paradigms. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: form and acquisition* (pp. 54–82). Oxford: Oxford University Press.
- Anderson, J. L. (1989). *Studies in Chinantec languages: Vol. 3. Comaltepec Chinantec syntax*. Arlington: The Summer Institute of Linguistics and the University of Texas.
- Aronoff, M. (1994). *Linguistic inquiry monograph: Vol. 22. Morphology by itself: stems and inflectional classes*. Cambridge: MIT Press.
- Baerman, M. (2012). Paradigmatic chaos in Nuer. *Language*, 88(3), 467–494.
- Baerman, M. (2013a). Inflection class interactions. In N. Hathout, F. Montermini, & J. Tseng (Eds.), *LINCOM studies in theoretical linguistics: Vol. 51. Morphology in Toulouse. Selected proceedings of Décembrettes 7*. Munich: Lincom.
- Baerman, M. (2013b). *Inflectional classes in the languages of the Americas: an overview*. Paper read at the 87th annual meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA), Boston, 3–6 January 2013.
- Baerman, M. (2014a). Inflectional interactions in Oto-Manguean. In J.-L. Léonard & A. Kihm (Eds.), *Issues in meso-American morphology* (pp. 15–25). Paris: Michel Houdiard.
- Baerman, M. (2014b). Covert systematicity in a distributionally complex system. *Journal of Linguistics*, 50(1), 1–47.
- Baerman, M., & Palancar, E. L. (forthcoming). The organization of Chinantec tone paradigms. *Cahiers de Grammaire*.
- Baerman, M., Brown, D., & Corbett, G. G. (2010). *Morphological complexity: a typological perspective*. Position paper, MS.
- Bonami, O. (Ed.) (2012). Stems. *Word Structure*. Special issue
- Brown, D., & Hippisley, A. (2012). *Network morphology: a defaults-based theory of word structure*. Cambridge: Cambridge University Press.
- Brown, D., & Evans, R. (2013). *Discovering classes in Tlaltepec Chinantecan*. Paper presented at workshop ‘Theoretical and Computational Morphology (TACMO): new trends and synergies’, 19th international congress of linguists, Geneva, July 21.
- Corbett, G. G. (2006). *Agreement*. Cambridge: Cambridge University Press.
- Corbett, G. G. (2007). Canonical typology, suppletion and possible words. *Language*, 83, 8–42.
- Corbett, G. G. (2009). Canonical inflectional classes. In F. Montermini, G. Boyé, & J. Tseng (Eds.), *Selected proceedings of the 6th Décembrettes: morphology in Bordeaux* (pp. 1–11). Somerville: Cascadilla Proceedings Project.
- Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Corbett, G., & Baerman, M. (2006). Prolegomena to a typology of morphological features. *Morphology*, 16(2), 231–246.
- Corbett, G., & Fraser, N. (1993). Network morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics*, 29, 113–142.
- Evans, N. (2003). Typologies of agreement: some problems from Kayardild. *Transactions of the Philological Society*, 101(2), 203–234.
- Finkel, R., & Stump, G. T. (2007). Principal parts and morphological typology. *Morphology*, 17(1), 39–75.
- Finkel, R., & Stump, G. T. (2009). Principal parts and degrees of paradigmatic transparency. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: form and acquisition* (pp. 13–53). Oxford: Oxford University Press.
- Foris, D. P. (2000). *Studies in Chinantec languages: Vol. 6. A grammar of Sochiapam Chinantec*. Dallas: SIL International and the University of Texas at Arlington.
- Fraser, N., & Corbett, G. (1994). Gender, animacy, and declensional class assignment: a unified account for Russian. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1994* (pp. 123–150). Dordrecht: Kluwer Academic.
- Givón, T. (1971). Historical syntax and synchronic morphology: an archeologist’s fieldtrip. *Proceedings of the Chicago Linguistic Society*, 7, 394–415.
- Merrifield, W. R. (1968). *Serie científica: Vol. 9. Papeles de la Chinantla: Vol. 5. Palantla Chinantec grammar*. Mexico City: Museo Nacional de Antropología.
- Mugele, R. L. (1982). *Tone and ballistic syllables in Lalana Chinantec*. Unpublished Ph.D. dissertation at the University of Texas at Austin.
- Müller, G. (2007). Notes on paradigm economy. *Morphology*, 17, 1–38.

- Müller, G. (2004). On decomposing inflection class features: syncretism in Russian noun inflection. In G. Müller, L. Gunkel, & G. Zifonun (Eds.), *Explorations in nominal inflection* (pp. 189–227). Berlin: de Gruyter.
- Palancar, E. L. (2014). Revisiting the complexity of the Chinantecan verb conjugation classes. In J.-L. Léonard & A. Kihm (Eds.), *Patterns in meso-American morphology* (pp. 77–102). Paris: Michel Houdiard.
- Polinsky, M. (2003). Non-canonical agreement is canonical. *Transactions of the Philological Society*, 101(2), 279–312.
- Rensch, C. (1978). Ballistic and controlled syllables in Otomanguean languages. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 85–92). Amsterdam: North Holland.
- Rupp, J., & Rupp, N. (1996). *Vocabularios y diccionarios indígenas “Mariano Silva y Aceves”: Vol. 35. Diccionario chinanteco de San Juan Lealao, Oaxaca*. Mexico: Instituto Lingüístico de Verano.
- Rupp, J. E. (1989). *Studies in Chinantec languages: Vol. 2. Lealao Chinantec syntax*. Dallas: SIL International and the University of Texas at Arlington.
- Rupp, J. E. (1990). The Lealao Chinantec syllable. In W. R. Merrifield & C. R. Rensch (Eds.), *Studies in Chinantec languages: Vol. 4. Syllable, tones and verb paradigms* (pp. 63–74). Dallas: SIL International and the University of Texas at Arlington..
- Silverman, D. (1994). A case study in acoustic transparency: [spread glottis] and tone in Chinantec. In *Proceedings of NELS 24*. Amherst: University of Massachusetts.
- Stump, G. (2006). Heteroclisys and paradigm linkage. *Language*, 82, 279–322.
- Stump, G., & Finkel, R. A. (2013). *Morphological typology: from word to paradigm*. Cambridge: Cambridge University Press.
- Wurzel, W. U. (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie Verlag.

Enrique L. Palancar* and Timothy Feist

Agreeing with subjects in number: The rare split of Amuzgo verbal inflection

DOI 10.1515/lingty-2015-0011

Received July 4, 2015; revised October 18, 2015

Abstract: Verbs in San Pedro Amuzgo, an Oto-Manguean language of Mexico, often have two different stems in the paradigm, one used with singular subjects and the other with plural subjects. This split motivated by number is typologically interesting due to its rarity, since number splits are commonly along the S/O vs. A distinction, not the S/A vs. O distinction. Apart from at stem level, the split is also manifested in the incomplete of an inflectional class of verbs. At stem level the plural stem is derived in a variety of unproductive ways, making the relation between singular and plural stems, synchronically, one of suppletion. In this article, we study the distribution and the morphological properties of this split in depth, using a sample of almost 600 fully inflected verbs from a large database compiled by native linguist Fermín Tapia and now publicly accessible on the Surrey Morphology Group's website. We also place it in a typological context, relating it to other systems we have observed.

Keywords: agreement, Amuzgo, inflection, morpheme, number, Oto-Manguean, relational alignment, suppletion, syntax, tone

[T]he formal expression of these categories involves a great deal of suppletion and morphologically conditioned internal modification and fusion, resulting in an extreme degree of irregularity. Indeed, Amuzgo morphology is so irregular that we have been tempted to call it a lexical language; that is, a language where the ideal seems to be for each form to have an idiosyncratic individuality rather than for it to be productively generatable [sic].

(Smith-Stark & Tapia 1986)

1 Introduction

In this article we introduce a typologically rare case of number agreement that is found in the verbal inflection of San Pedro Amuzgo (henceforth SP Amuzgo), an

*Corresponding author: Enrique L. Palancar, Structure et Dynamique des Langues (SeDyL – UMR8202), CNRS, 7 rue Guy Môquet, 94800 Villejuif, France, E-mail: palancar@vjf.cnrs.fr
Timothy Feist, Surrey Morphology Group, University of Surrey, Guildford, Surrey, GU2 7XH, UK, E-mail: t.feist@surrey.ac.uk

Oto-Manguéan language of Mexico. The paradigm of many verbs in this language displays a split, in other words, a division into two parts. The split is manifested in such a way that one of the portions of the paradigm is used when the verb agrees with a singular subject and the other when it agrees with a plural subject. The phenomenon is illustrated in Table 1 with the past progressive subparadigm of the two verbs *meʔ^l* ‘lie down’ and *tsaʔ^l* ‘hobble’,¹ where we use shading to distinguish the plural section of the subparadigm from the singular one. (In the segmentation, the plus sign represents a stem formative: in this case the formative *ndi⁵*+, found in the inflection of many verbs.)

Table 1: Two verbs with a subject number split in SP Amuzgo.

PST.PRG	‘lie down’	‘hobble’
1SG	<i>to³-meʔ^l=ha³</i>	<i>to³-ndi⁵+tsaʔ^l=ha³</i>
2SG	<i>to³-meʔ^l=ʔu³</i>	<i>to³-ndi⁵+tsaʔ^l=ʔu³</i>
3SG	<i>to³-meʔ^l=e⁵</i>	<i>to³-ndi⁵+tsaʔ^l=e⁵</i>
1PL.INCL	<i>to³-toēʔ³=ha⁵a⁵</i>	<i>to³-ndi⁵+ndaʔ^l=ha⁵a⁵</i>
1PL.EXCL	<i>to³-toēʔ³=ha⁵l</i>	<i>to³-ndi⁵+ndaʔ^l=ha⁵l</i>
2PL	<i>to³-toēʔ³=ʔo³</i>	<i>to³-ndi⁵+ndaʔ^l=ʔo³</i>
3PL	<i>to³-toēʔ³=e³</i>	<i>to³-ndi⁵+ndaʔ^l=e³</i>

In Table 1, the split is manifested by the occurrence of two different stems for each verb: *meʔ^l* vs. *toēʔ³* for ‘lie down’ and *tsaʔ^l* vs. *ndaʔ^l* for ‘hobble’. The selection of the stems responds to the presence of a singular or a plural subject, respectively. How shall we interpret this? One way to start is by taking a look at basic split types.

In a recent article, Corbett (2015) proposes that splits in inflectional paradigms can be of two main types, labeled “motivated” and “morphomic”. The split in Amuzgo is an instance of a motivated one. Motivated splits respond to the existence of a natural class, commonly in the form of a coherent bundle of feature values from the syntax or the semantics, while morphomic ones do not. A typical example of a morphomic split is the suppletive pattern found in the present indicative of the French verb *aller* ‘go’ in Table 2.² The split separates the portion of the paradigm for 1st and 2nd person plural from the rest, but these person and number subsets do not form natural classes definable through single features (hence the concept of a “morphome”).

1 SP Amuzgo is a tone language. There are eight contrastive tones, which are represented here with numbers, 1 being low. For more details see Section 2.3.

2 A split of this kind is pervasive in Romance languages; for more details see Maiden (2005).

Table 2: The morphomic split of *aller* ‘go’ in French.

PRS.IND	‘go’
1SG	<i>vais</i>
2SG	<i>vas</i>
3SG	<i>va</i>
1PL	<i>allons</i>
2PL	<i>allez</i>
3PL	<i>vont</i>

In contrast, motivated splits respond to natural classes, for example an aspect distinction as shown in Table 3 by the verb $ba^2\check{s}t\check{\alpha}^2$ ‘wrap’ in Chiquihuitlán Mazatec, another Oto-Manguean language, adapted from Jamieson (1982). The table shows the forms for the so-called neutral and the incomplete aspect.

Table 3: The motivated TAM split of $ba^2\check{s}t\check{\alpha}^2$ ‘wrap’ in Chiquihuitlán Mazatec.

	‘wrap’	
	NEUTRAL	INCL
1SG	$ba^1\check{s}t-\check{\alpha}^1$	$ku\check{\alpha}^1\check{s}t-\check{\alpha}^1$
2SG	$ba^2\check{s}t-e^2$	$ku\check{\alpha}^2\check{s}t-e^2$
3	$ba^2\check{s}t-\check{\alpha}^2$	$ku\check{\alpha}^4\check{s}t-\check{\alpha}^2$
1PL.INCL	$ba^2\check{s}t-e^2$	$ku\check{\alpha}^2\check{s}t-e^2$
1PL.EXCL	$ba^2\check{s}t-j^{24}$	$ku\check{\alpha}^4\check{s}t-j^{24}$
2PL	$ba^2\check{s}t-y^2$	$ku\check{\alpha}^2\check{s}t-y^2$

The split in SP Amuzgo in Table 1 treats the encoding of 1PL.INCL, 1PL.EXCL, 2PL, and 3PL as a unified class by virtue of the fact they all share the value of plurality. This separation leaves the treatment of 1SG, 2SG, and 3SG as a natural class too, sharing the value of singular. Because this separation is based on a number distinction grouping singular and plural forms as natural classes, this split is also an instance of a motivated split. This has consequences for the way in which we might interpret the syntax–morphology interface involving the number of the subject in Amuzgo morphosyntax, because motivated splits are not like morphomic ones which only involve morphological patterns. As a consequence of the split, in order to inflect a verb in SP Amuzgo, it is necessary to first pay heed to the number of the subject in order to select the right stem {SG vs. PL} and once this is done, attend to person-number values (and clusivity) as classes {1SG} vs. {1PL.INCL}, etc.

The split shown in Table 1 is not something that occurs with only a couple of irregular verbs in SP Amuzgo, as is the case with the French verb *aller*, but rather it affects hundreds of verbs in the lexicon, so it is a defining characteristic of Amuzgo inflection and thus of its morphosyntax. On the formal level, the split has its intricacies because, although there are instances where the contrast between the two stems is achieved by a transparent morphological operation (albeit not a productive one), in most cases the relation between them is one of suppletion, at least synchronically. In light of this, the purpose of the article is to present a detailed description of the factors that contribute to the singular–plural subject split in SP Amuzgo and to place it in a typological perspective in order to reveal what it can tell us about agreement patterns. To do this, a word about the language and the basics of its verbal inflection is in order.

2 Amuzgo and its verbal inflection

2.1 The Amuzgo family

Amuzgo is a small language family from Mexico that belongs to the Oto-Manguean phylum. The family is thought to be composed of three languages (although a rigorous dialectal survey is still required to corroborate this): the language spoken in the state of Guerrero with its epicentre in the town of Xochistlahuaca (30,000 speakers);³ and two other languages in the state of Oaxaca: the language spoken in and around Santa María Ipalapa (900) and the language spoken in and around the village of San Pedro Amuzgos (6,500). We base our analysis on the San Pedro Amuzgos variety, located in Map 1, but the existing materials on Xochistlahuaca Amuzgo in Apóstol (2014) and Buck (2012) confirm that verbal inflection is very similar in structural terms, indicating that the phenomenon we study here should be taken to be a characteristic trait of Amuzgo as a family.

2.2 Source of the data

For our analysis, we use a sample of 579 verbs from the Amuzgo language of San Pedro Amuzgos, which we will refer to as SP Amuzgo. We sometimes refer to it just as “Amuzgo” when the phenomenon under discussion is a trait of Amuzgo

³ Census by the Instituto Nacional de Lenguas Indígenas (INALI) in 2000.



Map 1: San Pedro Amuzgos, Oaxaca, Mexico.

as a system and can be found in the other Amuzgo languages independently of differences in form.

The most important source of our data is the inflectional database compiled by the native linguist Fermín Tapia, which is publicly accessible online via the Surrey Morphology Group site at <http://www.oto-manguean.surrey.ac.uk>. Fermín Tapia's complete database consists of about 500 inflected verbs. However, for present purposes, we have only used 311 of them. The ones we include have been revised by Fermín Tapia together with linguist Yuni Kim from the University of Manchester.⁴ We have also left out a few compound verbs that show inflectional behavior identical to those of the simple verbs on which they are based.

Fermín Tapia's sample (henceforth FT) has an overwhelming number of intransitive verbs: only 35 out of the 311 verbs in our sample are transitive. To compensate for the lack of transitive verbs, we have also included in our sample another 268 transitive verbs from Stewart & Stewart's (2000) dictionary (henceforth S&S), which are also available in the online database. We have excluded all transitive verbs in S&S that were already present in FT's database, together with a couple of compounds which were inflectionally uninformative. It should be noted that the data in the two samples differ as to the type and scope of the inflectional information included, thus allowing us to use the two sources in different ways. The focus of the analysis is on FT's data because it is more complete: it gives all inflectional forms for every verbal entry. Every verb in SP Amuzgo has a paradigm consisting of 42 cells. This means that our analysis has

⁴ In the online database, unrevised verbal entries are indicated by two asterisks.

been based primarily on the observation of 13,062 inflected forms. To complete the picture of Amuzgo verbal inflection we have also used data from Buck (2000), which served as the basis for the grammatical categorization of lexemes in the entries in S&S. The figures in Table 4 provide a summary of how the verbs are distributed in our sample.

Table 4: Verbs in the sample.

	Intransitive	Transitive	Total
Fermín Tapia's sample	276	35	311
Stewart & Stewart (2000)	0	268	268
Total	276	303	579

With the addition of S&S's data, the number of intransitive and transitive verbs in the sample is practically even. However, Amuzgo is an intransitive-oriented language; this means that the majority of basic verbs in the lexicon are intransitive, while the vast majority of the transitive verbs are causative verbs (60% of the verbs in our sample), although the causative morphology they display is unproductive synchronically (see Section 5.6 for more details).

2.3 Basics of the verbal inflection of Amuzgo

In this section we introduce a few relevant characteristics of the verbal inflection of SP Amuzgo. This is done with the intention of creating a background for understanding what will follow. Verbs in Amuzgo inflect for TAM and for the person of the subject. As for TAM, verbs inflect for six values: incomplete, complete, past progressive, subjunctive, irrealis, and future progressive. The TAM paradigm of the verb $ndi^5ts\tilde{a}\tilde{\rho}^3$ 'go crazy' is given in Table 5. As a general

Table 5: The TAM of the verb $ndi^5ts\tilde{a}\tilde{\rho}^3$ 'go crazy' in three persons.

'go crazy'			
	1SG	3SG	3PL
INCPL	$ba^5\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$ba^5\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$ba^5\cdot ndi^5+na\tilde{\rho}^3=\tilde{e}^3$
CPL	$t\cdot(nd)i^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$t\cdot(nd)i^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$t\cdot(nd)i^5+na\tilde{\rho}^3=\tilde{e}^3$
PST.PRG	$to^3\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$to^3\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$to^3\cdot ndi^5+na\tilde{\rho}^3=\tilde{e}^3$
SUBJV	$kw\cdot(nd)i^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$kw\cdot(nd)i^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$kw\cdot(nd)i^5+na\tilde{\rho}^3=\tilde{e}^3$
IRR	$n^5\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$n^5\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$n^5\cdot ndi^5+na\tilde{\rho}^3=\tilde{e}^3$
FUT.PRG	$n^5ngo^3\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=ha^3$	$n^5ngo^3\cdot ndi^5+ts\tilde{a}\tilde{\rho}^3=\tilde{e}^5$	$n^5ngo^3\cdot ndi^5+na\tilde{\rho}^3=\tilde{e}^3$

rule, the marking of TAM is independent from that of person and number. To illustrate this, in Table 5 we give the forms for the 1st person singular alongside those of the 3rd person singular and plural, which shows that there is no change affecting prefixal marking. We have set the prefixes marking TAM values in boldface.

For the marking of the person of the subject, verbs fall into two classes that are called “inactive” and “active” in Smith-Stark & Tapia (1986, 2002). Members of the inactive class have a full set of segmental enclitics to mark person-number. All verbs given in the examples so far are inactive verbs. In contrast, for verbs belonging to the active class, the person of the subject is realized by a combination of tone patterns, glottalization patterns, and person suffixes.⁵ The two classes are illustrated in Table 6 with the incompletive forms of the verbs ‘get upset’ and ‘get sick’. Notice that inactive verbs do not show variation in stem tone, while active verbs may and often do.⁶ In Table 6, we have also included the pronominal paradigm used for each class, following Smith-Stark & Tapia (2002). Shading is used for the values where the same form is used in both classes.

Table 6: Inactive/active intransitive verbs.

INCL		Inactive 'get upset'	Active 'get sick'	Inactive	Active
1SG		\emptyset -xɾē ⁵ =ha ³	\emptyset -bɾi ¹²	=ha ³	–
2SG		\emptyset -xɾē ⁵ =ɾu ³	\emptyset -bi(ɾ) ⁵¹ -ɾ	=ɾu ³	-ɾ
3SG	HUM	\emptyset -xɾē ⁵ =ē ⁵	\emptyset -biɾ ¹² =hu ⁵	=ē ⁵	=hu ⁵
	HON	\emptyset -xɾē ⁵ =ē ³	\emptyset -biɾ ¹² =hō ⁵³	=ē ³	=hō ⁵³
	ANIMAL	\emptyset -xɾē ⁵ =oɾ ⁵ /āɾ ⁵	\emptyset -biɾ ¹² =oɾ ⁵	=oɾ ⁵ /=āɾ ⁵	=oɾ ⁵
	INANIMATE	\emptyset -xɾē ⁵ =āɾ ¹	\emptyset -biɾ ¹² =āɾ ¹	=āɾ ¹	=āɾ ¹
1PL.INCL		\emptyset -nkɾē ⁵ =ha ⁵ ⟨a⟩ ⁵	\emptyset -bɾi ¹² ⟨i⟩ ⁵	=ha ⁵ ⟨a⟩ ⁵	–
1PL.EXCL		\emptyset -nkɾē ⁵ =ha ⁵¹	\emptyset -bɾi ⁵¹	=ha ⁵¹	–
2PL		\emptyset -nkɾē ⁵ =oɾ ³	\emptyset -bɾi ¹² =oɾ ³	=oɾ ³	=oɾ ³
3PL	HUM	\emptyset -nkɾē ⁵ =ē ³	\emptyset -bɾi ¹² =ho ⁵	=ē ³	=ho ⁵
	HON	\emptyset -nkɾē ⁵ =ā ³⁴	\emptyset -biɾ ¹² =ā ³⁴	=ā ³⁴	=ā ³⁴

The inactive class consists almost entirely of intransitive verbs, while the active class contains both intransitive and transitive verbs. Taking into account intransitive verbs only, the two classes are distributed in our sample as

⁵ A vowel marked as ⟨v⟩ represents either an epenthetic echo vowel or a powered vowel.

⁶ The inactive verb ‘get upset’ also shows the number split we have presented in Table 1, while the verb ‘get sick’ is an example of a verb with glottalization (see further below).

Table 7: Distribution of inactive/active intransitive verbs.

	Inactive	Active	Total
Fermín Tapia's sample	119	157	276
Stewart & Stewart (2000)	260	178	438

indicated in Table 7. For convenience, we also include information about the intransitive verbs present in S&S's dictionary but which are not considered in our sample.

Both classes have a semantic core. In other words, most of their members share a bundle of semantic properties typical of systems with semantic alignment (Mithun 1991; Donohue & Wichmann (eds.) 2008, etc.). Within the active class (disregarding transitive verbs), one finds verbs referring to activities (*kwaʔ³* 'eat', etc.), communication (*ʔma³* 'talk', etc.), manner of speaking (*ta³* 'sing', *tiu¹* 'whistle', etc.), motion verbs (*wa⁵ka⁵* 'walk', etc.), and posture and grooming verbs (*wa³⁵* 'lie down', *n⁵de³⁵ʔ* 'bathe', etc.). Within the inactive class we find verbs for natural phenomena (*waʔ¹* 'rain', etc.), aspectual predicates (*nde¹* 'for something to finish', etc.), patient-oriented verbs (*tyheʔ¹* 'boil', *ʔō⁵* 'sprout', *toʔ⁵* 'get filled up', etc.), inchoative verbs derived from adjectives (*ndi⁵+si¹* 'get wrinkled', *ndi⁵+hā⁵* 'get dark', etc.), and so on. However, Smith-Stark & Tapia (1986, 2002) point out that despite such semantic alignment tendencies, membership of the classes is not predictable from the semantics because there are also verbs which are in one class but would in principle be expected to belong to the other, i.e., the verbs *biʔ¹²* 'get sick' in Table 6, *hmeʔ³⁴* 'be hot', *beʔ³* 'die', and *be³+hndyaʔ⁵* 'get tired' are found in the active class, just as *huʔ¹* 'jump' and other causative based verbs such as *tsi⁵+nthaʔ⁵³* 'quarrel' and *tsi⁵+tiu¹* 'call somebody by whistling' (from the active verb *tiu¹* 'whistle') are members of the inactive class. This lack of predictability suggests that the most suitable analysis of the classes is in terms of inflectional classes rather than just semantic classes. A similar position has been argued convincingly by Apóstol (2014) for the counterpart classes in the Amuzgo language of Xochistlahuaca.

As we will see further below, the distinction between inactive and active verbs is relevant for the split we are studying. Active verbs have intricate morphology. For example, they use tone as a marker of person of the subject. As for tone, SP Amuzgo has eight tones involved in lexical contrasts (some of which are also found in verbal inflection). Following Smith-Stark & Tapia's (1986, 2000) notation, we represent tone using numbers. In the scale, number

Table 8: The eight tones of SP Amuzgo.

Kim (in press)		Buck (2000)
$\eta\epsilon^1$	‘grass’	L
$\eta\epsilon^{12}$	‘vegetation’	LL
$\eta\epsilon^3$	‘cheap’	M
$\eta\epsilon^{34}$	‘shudder.CPL.3SG’	MM
$\eta\epsilon^{35}$	‘ripe’	MH
$\eta\epsilon^5$	‘dust’	H
$\eta\epsilon^{51}$	‘already’	HL
$\eta\epsilon^{53}$	‘sell.CPL.1SG’	HM

‘1’ represents low tone and ‘5’ represents high tone. The minimal octuplet in Table 8 is from Kim (in press) to illustrate the eight tones.⁷

In the octuplet in Table 8, the words with tones /34/ and /53/ are inflected forms for the completive of 3rd and 1st person, respectively. We treat the tone pattern in 3rd person forms, singular or plural, as representing lexical tone. The 2nd person plural form always has the same tone as 3rd person plural, while 1st person plural exclusive is always marked by /51/. For the marking of the remaining three person values (i.e., 1st person singular, 2nd person singular, and 1st person plural inclusive), there is tonal allomorphy across the different paradigms. Such allomorphy can be seen as dividing the lexicon into tone classes whose membership is dictated by the lexicon; for details see Kim (in press).

Apart from tone, most verbs whose stems end in /e/ or /o/ also undergo vowel lowering to /ɛ/ and /ɔ/, respectively, in the 1st person singular and 1st person plural inclusive. In the realization of 1st person plural inclusive, there is often an echo vowel as well, which can be seen for example in the form $b\eta^{i2}i^5$ ‘we (INCL) get sick’ in Table 6, contrasting with $b\eta^{i51}$ ‘we (EXCL) get sick’ without an echo vowel and with a different tone.

Glottalization is yet another important parameter in the building of inflectional paradigms for SP Amuzgo. While the 2nd person singular is always marked with a suffixed -ʔ, there is at times a high degree of stem allomorphy involving glottalization patterns which Kim (to appear) proposes to organize into five morphophonological classes. For example, the verb $bi\eta^{i2}$ ‘get sick’ in Table 6 has two different stems involving a glottal: /CʔV/ and /CVʔ/. For the glottalization class to which this verb belongs, /CʔV/ is used for 1st person singular and

⁷ In Table 8, we follow Kim’s (to appear) treatment of tone /31/ in FT’s database as /51/, following Buck’s (2000) treatment of the same tone as /HIGH-LOW/.

all persons in the plural, while /CVʔ/ is used for both 2nd and 3rd person singular.⁸

Having presented the basics of Amuzgo verbal inflection, we are now in a position to consider the number split in greater depth. We begin by presenting the general characteristics of this split and then go on to investigate the way it is realized by the morphology.

3 The number split in Amuzgo: Singular vs. plural subject

3.1 The split illustrated

As we have seen in Table 1, a verb in SP Amuzgo may present a split in its paradigm between the forms that index a singular subject and those indexing a plural subject. The split only involves the number of the subject, i.e., it is insensitive to the number of the object. The association with the subject is shown in examples (1a, b) with the two different stems of the transitive verb *ʔwa¹* ‘cut with a machete’. Examples are from Buck (2000: 52, 72), but glosses, tones, the modified orthography, and the translation are our own. Example (1b) additionally shows that the language has a VSO word order.

- (1) a. *t⁵kiu²12* *ta²ʔ³* [*lu⁵-chh⁵ʔ³*]_{SUBJ.SG} *na³*
 a.while.ago CPL.start[3SG] boy-small=DET.SG COMP
i⁵-ʔwa¹=hu⁵ [*ndu¹*]_{OBJ.SG}
 3SG.INCPL-cut.SG[3SG]=PRO3SG.HUM_i wood
 ‘The boy started cutting wood a while ago.’
- b. *hndε³⁵* *ko³-tʔwa³* [*nã³⁴=min²ʔ³*]_{SUBJ.PL} [*hndε¹²-tke³=ba²ʔ³*]_{OBJ.SG}
 already PL.INCPL-cut.PL[3PL] people=DEM.PL woodland-old=DEM.SG
 ‘Those people are already clearing out (i.e., cutting) this old woodland.’

These examples show that a different stem of the verb is required when the subject is singular than when it is plural, i.e., *-ʔwa¹* is used in (1a) for a singular subject (note that the two clauses in (1a) have the same subject; the pronominal clitic =*hu⁵* in the subordinated clause cross-references the NP referring to the

⁸ Suffixing *-ʔ* for 2nd person to a stem that already bears a glottal in the coda like /CVʔ/ results in the elision of one of the glottals.

‘boy’ in the main clause) and *-tɔwa³* is used in (1b) for a plural subject. The crucial thing is that in both clauses the object remains invariably singular, which provides evidence that the different stems reflect the number of the subject and not that of the object. As the split in Amuzgo is concerned with number and it involves verbs, it could be taken to be an instance of verbal number, but in reality it is not, as we argue in the next section.

3.2 Verbal number typologically and in Amuzgo

Verbal number has received recent attention in linguistic typology given the diversity of grammatical and lexical constructions involving plurality and verbs across languages (for a recent comprehensive review see, for example, Cabredo Hofherr (2010) and Cabredo Hofherr & Laca (2012)). “Verbal number” is commonly used as an umbrella term to cover verbal constructions elaborating on either the repetition or the multiplicity of events or the plurality of affected participants in events. Following Corbett’s (2000) taxonomy of verbal number, the first type is commonly treated as “event number” and the markers involved in its encoding are known as “pluractionals”,⁹ while the second type is known as “participant number”. The number split of Amuzgo is different from both of these, as we will show next.

Event number is illustrated by the contrasting examples in Hausa presented in (2), cited in Corbett (2000: 246) from Eulenberg (1971: 73–74), with glossing inspired by Caron (2012) and tones added courtesy of Bernard Caron (personal communication).¹⁰

- (2) a. *naa áíkèè sù*
 1SG.PRF send 3PL.OBJ
 ‘I sent them.’
 b. *naa á’àikéé sù*
 1SG.PRF RED/send 3PL.OBJ
 ‘I sent them.’

⁹ The term “pluractional” is from Newman (1980) where it describes a particular class of derived verb stems in Chadic languages, as in example (2) from Hausa. In the languages of the Americas, pluractionals are also found. See, for example, Henderson (2012) for the Mayan languages and Gillon & Mailhammer (2015) for Yuman.

¹⁰ The transitive verb is inflected in Grade 4 indicating that the action is finished and the object is fully affected (Bernard Caron, personal communication).

In contrast to (2a), the verb *aikaa* ‘send’ appears in (2b) in a reduplicated form which expresses plurality of actions. Both examples mean ‘I sent them’, both involve a plural object, and both can be used for a variety of event situations, e.g., the sending was at the same time to different places, at different times to the same place, or at different times to different places. However, only (2a) can be used for a situation where the sending was at the same time to the same place for which (2b) is ungrammatical. This restriction means that the reduplicating construction in (2b) is only used when there is some multiplicity of acts of sending (whether in the initial stages, i.e., the packaging, or in the final stages, i.e., the delivering).

Participant number, meanwhile, involves plurality of participants and often comes in the form of verbal suppletion. This is illustrated in Hiw, an Oceanic non-Polynesian language of Vanuatu, in the lexical contrast between (3a) and (3b), from François (2009).

- (3) a. *nine sō* (**nine iw*)
 3SG fall.SG
 ‘S/he fell.’
 b. *sise iw* (**sise sō*)
 3PL fall.PL
 ‘They fell.’

Participant number as found in Hiw bears a strong resemblance to the split we observe in Amuzgo verbs, but it is different. In languages with typical participant number, like Hiw, the selection of a verb in a number pair (or of the relevant plural or dual inflection) responds to the number of the subject with intransitive verbs, but to the number of the object with transitive verbs. This is shown in the transitive pair in (4), also from Hiw (François 2009).

- (4) a. *temāṛēṛē peon not* [*i noke*]_{OBJ.SG}
 old.man FUT kill.SG OBJ 1SG
 ‘The ogre is going to kill me!’
 b. *temāṛēṛē peon qētṅog* [*i tite*]_{OBJ.PL}
 old.man FUT kill.PL OBJ 1PL.INCL
 ‘The ogre is going to kill us!’

This pattern of alignment is accounted for by thinking of both S and O as the most affected actants in the dynamism of the event encoded by the verb, while the A actant would remain a non-affected, external participant. Participant number, as present in Hiw, is well-documented in the Polynesian and the

Uto-Aztec language families, but it is also found in many other languages with a more sporadic distribution, for example, in isolate Ainu (Tamura 2000), in Koasati (Muskogean) (Mithun 1999), and in Shipibo-Conibo (Valenzuela 1997).

In languages with participant number, the phenomenon involves a restricted set of lexical pairs which rarely amounts to more than 30.¹¹ Within the Oto-Manguean phylum to which Amuzgo belongs, we also find participant number. In Lealao Chinantec, possibly the Oto-Manguean language with the greatest number of such pairs, the list amounts to 12 pairs (at least considering the data in the dictionary in Rupp & de Rupp 1996). In SP Amuzgo, participant number is restricted to the three pairs (a–c) in Table 9, i.e., it only involves six transitive verbs in the language in total, and the list is exhaustive. Notice that the six verbs in Table 9 also display the split involving number of the subject. This is strong evidence that the two splits represent two different and mutually independent phenomena.

Table 9: The six verbs with participant number in SP Amuzgo.

	SG.SUBJ	PL.SUBJ	
a.	<i>tskwe</i> ⁵	<i>nãʔ¹kwe</i> ⁵	‘kill (SG.OBJ)’
	<i>tskwhe</i> ⁵	<i>nãʔ¹kwhe</i> ⁵	‘kill (PL.OBJ)’
b.	<i>bʔa</i> ³	<i>tʔa</i> ¹²	‘grab (SG.OBJ)’
	<i>bha</i> ³	<i>tha</i> ¹²	‘grab (PL.OBJ)’
c.	<i>yʔõ</i> ³⁵	<i>yʔõ</i> ¹	‘take (SG.OBJ)’
	<i>chu</i> ⁵	<i>cho</i> ¹	‘take (PL.OBJ)’

3.3 The distribution of the number split of Amuzgo

Although in the split the selection of the two morphological forms (i.e., the two stems) is triggered by the syntax (i.e., a rule of subject agreement in number), not all verbs display it. To illustrate this, compare the four active verbs in Table 10. As shown in Table 10, the verbs *nchhe*³ ‘smash’ and *xke*³ ‘(for a woman to) grow old’ have invariant stems throughout their paradigms. This contrasts with the suppletion we find for verbs *nchht*⁵³~*ngio*¹ ‘know’ and *xʔue*³⁵~*hndyʔue*³⁵ ‘chew’. The distribution of the number split in our sample of 579 verbs is given in Table 11, where we show that almost 70% of the verbs in the sample manifest the split in some way or another.

¹¹ Commonly, the verbs in question mainly involve motion, location, and a few highly manipulative actions.

Table 10: Verbs with and without the number split.

CPL	'smash'	'grow old'	'know, feel'	'chew'
1SG	<i>nchh</i> ⟨ε⟩ ⁵³	<i>t-(nd)</i> ⁱ⁵ + <i>xk</i> ⟨ε⟩ ³	<i>t-(nd)</i> ⁱ⁵ + <i>nchhi</i> ⁵³	<i>xʔu</i> ⟨ε⟩ ⁵³
2SG	<i>nchhe</i> ^{3-ʔ}	<i>t-(nd)</i> ⁱ⁵ + <i>xke</i> ^{3-ʔ}	<i>t-(nd)</i> ⁱ⁵ + <i>nchhi</i> ^{53-ʔ}	<i>xʔue</i> ^{53-ʔ}
3SG	<i>nchhe</i> ³	<i>t-(nd)</i> ⁱ⁵ + <i>xke</i> ³	<i>t-(nd)</i> ⁱ⁵ + <i>nchhi</i> ³⁵	<i>xʔue</i> ³⁵
1PL.INCL	<i>nchh</i> ⟨ε⟩ ³⁴	<i>t-(nd)</i> ⁱ⁵ + <i>xk</i> ⟨ε⟩ ³⁴	<i>t-(nd)</i> ⁱ⁵ + <i>ngi</i> ⟨ɔ⟩ ¹²	<i>hndyʔu</i> ⟨ε⟩ ³⁵ ⟨ε⟩ ⁵
1PL.EXCL	<i>nchhe</i> ⁵¹	<i>t-(nd)</i> ⁱ⁵ + <i>xke</i> ⁵¹	<i>t-(nd)</i> ⁱ⁵ + <i>ngio</i> ⁵¹	<i>hndyʔue</i> ⁵¹
2PL	<i>nchhe</i> ^{3=ɔʔ3}	<i>t-(nd)</i> ⁱ⁵ + <i>xke</i> ^{3=ɔʔ3}	<i>t-(nd)</i> ⁱ⁵ + <i>ngio</i> ^{1=ɔʔ3}	<i>hndyʔue</i> ^{35=ɔʔ3}
3PL	<i>nchhe</i> ^{3(=ho5)}	<i>t-(nd)</i> ⁱ⁵ + <i>xke</i> ^{3(=ho5)}	<i>t-(nd)</i> ⁱ⁵ + <i>ngio</i> ^{1(=ho5)}	<i>hndyʔue</i> ^{35(=ho5)}

Table 11: Distribution of the split in the sample.

SG-PL split	No split	Total
393	186	579
68%	32%	100%

Whether or not a verb will display the split is not predictable. We only find tendencies. For example, the split is typically associated with active verbs. This is shown in Table 12(a), where more than 95% of the verbs with the split are active. However, Table 12(b) shows that this does not mean that all active verbs have the split: almost 20% of them do not. Similarly, while 85% of the inactive verbs do NOT exhibit a split, we still have 18 verbs in the sample that do. These tendencies suggest that the split is linked to the division of verbs into inactive and active classes, but still somehow independent from it.

Table 12: The distribution of the number split across inactive/active verbs.

a.	SG-PL split		No split	
Active	375	95.5%	84	45%
Inactive	18	4.5%	102	55%
Total	393	100%	186	100%
b.	Active		Inactive	
SG-PL split	375	81.5%	18	15%
No split	84	18.5%	102	85%
Total	459	100%	120	100%

Table 13: The distribution of the number split in relation to transitivity.

	Transitive		Intransitive	
SG-PL split	296	98%	97	35%
No split	7	2%	179	65%
Total	303	100%	276	100%

The number split also shows a relevant correlation with transitivity: 98% of transitive verbs have the split, whereas it is only observed in 35% of intransitive verbs (Table 13).

We may look at the same figures from a different point of view, but we obtain similar results. This is shown in Table 14, where we can see that three quarters of the verbs that display the split are transitive (i.e., 296 verbs out of 393), while 96% of the ones that do not are intransitive (179 out of 186). The fact that the split is strongly associated with transitive verbs equally shows that we are dealing with a very different phenomenon to participant number.

Table 14: The distribution of transitivity in relation to the number split.

	SG-PL split		No split	
Transitive	296	75%	7	4%
Intransitive	97	25%	179	96%
Total	393	100%	186	100%

3.4 Domains where the number split is observed

The split as such may be observed in two different domains: (i) in the domain of the marking of TAM, where it is found in the encoding of the incomplete within a particular class of verbs; and (ii) in the domain of the verbal stem, where it is exploited in various ways, as we will show in detail in Section 5. All examples given so far that show the split (in Tables 1, 6, and 10) are instances of stem-based splits. While the two domains work independently of each other, when all verbs are taken into account almost 80% of them display the singular-plural split twice, i.e., in both domains (examples of a doubly-marked split are given in Tables 22, 25, and 53 below). This is shown in Table 15.

Table 15: Verbs according to the type of their number split.

TAM-based split only	33	8.5%
Stem-based split only	55	14.0%
Both TAM-based & stem-based split	305	77.5%
Total	393	100%

When transitivity is factored in, however, transitive verbs have an overwhelming preference for double marking: not only are 92% of the verbs that mark the split twice transitive, as shown in Table 16(a), but also 94.5% of all transitive verbs mark it twice, as shown in Table 16(b). Similarly, intransitive verbs prefer a stem-based only split.

Table 16: Number split type according to transitivity.

a.	TAM-based		Stem-based		Both	
Transitive	12	36.5%	4	7%	280	92%
Intransitive	21	63.5%	51	93%	25	8%
Total	33	100%	55	100%	305	100%
b.	Transitive		Intransitive			
TAM-based only	12	4.0%	21	21.5%		
Stem-based only	4	1.5%	51	52.5%		
Both	280	94.5%	25	26%		
Total	296	100%	97	100%		

But despite such correlations, since the split is not predictable one cannot overlook the fact that there is a breach in Amuzgo grammar affecting the mapping between what is required by the syntax (i.e., agreement in number of subject) and what is actually reflected by the morphology. This breach has consequences for the morphosyntax, as it has to be filtered out through information stored in the lexicon. The total number of verbs in our sample undergoing some type of stem change is 360. Of these, as we will see in Section 5.6, the only ones for which the form of the plural stem is predictable from the singular and vice versa are verbs bearing the stem formative pair $ts(i)^5+/nã^2+$. Such verbs amount to 190 in our sample. For the remaining 170 verbs, speakers need to remember two stems independently.¹²

¹² Of course this only applies to the verbs of our sample. If more verbs were added, the number of verbs for which remembering two stems is required would obviously increase.

In the following two sections, we introduce the way the number split is manifested in these two domains. But as the split at stem level is more intricate due to the existence of a great deal of allomorphy, we begin with the TAM-based split next.

4 The TAM-based number split

Based on FT's data, verbs in San Pedro Amuzgo can be classified as belonging to at least six different inflectional classes depending on how they mark the incomplete aspect. The phenomenon is illustrated by the seven verbs in Table 17. To appreciate the morphology involved in the encoding of the incomplete more clearly, we provide the forms for the complete for comparison. In the segmentation, composite elements in a verbal compound are indicated by an en-dash (-).

As shown in Table 17, the classes are unequal in size. We focus on the first three. Here the distribution of the markers for the incomplete reveals a singular vs. plural split, where the persons of the plural all receive the same marker *ko-*, while the persons of the singular receive different encodings. Class 1 verbs have a marker *ma*³- for both 1st and 2nd persons singular and *ʔ*⁵- for 3rd person singular. This contrasts with Classes 2 and 3 which do not differentiate person in the singular; they have either *ba-* or zero marking. It is interesting that the plural marker *ko-* of Classes 1–3 is homophonous to the generalized incomplete marker for verbs of Class 4.¹³ The same happens for the *ba-* and the zero marking of the singular of Classes 2 and 3, which are used as generalized incomplete markers for Classes 5 and 6, respectively.

The figures in Table 18 show how the verbs in the sample are distributed across the incomplete classes according to whether they display a TAM-based split or not and to whether they are inactive or active. Table 18 reveals that 58% of the verbs in the sample display a TAM-based split involving the marking for the incomplete. Of these, only three are inactive. This means that if a verb is inactive it will most likely not have a TAM-based split. Being active is, on the

¹³ One way to interpret this is to view inactive verbs as derived from Class 1 verbs. There is some evidence which points in this direction, although a proper study of this phenomenon needs to be left for future research. For example, *ʔ*⁵*wa*¹ 'cut' is a transitive verb of Class 1 (3SG.INCPL *ʔ*⁵-*ʔ*¹*wa*¹=*hō*⁵³ vs. 3PL.INCPL *ko*³-*t*-*ʔ**wa*²=*ho*²). Its middle counterpart 'cut oneself by accident' is instead a Class 4 inactive verb (3SG.INCPL *ko*³-*t*-*ʔ**wa*¹=*ē*³ vs. 3PL.INCPL *ko*³-*t*-*ʔ**wa*¹=*ē*³). Notice that the middle verb, as a Class 4 verb, does not only use *ko*³- in all the forms of the incomplete but also uses the plural stem for all forms.

Table 17: Inflectional classes of verbs in SP Amuzgo for the marking of the incomplete.

INCL	Class 1 'whistle' (active)	Class 2 'lose weight' (inactive)	Class 3 'come' (active)	Class 4 'bleed heavily' (active)	Class 5 'learn' (active)	Class 6 'be cold' (active)
1SG	$m^3-tiu^{51}\langle u \rangle^3$	$ba^5-ndi^5+ki^5hne^1=ha^3$	$\emptyset-ndy\langle u \rangle^{53}$	$ko^3-ndui^1-n\eta\langle u \rangle^3$	$ba^5-ndi^5+hn\eta a^5$	$\emptyset-xi\langle u \rangle^{53}$
2SG	$m^3-tiu^1-\eta$	$ba^5-ndi^5+ki^5hne^1=\eta^3$	$\emptyset-ndy\langle o \rangle^3-\eta$	$ko^3-ndui^1-nio^{51}-\eta$	$ba^5-ndi^5+hna^{51}-\eta$	$\emptyset-xio^{12}-\eta$
3SG	η^5-tiu^1	$ba^5-ndi^5+ki^5hne^1=e^5$	$\emptyset-ndyo^1$	$ko^3-ndui^1-nio\eta^{34}$	$ba^5-ndi^5+hna\eta^5$	$\emptyset-xio^3$
1PL.INCL	ko^3-tiu^{34}	$ko^3-ndi^5+ki^5hne^1=ha^5\langle a \rangle^5$	$ko^3-nki\langle o \rangle^{34}$	$ko^3-ndui^1-n\eta\langle u \rangle^{34}$	$ba^5-ndi^5+hn\eta a^5\langle \bar{a} \rangle^5$	$\emptyset-xi\langle u \rangle^{35}\langle u \rangle^5$
1PL.EXCL	ko^3-tiu^{51}	$ko^3-ndi^5+ki^5hne^1=ha^{51}$	$ko^3-nki\langle o \rangle^{51}$	$ko^3-ndui^1-n\eta\langle io \rangle^{51}$	$ba^5-ndi^5+hn\eta a^5$	$\emptyset-xio^{51}$
2PL	$ko^3-tiu^3=0\eta^3$	$ko^3-ndi^5+ki^5hne^1=\eta^3$	$ko^3-nki\langle o \rangle^3=0\eta^3$	$ko^3-ndui^1-n\eta\langle io \rangle^{34}=0\eta^3$	$ba^5-ndi^5+hn\eta a^5=0\eta^3$	$\emptyset-xio^3=0\eta^3$
3PL	$ko^3-tiu^3=ho^5$	$ko^3-ndi^5+ki^5hne^1=e^3$	$ko^3-nki\langle o \rangle^5=ho^5$	$ko^3-ndui^1-n\eta\langle io \rangle^{34}=ho^5$	$ba^5-ndi^5+hn\eta a^5=ho^5$	$\emptyset-xio^3=ho^5$
CPL						
1SG	$tiu^{51}\langle u \rangle^3$	$t-\eta^5+ki^5hne^1=ha^3$	$hndy\langle u \rangle^{53}$	$hndui^1-n\eta\langle u \rangle^3$	$t-\eta^5+hn\eta a^5$	$t-\eta^5+xi\langle u \rangle^{53}$
2SG	$tiu^1-\eta$	$t-\eta^5+ki^5hne^1=\eta^3$	$hndy\langle o \rangle^3-\eta$	$hndui^1-nio^{51}-\eta$	$t-\eta^5+hna^{51}-\eta$	$t-\eta^5+xio^{12}-\eta$
3SG	tiu^1	$t-\eta^5+ki^5hne^1=e^5$	$hndyo^1$	$hndui^1-nio\eta^{34}$	$t-\eta^5+hna\eta^5$	$t-\eta^5+xio^3$
1PL.INCL	tiu^{34}	$t-\eta^5+ki^5hne^1=ha^5\langle a \rangle^5$	$ty\eta^3-nki\langle u \rangle^{34}$	$hndui^1-n\eta\langle u \rangle^{34}$	$t-\eta^5+hn\eta a^5\langle \bar{a} \rangle^5$	$t-\eta^5+xi\langle u \rangle^{35}\langle u \rangle^5$
1PL.EXCL	tiu^{51}	$t-\eta^5+ki^5hne^1=ha^{51}$	$ty\eta^3-nki\langle o \rangle^{51}$	$hndui^1-n\eta\langle io \rangle^{51}$	$t-\eta^5+hn\eta a^5$	$t-\eta^5+xio^{51}$
2PL	$tiu^3=0\eta^3$	$t-\eta^5+ki^5hne^1=\eta^3$	$ty\eta^3-nki\langle o \rangle^3=0\eta^3$	$hndui^1-n\eta\langle io \rangle^{34}=0\eta^3$	$t-\eta^5+hn\eta a^5=0\eta^3$	$t-\eta^5+xio^3=0\eta^3$
3PL	$tiu^3=ho^5$	$t-\eta^5+ki^5hne^1=e^3$	$ty\eta^3-nki\langle o \rangle^5=ho^5$	$hndui^1-n\eta\langle io \rangle^{34}=ho^5$	$t-\eta^5+hn\eta a^5=ho^5$	$t-\eta^5+xio^3=ho^5$
Total	324	2	12	50	72	119

Table 18: The distribution of the TAM-based split across inactive/active verbs.

		Incomplete inflectional class						Total
		1	2	3	4	5	6	
TAM-based split		SG-PL split			No split			
		338			241			
		58%			42%			
Inactive	Intransitive	2	1	0	38	46	31	118
	Transitive	0	0	0	1	0	0	1
Active	Intransitive	40	2	1	10	26	79	158
	Transitive	282	0	10	1	0	9	302
Total		324	3	11	50	72	119	579

other hand, uninformative: out of the 460 active verbs, while 335 display a TAM-based split, there still remain 125 that do not. In contrast, we can see a strong correlation with transitivity: 282 out of the 302 transitive (active) verbs in the sample are members of Class 1 and thus display the TAM-based split. The profile of Classes 4–6, not having the split, is more heterogeneous.

Table 19 gives an indication of how the incomplete classes correlate with the 360 verbs that have a stem-based split. The overall majority corresponds to Class 1 verbs (81%), and since Class 1 verbs by definition have a TAM-based split, they end up showing the split at least twice in the relevant forms. If we sum up all verbs of Classes 4–6 in Table 19 we obtain 55 of them; most significantly, 51 of them (93%) are intransitive.

Table 19: Correlation between the TAM-based split and the stem-based split.

Incomplete inflectional class		Stem-based split	Intransitive	Transitive	
SG-PL split	Class 1	292	81%	22	270
	Class 2	2	0.7%	2	0
	Class 3	11	3%	1	10
No split	Class 4	1	0.3%	1	0
	Class 5	11	3%	11	0
	Class 6	43	12%	39	4
Total		360	100%	76	284

We end this section by noting that the incomplete as an aspectual value is not a category which is intuitively related to the agreement in number with the subject. Consequently, a TAM-based split as witnessed by the contrasts involved in Classes 1–3 could be seen as epiphenomenal to number agreement. As a

result, showing that Amuzgo verbs also rely on stem changes to produce the number split is fundamental to our argumentation. In Table 19 we are reminded that 360 verbs in the sample display a stem-based split. In the following section we delve into the morphological intricacies of such stems.

5 The split at the level of the stem

5.1 Stem doublets

Verbs in Amuzgo may also display a number split in the domain of the verbal stem. For this, verbs have two stems, one used to build the inflected forms when the subject is singular and another when the subject is plural. In our sample of 579 verbs, 360 (62%) display the number split at the level of the stem in some way or another. Whether a given verbal lexeme will display the split or not is not predictable, and in principle, speakers need to know which lexemes do and which do not. But to make matters more interesting from a morphological point of view, for those that do, speakers also need to know how the two stems differ from one another.

In this respect, we have identified a set of formal ways or strategies by which the two stems may differ, namely: (i) by means of glottalization; (ii) by replacing a final vowel; (iii) by adding the prefix *t-*; (iv) by adding the prefix *n-*; (v) by means of stem formatives; (vi) by a combination of at least two of the aforementioned mechanisms; (vii) by having a suppletive stem; or (viii) by a change in tone. In Table 20, we give an indication of the frequency of such operations in our sample together with an example.

Table 20: Operations producing the stem-based split.

	Strategy	#	%	Singular	Plural	
(i)	Glottalization	18	5%	<i>hna³no³</i>	<i>hnaʔ¹no¹</i>	‘flee’
(ii)	Final vowel	11	3%	<i>ndy-i³⁵</i>	<i>ndy-e¹</i>	‘hear’
(iii)	Prefix <i>t-</i>	69	19%	<i>ki⁵</i>	<i>t-ki⁵</i>	‘eat’
(iv)	Prefix <i>n-</i>	12	3.5%	<i>tykwāʔ³</i>	<i>n-tykwāʔ³</i>	‘have a handicap’
(v)	Stem formatives	190	53%	<i>tsi⁵+tyʔue³</i>	<i>nāʔ¹+tyʔue³</i>	‘scare’
(vi)	Combination	9	2.5%	<i>skuʔ^{1,2}</i>	<i>n-t-ku^{1,2}</i>	‘have a wife’
(vii)	Suppletion	24	6.5%	<i>tskuʔ¹</i>	<i>tɛʔ³</i>	‘close’
(viii)	Tone only	27	7.5%	<i>nduiʔ¹</i>	<i>nduiʔ³</i>	‘go out’
	Total	360	100%			

The strategy involving stem formatives is the most frequent and is also the most transparent: knowing that a verb in the singular has a stem formative of a certain type would enable us to predict what stem formative the plural forms take. For the other cases, making predictions is a little less obvious. All operations are unproductive and we suggest that treating the stem contrasts as cases of suppletion is perhaps the best way to characterize the variance.

Another relevant factor that reinforces the relation between the singular and the plural stems as being one of suppletion, whether partial or total, is the intricate use of tone. In Table 20, we observe that there are 27 verbs whose stems contrast in tone only (we will further see in Section 5.9 that they fall into different tone classes). However, at least 106 of the remaining verbs, whose stems contrast by means of one of the different segmental encoding strategies (i–vii), also use tone as an additional encoding strategy. The details are given in Table 21.

Table 21: Distribution of tone in stem-based splits.

Without tone change	225	63%			
With tone change	133	37%	Tone only	27	20%
			Tone change in addition to another segmental change	106	80%
Total	360	100%	Total	133	100%

Besides, as we have already pointed out in Table 15 above, 305 (85%) out of the 360 verbs that have a stem-based split belong to one of the three classes that also reflect the number split in the incomplete. As a result, when a verb of these classes is inflected in the incomplete, the plural forms manifest the split twice, that is, in two domains: in both the affixal and the stem domains. This is shown in Table 22a with the verb *chʔue*³⁵ ‘steal, pick’ which has its plural stem

Table 22: Verbs manifesting the singular-plural split by multiple means.

INCP	a. ‘steal, pick’	b. ‘tread’	c. ‘pick up’
1SG	<i>ma</i> ³ - <i>chʔu</i> ⟨ <i>ε</i> ⟩ ⁵³	<i>ma</i> ³ - <i>ndyu</i> - <i>i</i> ⁵³	<i>ma</i> ³ - <i>ba</i> ⁵ <i>nkhe</i> ⟨ <i>ε</i> ⟩ ^{ʔ35} ⟨ <i>ε</i> ⟩ ³
2SG	<i>ma</i> ³ - <i>chʔue</i> ⁵³ - <i>ʔ</i>	<i>ma</i> ³ - <i>ndyu</i> - <i>i</i> ⁵³ - <i>ʔ</i>	<i>ma</i> ³ - <i>ba</i> ⁵ <i>nkhe</i> ³¹ - <i>ʔ</i>
3SG	<i>ʔ</i> ⁵ - <i>chʔue</i> ³⁵	<i>ʔ</i> ⁵ - <i>ndyu</i> - <i>i</i> ³⁵	<i>ʔ</i> ⁵ - <i>ba</i> ³ <i>nkhe</i> ^{ʔ35}
1PL.INCL	<i>ko</i> ³ - <i>n-dyʔu</i> ⟨ <i>ε</i> ⟩ ³⁵ ⟨ <i>ε</i> ⟩ ⁵	<i>ko</i> ³ - <i>ndyu</i> -⟨ <i>ε</i> ⟩ ¹²	<i>ko</i> ³ - <i>t</i> -(<i>b</i>) <i>a</i> ¹ <i>ʔnkhe</i> ⟨ <i>ε</i> ⟩ ^{ʔ35} ⟨ <i>ε</i> ⟩ ⁵
1PL.EXCL	<i>ko</i> ³ - <i>n-dyʔue</i> ⁵¹	<i>ko</i> ³ - <i>ndyu</i> - <i>e</i> ⁵¹	<i>ko</i> ³ - <i>t</i> -(<i>b</i>) <i>a</i> ¹ <i>ʔnkhe</i> ^{ʔ51} <i>ε</i> ³
2PL	<i>ko</i> ³ - <i>n-dyʔue</i> ³⁵ = <i>oʔ</i> ³	<i>ko</i> ³ - <i>ndyu</i> - <i>e</i> ¹ = <i>oʔ</i> ³	<i>ko</i> ³ - <i>t</i> -(<i>b</i>) <i>a</i> ¹ <i>ʔnkhe</i> ^{ʔ35} = <i>oʔ</i> ³
3PL	<i>ko</i> ³ - <i>n-dyʔue</i> ³⁵ = <i>ho</i> ⁵	<i>ko</i> ³ - <i>ndyu</i> - <i>e</i> ¹ = <i>ho</i> ⁵	<i>ko</i> ³ - <i>t</i> -(<i>b</i>) <i>a</i> ¹ <i>ʔnkhe</i> ^{ʔ35} = <i>ho</i> ⁵

marked with a nasal prefix (plus a predictable morphophonological adjustment to the root) and the incompletive prefix *ko-*. Note that tone does not change; we find the same tone /35/ in the stem of both forms for the 3rd person. When the verb employs a combination of stem-based strategies, the inflected forms may display the split in various places. We illustrate this with the verb *ndyui*³⁵ ‘tread’ in Table 22, whose plural stem contrasts with the singular stem in both the final vowel and in tone. In more extreme cases, like the verb *ba*³*nkhe*²³⁵ ‘pick up’, we may find plural subject marked in four places in an inflected form: by (i) the presence of the prefix *t-*, (ii) the change in tone in the /ba/ segment in the first syllable of the stem; (iii) the glottalization at root level; and (iv) the incompletive prefix *ko-*.

In order to understand the nature of the contrasts between the singular and the plural stems, we take here the singular as a point of departure. The rampant arbitrariness of the different possibilities makes it clear that the derivation does not obey any transparent or productive rule, at least synchronically. In the following sections, we will present each strategy in some detail.

5.2 Stems contrasting in glottalization

Some verbs in Amuzgo have two stems that contrast in glottalization: one of them is used for the singular and the other for the plural. Examples are given in Table 23. The distribution of the stems in these examples is according to the number of the subject. This is a different phenomenon from the glottalization classes mentioned in Section 2.3, as proposed in Kim (to appear), where the two stems (glottalized and non-glottalized) are distributed according to more morphological (i.e., less meaningful) patterns, as seen, for example, with the verb *bi*²¹² ‘get sick’ in Table 6.

Table 23: Two verbs with stems contrasting in glottalization.

CPL	‘flee’	‘get scared’
1SG	<i>hna</i> ³ <i>n</i> ⟨ɔ⟩ ⁵³	<i>tyʔu</i> ⟨ɛ⟩ ¹
2SG	<i>hna</i> ³ <i>no</i> - <i>ʔ</i> ¹²	<i>tyʔue</i> - <i>ʔ</i> ³
3SG	<i>hna</i> ³ <i>no</i> ³	<i>tyʔue</i> ³
1PL.INCL	<i>hna</i> ^{ʔ1} <i>n</i> ⟨ɔ⟩ ¹²	<i>tyue</i> ³⁴
1PL.EXCL	<i>hna</i> ^{ʔ1} <i>no</i> ⁵¹	<i>tyue</i> ⁵¹
2PL	<i>hna</i> ^{ʔ1} <i>no</i> ¹ <i>o</i> ^{ʔ3}	<i>tyue</i> ³ <i>o</i> ^{ʔ3}
3PL	<i>hna</i> ^{ʔ1} <i>no</i> ¹ (<i>ho</i> ⁵)	<i>tyue</i> ³ (<i>ho</i> ⁵)

In Table 23, the singular–plural split is manifested by a contrast in the glottalization of the stem. There are 18 verbs like this in FT’s sample, given in Table 24. Of these, 15 others are like *hna³no³* ‘flee’ in having the glottalized stem in the plural while having a concomitant tone change /3/ vs. /1/, i.e., *hna³no³* vs. *hna¹no¹*.¹⁴ In contrast, the two verbs listed under (b) in Table 24 have the glottalized stem in the singular. The last verb, *nde[?]* ‘bathe’, shows glottal metathesis rather than the addition of a glottal.

Table 24: Verbs with stems contrasting in glottalization.

	Singular	Plural	
a.	<i>hna³no³</i>	<i>hna¹no¹</i>	‘flee’
	<i>te³hndi⁵</i>	<i>te¹hndi⁵</i>	‘support, help’
	<i>chui³nʔo⁵³</i>	<i>chui¹nʔo⁵³</i>	‘get confused’
	<i>be³+hndya¹</i>	<i>be¹+hndya¹</i>	‘cross’
	<i>be³+tʔio¹</i>	<i>be¹+tʔio¹</i>	‘cross’
	<i>be³+ndo³</i>	<i>be¹+ndo³</i>	‘wait’
	<i>ke³+nā³⁵</i>	<i>ke¹+nā³⁵</i>	‘remove’
	<i>ke³+ntho¹²</i>	<i>ke¹+ntho¹²</i>	‘run over’
	<i>ndyo³-tska³⁵</i>	<i>ndyo¹-tska³⁵</i>	‘drag’
	<i>ndo³nde³⁴</i>	<i>ndo¹-nde³⁴</i>	‘look up’
	<i>tyi³-xe³⁴</i>	<i>tyi¹xe³⁴</i>	‘feed’
	<i>ndyo³ndyi³⁴</i>	<i>ndyo¹ndyi³⁴</i>	‘lean out’
	<i>ndyo³tsʔa³⁵</i>	<i>ndyo¹tsʔa³⁵</i>	‘hug’
	<i>na³+no³</i>	<i>na¹+no¹</i>	‘run, flee’
	<i>na³+ba³-ndye¹</i>	<i>na¹+ba¹-ndye¹</i>	‘jump on the spot’
	<i>na³+ki³-ntyha³</i>	<i>na¹+ki¹-ntyha¹</i>	‘get up’
b.	<i>tyue³</i>	<i>tyue³</i>	‘get scared’
	<i>nō³</i>	<i>nō¹</i>	‘nominate’
c.	<i>nde[?]</i>	<i>ndʔe¹²</i>	‘bathe’

5.3 Stems contrasting in the final vowel

Eleven verbs in our sample have stems which contrast in the singular and plural with regard to their final vowel. Three representative examples are given in

¹⁴ We have taken seven verbs from S&S’s sample that display both glottalization and the tone change /3/ vs. /1/ as instances of this type of contrast despite the fact that we don’t have evidence from their entire paradigm to discard the possibility that they are members of a glottalization class (as described by Kim (to appear)).

Table 25: Three verbs with stems contrasting in the final vowel.

INCP.L	‘hear’	‘handle’	‘get drunk’
1SG	<i>ma³-ndy-i⁵³</i>	<i>ma³-ndyʔ-u⁵³</i>	<i>∅-ki⁵ndy-ɔ³</i>
2SG	<i>ma³-ndy-i⁵³-ʔ</i>	<i>ma³-ndy-u⁵³-ʔ</i>	<i>∅-ki⁵ndy-u³-ʔ</i>
3SG	<i>ʔ⁵-ndy-i³⁵</i>	<i>ʔ⁵-ndy-u³⁵</i>	<i>∅-ki⁵ndy-i⁵</i>
1PL.INCL	<i>ko³-ndy-ɛ¹²</i>	<i>ko³-ndyʔ-ɔ¹²</i>	<i>∅-ki⁵ndy-ɛ⁵ɛ⁵</i>
1PL.EXCL	<i>ko³-ndy-e⁵¹</i>	<i>ko³-ndyʔ-o⁵¹</i>	<i>∅-ki⁵ndy-e⁵¹</i>
2PL	<i>ko³-ndy-e¹ oʔ³</i>	<i>ko³-ndyʔ-o¹² oʔ³</i>	<i>∅-ki⁵ndy-e⁵ oʔ³</i>
3PL	<i>ko³-ndy-e¹ (ho⁵)</i>	<i>ko³-ndyʔ-o¹² (ho⁵)</i>	<i>∅-ki⁵ndy-e⁵ (ho⁵)</i>

Table 25.¹⁵ The verbs ‘hear’ and ‘handle’ have two stems: *ndyⁱ³⁵* and *ndyʔu⁵³* in the singular, respectively, and *ndye¹* and *ndyʔo¹²* in the plural. The two stems also contrast in tone, but we will return to this phenomenon in Section 5.9. The verb ‘get drunk’ is similar to these two verbs (i.e., it has *ki⁵ndyⁱ⁵* in the singular and *ki⁵ndye⁵* in the plural), but it has two further stems for 1st and 2nd persons singular: *ki⁵ndyo³* and *ki⁵ndyu³*, respectively. Related person splits are also attested elsewhere; we will review them in Section 5.10. The 11 verbs in our sample like the ones in Table 25 are given in Table 26.

Table 26: Verbs with stems contrasting in the final vowel.

Singular	Plural	
<i>ndy-i³⁵</i>	<i>ndy-e¹</i>	‘hear’
<i>tu-i³⁵</i>	<i>tu-e¹</i>	‘put in’
<i>ndyu-i³⁵</i>	<i>ndyu-e¹</i>	‘kick’, ‘tread’
<i>tʔu-i³⁵</i>	<i>tʔu-e¹</i>	‘grab’, ‘touch’
<i>ʔndy-i³⁵</i>	<i>ʔndy-e¹</i>	‘leave’
<i>ki⁵ndy-i⁵</i>	<i>ki⁵ndy-e⁵</i>	‘get drunk’
<i>ndi⁵+ʔhnd-i⁵</i>	<i>ndi⁵+ʔhnd-ɛ³</i>	‘be fat’, ‘be strong’
<i>ʔ⁵hnd-i</i>	<i>ʔ³hnd-ɛ⁵</i>	‘have strength’, ‘be fat’
<i>ndyo¹ch-u⁵</i>	<i>ndyo³ch-o¹</i>	‘bring’
<i>ndy-u^{ʔ35}</i>	<i>ndyʔ-o¹²</i>	‘handle’, ‘tempt’
<i>tyion³ na³ch-u⁵ hna⁵¹</i>	<i>tyion³ na³ch-o⁵ hna⁵¹</i>	‘pay a debt’

¹⁵ The verbs in Table 25 receive different marking for the incomplete because they belong to different classes, *ʔ⁵ndyⁱ³⁵* ‘hear’ and *ʔ⁵ndyu^{ʔ35}* ‘handle’ are from Class 1, while *ki⁵ndyⁱ⁵* ‘get drunk’ is from Class 6 (see Table 17 for more details).

5.4 Plural stems taking the prefix *t-*

There are 69 verbs in our sample whose plural stems bear a prefix *t-*. The mechanism is typically associated with transitive verbs: 56 of the 69 verbs are transitive. Examples of verbs bearing this prefix are presented in Table 27.

Most verbs bearing an alveolar prefix in the plural stem have some form of the prefix *t-*, but in FT's sample there are two verbs which have a syllabic prefix *ti³-*, like *tha⁵³* 'be missing'. Affixing *t-* produces a series of regular morphophonological adjustments to the onset of the stem, which are also indicated in Table 27. Such adjustments are similar to the ones observed in the forms for the completive, an aspect that is also marked by a homophonous prefix *t-*. There, the completive prefix is dropped before stems with an onset in /t/: e.g., 3SG.INCPL *t⁵-tiu¹* 'whistle' but 3SG.CPL *t-tiu¹* → *tiu¹*. Taking this into consideration, we have assumed that at least nine verbs which have an onset in /t/ in our sample are also examples of plural stems with *t-*. But to be on the safe side, we have only considered those that also display a tonal change. We have proceeded in this way based on the observation that of the other 60 verbs, no less than 48 also involve a tonal change (80%). An example of such a verb would be 'whistle': SG *tiu¹* vs. PL *t-tiu³* → *tiu³*. As we will see in Section 5.9, tonal changes are independent of the prefix *t-*, but their co-occurrence with this prefix is particularly common to justify taking these nine verbs as instances of the prefix *t-*.

Table 27: Verbs with the prefix *t-* in the plural.

Singular	Plural		→		
<i>ki⁵</i>	{ <i>t-ki⁵</i> }	/t-X/	→	<i>tki⁵</i>	'eat'
<i>bʔa³</i>	{ <i>t-bʔa¹²</i> }	/t-bX → t-X/	→	<i>tʔa¹²</i>	'grab', 'collect'
<i>ma¹</i>	{ <i>t-ma³</i> }		→	<i>toā³</i>	'wash'
<i>meʔ¹ ndaʔ</i>	{ <i>t-meʔ³</i> } <i>ndaʔ</i>	/t-mV → t-oV̄/	→	<i>toē³ ndaʔ</i>	'hit with one's hand'
<i>khaʔ³⁵</i>	{ <i>t-khaʔ³⁵</i> }	/t-kX → t-X/	→	<i>thaʔ³⁵</i>	'sting'
<i>kwhaʔ³⁵</i>	{ <i>t-kwhaʔ³⁵</i> }	/t-kwX → t-X/	→	<i>thaʔ³⁵</i>	'repair'
<i>ndo³ʔbiʔ¹</i>	{ <i>t-ndo¹²ʔbiʔ¹</i> }	/t-NX → NX/	→	<i>hndo¹²ʔbiʔ¹</i>	'curse (witchcraft)'
<i>tha⁵³</i>	{ <i>ti³-tha⁵³</i> }	/ti ³ -X/	→	<i>ti³tha⁵³</i>	'be missing'

5.5 Plural stems taking the prefix *n-*

Twelve verbs in our sample have a plural stem marked with a nasal prefix. These are given in Table 28. Notice the different stem adjustments in the plural, which are due to regular morphophonological rules. The nasal prefix is the

Table 28: Verbs with the prefix *n-*.

Singular	Plural		
<i>tykwāŋ³</i>	/nt ⁱ CV/	<i>n-tykwāŋ³</i>	‘have a handicap’, ‘be one-armed’
<i>ndi⁵+tsaŋ³</i>	/ntsV → n__V/	<i>ndi⁵+n-aŋ³</i>	‘go crazy’
<i>ndi⁵+tsŋ³</i>	/ntsŋV → n__ŋV/	<i>ndi⁵+n-ŋo³</i>	‘go stiff’
<i>tshwēŋ³</i>	/nts ^h V → n__hV/	<i>n-hwēŋ¹</i>	‘own an animal’
<i>xēŋ¹²</i>	/nŋV → nkV/	<i>n-kē¹²</i>	‘have a father- or brother-in-law’
<i>xhēŋ⁵</i>	/nŋ ^h V → nk ^h V/	<i>n-khē⁵</i>	‘have recently given birth’
<i>xŋē⁵</i>	/nŋŋV → nkŋV/	<i>n-kŋē⁵</i>	‘bother’
<i>ndi⁵+ki⁵.xŋē⁵</i>		<i>ndi⁵+n-kŋē⁵</i>	‘bother’
<i>ñeŋ¹+ki³.chiuŋ³⁵</i>	/ntŋV → nd ^h V/	<i>ñeŋ¹+ki³.n-dyŋiu¹²</i>	‘feel like urinating’
<i>chiuŋ³⁵</i>		<i>n-dyŋiu¹²</i>	‘urinate’
<i>chŋue³⁵</i>		<i>n-dyŋue³⁵</i>	‘rob’
<i>ndi⁵+chio³⁵</i>		<i>ndi⁵+n-dyio³⁵</i>	‘go pale’

Table 29: Plural nouns with the prefix *n-*.

Singular	Plural		
<i>tye³⁴</i>	/nt ⁱ /	<i>n-tye³⁴</i>	‘priest’
<i>tsōŋ⁵¹</i>	/ntsŋ̃V → n__V ^a /	<i>n-oŋ⁵¹</i>	‘drum’
<i>tshāŋ³</i>	/nts ^h V → n__hV/	<i>n-hāŋ³</i>	‘sweet potato’
<i>xēŋ¹</i>	/nŋV → nkV/	<i>n-kēŋ¹</i>	‘back’
<i>xho¹</i>	/nŋ ^h V → nk ^h V/	<i>n-kho¹</i>	‘metal’
<i>xŋō⁵³</i>	/nŋŋV → nkŋV/	<i>n-kŋō⁵³</i>	‘boil’
<i>chu³</i>	/ntŋV → nd ^h V/	<i>n-dyu³</i>	‘year’

a. The vowel of the plural stem loses its nasal component.

plural marker for nouns and similar regular morphophonological adjustments are found in nouns, see Buck (2000: 394–395). This is illustrated in Table 29. While the verbs which receive a nasal prefix in the plural stem remain largely arbitrary, it is noteworthy that out of the 18 inactive verbs in our sample that have the number split, 8 of them select this prefix. Given the link to nominal number, the occurrence of the prefix *n-* in these verbs may point to an old denominal origin of these inactive verbs.

5.6 Stems contrasting in stem formatives

Many verbs in Amuzgo manifest the singular–plural split by means of a stem contrast that involves stem formatives. There are two main types. The first type is

Table 30: A verb with the stem formatives $tsi^5+ \sim n\tilde{a}\rho^1+$.

INCPL	'buy'	Stems
1SG	$ma^3-tsi^5+hnda^{12}$	tsi^5+hnda^1
2SG	$ma^3-tsi^5+hnda^1-\rho$	
3SG	$\rho^5-tsi^5+hnda^1$	
1PL.INCL	$ko^3-n\tilde{a}\rho^1+hnda^{12}$	$n\tilde{a}\rho^1+hnda^1$
1PL.EXCL	$ko^3-n\tilde{a}\rho^1+hnda^{51}$	
2PL	$ko^3-n\tilde{a}\rho^1+hnda^1 o\rho^3$	
3PL	$ko^3-n\tilde{a}\rho^1+hnda^1 (ho^5)$	

shown in Table 30, with the verb tsi^5hnda^1 'buy', adapted from Buck (2000: 441). Stem formatives are segmented from the root and indicated by the plus sign.

The verb tsi^5hnda^1 'buy' has two stems: one used with singular subjects (tsi^5+hnda^1) and another with plural subjects ($n\tilde{a}\rho^1+hnda^1$). The singular stem bears the formative tsi^5+ and the plural $n\tilde{a}\rho^1+$. This formative pair served as the morphological means to produce causative verbs in Amuzgo. It is no longer productive, but the mechanism has left behind a large number of transitive verbs in the lexicon (185 in our corpus) together with a few active intransitives with agentive semantics (3 in our sample). For the singular, the formative tsi^5+ is reduced to $ts+$ before a stem with a vocalic onset or beginning with /k/, and sometimes before /b/. Examples are given in Table 31.

Table 31: Examples of causative verbs.

Singular	Plural	
$tsi^5+ty\rho ue^3$	$n\tilde{a}\rho^1+ty\rho ue^3$	'scare'
$tsi^5+k\rho ua^{35}$	$n\tilde{a}\rho^1+k\rho ua^{35}$	'make noise'
$tsi^5+nchu\rho^{51}$	$n\tilde{a}\rho^1+nchu\rho^{51}$	'fix'
$tsi^5+ts\tilde{a}\rho^3$	$n\tilde{a}\rho^1+ts\tilde{a}\rho^3$	'drive crazy'
$tsi^5+ba^3/ts+ba^{53}$	$n\tilde{a}\rho^1+ba^3$	'lift'
$ts+ia^{51}$	$n\tilde{a}\rho^1+ya^1$	'make'
$ts+k\tilde{a}^5$	$n\tilde{a}\rho^1+k\tilde{a}^5$	'dry'
$ts+ke\rho^{34}$	$n\tilde{a}\rho^1+ke\rho^{34}$	'hurt, wound'
$ts+kwa^3-nd\varepsilon\rho^{51}$	$n\tilde{a}\rho^1+kwa^3-nd\varepsilon\rho^{51}$	'turn around'

Causative verbs in Amuzgo are fossilized structures, but they originated from a light verb construction consisting of a series of two verbs. The morphs tsi^5+ and $n\tilde{a}\rho^1+$ are the reflexes of the two stems of the verb $tsia^{51}/n\tilde{a}\rho^1ya^1$ 'make'

used as a causative light verb. Besides, about 46 of the verbs of the $tsi^5+~n\ddot{a}\gamma^1$ type in S&S's sample still bear the morph ki^3 , which is doubtless the relic of the inflectional prefix ki^3 -, which is used for the subjunctive.¹⁶ We mark this by an underscore in Table 32.

Table 32: Examples of causative verbs with the old prefix ki^3 -.

Singular	Plural			Source intransitive verb
$tsi^5+ki^3_nde^35$	$n\ddot{a}\gamma^1+ki^3_nde^35$	'ripen'	$\sim nde^35$	'be ripe'
$tsi^5+ki^3_ndya^3$	$n\ddot{a}\gamma^1+ki^3_ndya^3$	'make spicy'	$\sim ndya^3$	'be spicy'
$tsi^5+ki^3_n\ddot{o}^3$	$n\ddot{a}\gamma^1+ki^3_n\ddot{o}^3$	'make run'	$\sim n\ddot{a}^3-n\ddot{o}^3$	'run off' (lit., 'exit-run')

Apart from causative verbs, there is a second type of verbs in Amuzgo that are lexicalized forms of old V_1V_2 compounds with the verb 'go' as V_1 . In the Surrey Morphology Group's database, there are 63 such verbs (42 from FT's dataset and 21 from S&S) but we have only included four of them in our sample (one from FT's data and three transitive verbs from S&S). The great majority of such verbs are intransitive, and most of them are motion or posture verbs, but at times it is difficult to identify semantic compositionality. The verbs in question could be characterized as consisting of two components. To illustrate this, consider the verb ba^5+kwa^3 'lie down' in Table 33.

Table 33: Three TAMs of the verb ba^5+kwa^3 'lie down'.

	Incomplete	Completive	Subjunctive
1SG	$\emptyset-ba^5+kwa^3$	$ty-\epsilon\gamma^1+kwa^3$	$k-o\gamma^1+kwa^3$
2SG	$\emptyset-ba^3+kwa^3-\gamma$	$t-(b)a^3+kwa^3-\gamma$	$k-(b)a^3+kwa^3-\gamma$
3SG	$\emptyset-ba^5+kwa^3$	$t-(b)a^5+kwa^3$	$k-(b)a^5+kwa^3$
1PL.INCL	$ko^3-tsa^5+ki^3_m\gamma a^34$	$\langle s \rangle a^5+ki^3_m\gamma a^34$	$ki^3-tsa^5+ki^3_m\gamma a^34$
1PL.EXCL	$ko^3-tsa^5+ki^3_m\gamma a^51$	$\langle s \rangle a^5+ki^3_m\gamma a^51$	$ki^3-tsa^5+ki^3_m\gamma a^51$
2PL	$ko^3-tsa^5+ki^3_ma\gamma^1 o\gamma^3$	$\langle s \rangle a^5+ki^3_ma\gamma^1 o\gamma^3$	$ki^3-tsa^5+ki^3_ma\gamma^1 o\gamma^3$
3PL	$ko^3-tsa^5+ki^3_ma\gamma^1 (ho^5)$	$\langle s \rangle a^5+ki^3_ma\gamma^1 (ho^5)$	$ki^3-tsa^5+ki^3_ma\gamma^1 (ho^5)$

The second component in such formations is the lexical element. In the example in Table 33, this lexical element already has a singular–plural split at the stem

¹⁶ The use of this prefix is shown in Table 33. It has allomorphy, e.g., notice the allomorph kw - in Table 5 above. In the complementation constructions of Oto-Manguean languages, which generally lack non-finite forms such as infinitives, it is very common to find the dependent clause marked with a special subordinating mood such as subjunctive, potential, or irrealis.

level: we have kwa^3 for the singular, and $ki^3_ma\gamma^1$ for the plural.¹⁷ The first component of the verb, meanwhile, has forms that are remnants of inflected forms of the verb bha^5 ‘go’. To see this, compare the inflection of the first component of the verb ba^5+kwa^3 ‘lie down’ in Table 33 with the forms of the verb bha^5 ‘go’ in Table 34 from equivalent sub-paradigms.

Table 34: Three TAMs of the verb bha^5 ‘go’.

	Incomplete	Completive	Subjunctive
1SG	$\emptyset-h\text{ɔ}^1$	$ty-\text{ɛ}\text{ɛ}^1$	$k-h\text{ɔ}^1$
2SG	$\emptyset-b_ha^{53-\gamma}$	$t-(b_ha)^{53-\gamma}$	$k-(b_ha)^{53-\gamma}$
3SG	$\emptyset-b_ha^5$	$t-(b_ha)^5$	$k-(b_ha)^5$
1PL.INCL	$ko^3-\gamma\text{ɔ}^5$	$ty-\gamma\text{ɛ}\text{ɛ}^5$	$k-\gamma\text{ɔ}^5$
1PL.EXCL	ko^3-tsa^{51}	$\text{ɛ}sa^{51}$	ki^3-tsa^{51}
2PL	$ko^3-\gamma\text{o}^5=\text{o}\gamma^3$	$ty-\gamma\text{e}^5=\text{o}\gamma^3$	$k-\gamma\text{o}^5=\text{o}\gamma^3$
3PL	$ko^3-\gamma\text{o}^5(=ho^5)$	$ty-\gamma\text{e}^5(=ho^5)$	$k-\gamma\text{o}^5(=ho^5)$

The first component of the plural stem of the verb in Table 33 has tsa^5+ , which we take to function synchronically as a stem formative. This morph originated from the plural stem of the verb ‘go’ that we still observe today in the form for the 1st person plural exclusive. For the singular, the forms remind us of equivalent inflected forms of the verb ‘go’; they show an almost identical suppletion. But there are slight differences that justify their treatment as independent forms. Such formal differences together with the overall lack of semantic compositionality justifies treating verbs of this type as verbs that have idiosyncratic inflectional properties rather than just verbal compounds based on ‘go’. All such verbs manifest the singular–plural split. This is achieved for all of them by means of the formal contrasts involving the first component, but for some of the verbs, for example the one in Table 33, the second component may also manifest the split independently. For such cases, we have double marking. This brings us to the next section where we show more instances of double marking at stemlevel.

¹⁷ For more about this stem see Note (b) in Table 36. In turn, the plural stem $ki^3_ma\gamma^1$ is the lexicalization of an inflected form bearing the subjunctive prefix ki^3 - plus the genuine old plural stem $ma\gamma^1$. A similar structure is found in many causative verbs, as illustrated in Table 32, but at least for this verb and others like it (e.g., $ba^5+kan^1/tsa^5+ki^3_t-(k)an^3$ ‘beg’) the subjunctive structure is only circumscribed to the plural, which can be taken as a further indication of the great depth of the split.

5.7 The plural stem combines different marking strategies

There are also verbs whose stems combine two of the strategies we have presented so far. The verbs that are found in our sample that behave in this way are given in Table 35. The existence of verbs like the ones in Table 35 can be taken as further evidence that the operations behind the building of plural stems have ceased to be productive, and are only connected to the lexicon as instances of partial suppletion. The last verb is an example of how the stem formative strategy can combine with root suppletion. This leads us to the next section where we present instances of suppletive stems which are more canonical.

Table 35: Verbs with exuberant plural marking.

Singular	Plural		
<i>ndi⁵+tshɔ³⁴</i>	<i>n- + t-</i>	<i>ndi⁵+n-t-hɔ³⁴</i>	'be slow', 'be docile'
<i>tsʔa³⁵</i>	<i>n- + t-</i>	<i>n-t-ʔa¹</i>	'do'
<i>skuʔ¹²</i>	<i>n- + t-</i>	<i>n-t-ku¹²</i>	'have a wife'
<i>ndi⁵+sku¹</i>	<i>n- + t-</i>	<i>ndi⁵+n-t-ku¹</i>	'be beautiful'
<i>ndi⁵+ki⁵</i>	<i>n- + -e</i>	<i>ndi⁵+n-k-e⁵</i>	'harden'
<i>ba³nkheʔ³⁵</i>	<i>t- + ʔ-</i>	<i>t-(b)a¹ʔnkheʔ³⁵</i>	'pick up, choose'
<i>tsj⁵+ki⁵+tsʔɔ³</i>	<i>nãʔ¹ + n-</i>	<i>nãʔ¹+ki⁵+n-ʔo³</i>	'bear pain'
<i>tsj⁵+ki⁵_ndy-i⁵</i>	<i>nãʔ¹ + -e</i>	<i>nãʔ¹+ki⁵_ndy-e⁵</i>	'make drunk'
<i>ba⁵+kwa³-ngio³</i>	<i>tša⁵ + Suppletion</i>	<i>tša⁵+ki³_maʔ¹-ngio³</i>	'lie face down'

5.8 Suppletive stems

Verbs with more canonical suppletive stems also occur in SP Amuzgo. The difference between the singular and plural stems can be one of total suppletion, like in Table 36(a) where the stems lack any element in common, suggesting they probably come from different lexical sources. We also have cases of partial suppletion shown in Table 36(b), where the plural stem is related to the singular stem, but only in obscure ways.

5.9 Stems contrasting in tone

More than a hundred verbs in our sample have a singular stem that contrasts in tone with the plural stem. For the instances in Table 37, the change in tone is

Table 36: Verbs with suppletive stems.

	Singular	Plural	
a.	<i>ki³-ndyi³⁴</i>	<i>ki⁵-to⁵</i>	'be included'
	<i>ñe^ɾ+ki⁵_tso⁵</i>	<i>ñe^ɾ+ki³_nda¹</i>	'be tired'
	<i>tsku^ɾ</i>	<i>tɛ^ɾ</i>	'close', 'prohibit'
	<i>tʔeo³⁵</i>	<i>tyue³⁵</i>	'cry'
	<i>ncchi³⁵</i>	<i>ngio¹</i>	'know', 'find out' ^a
	<i>nthi³⁵</i>	<i>ndio¹</i>	'find'
	<i>bʔi³⁵</i>	<i>bhe³⁵</i>	'get angry'
	<i>ndyo¹</i>	<i>nkio³</i>	'come'
	<i>ndi⁵+hi⁵³-ba³⁵</i>	<i>ndi⁵+hi⁵³-mi³</i>	'get bad, ugly'
b.	<i>nchkē^ɾ</i>	<i>ntykē¹²</i>	'grow weak', 'get bored'
	<i>tɔa^ɾ</i>	<i>nda^ɾ</i>	'limp'
	<i>ñe^ɾ+k_ue^ɾ</i>	<i>ñe^ɾ+k_whe³</i>	'want to die'
	<i>ñe^ɾ+k_ʔu⁵</i>	<i>ñe^ɾ+k_we¹</i>	'want to drink'
	<i>hnda³</i>	<i>nda³</i>	'have an adopted child'
	<i>ndi⁵+hndyo³</i>	<i>ndi⁵+ndyo³</i>	'grow old (man)'
	<i>ba⁵+kwa³</i>	<i>tɔa⁵+ki³_ma^ɾ</i>	'lie down' ^b

a. This verb occurs in seven compound verbs of emotion in our sample: *bi^ɾ-nchhi³⁵/bi^ɾ-ngio¹* 'love'; *hnde⁵¹-ya¹-nchhi³⁵/hnde⁵¹.ya¹-ngio¹* 'be content'; *ki³ndya^ɾ-nchhi³⁵/ki³ndya^ɾ-ngio¹* 'be upset'; *na³nko^{ɾ34}-nchhi³⁵/na³nko^{ɾ34}-ngio¹* 'be happy'; *ndi³+ʔnda^ɾ-nchhi³⁵/ndi³+ʔnda^ɾ-ngio¹* 'miss'; *tyia¹-nchhi³⁵/tyia¹-ngio¹* 'become sad'; *tyiō³ na³ki³ndya^ɾ-nchhi³⁵/tyiō³ na³ki³ndya^ɾ-ngio¹* 'get upset'.

b. For the verb *ba⁵+kwa³* 'lie down' one can still see traces of how the suppletion has emerged. The plural stem *ki³_ma^ɾ* surfaces alternatively as *kwa^ɾ*. In compounds, e.g., *ba⁵+kwa³-ngio³* 'lie face down' with plural stem *tɔa⁵+kwa^ɾ-ngio³*. Unstressed /ki/ becomes /k/ and /ma^ɾ/ labializes to /wa^ɾ/. One could interpret the component *kwa³* of the singular stem as originating equally from **ki³_ma³*. In such a case, the stems would only contrast in glottalization and tone.

Table 37: Verbs with predictable tone in the singular stem.

Singular	Plural	
<i>ts+ia⁵¹</i>	<i>nã^ɾ+ya¹</i>	'make'
<i>ts+kwe^ɾ⁵¹ hndyo¹²</i>	<i>nã^ɾ+kwe^ɾ¹ hndyoo¹²</i>	'change'
<i>ts+ba⁵³</i>	<i>nã^ɾ+ba³</i>	'lift up'
<i>ts+be⁵³</i>	<i>nã^ɾ+be³</i>	'keep'

predictable if we consider the tone of the root in the plural stem as lexical and the tone of the singular as emerging from the conflation of the lexical tone /5/ of the reduced formative *tsi*⁵+ with the root tone.

Some other verbs make the contrast between a singular and a plural stem by means of tone only. However, there is no simple form-function mapping. In other words, it is not possible to assign a function to a given tone value, e.g., tone *X* is a marker of function *Y*. There are a number of tonal patterns involved. The variance is better understood as organized into tone classes dictated by the lexicon. Examples of the tone classes we have identified are given in Table 38 with an indication as to the number of verbs instantiating each class. The tone classes are not labeled with numbers or letters, but by the tone pairs found in the singular and the plural, e.g., tone class /1~3/ means that the stem has tone /1/ in the singular, but tone /3/ in the plural.¹⁸

Table 38: Tone classes in verbs whose stems only contrast in tone.

Class	Singular	Plural		Number of verbs
/1~3/	<i>nde</i> ¹	<i>nde</i> ³	'copulate'	8
	<i>ndui</i> ¹	<i>ndui</i> ³	'go out'	
	<i>nā</i> ¹	<i>nā</i> ³	'remove seed (cotton)', 'remove'	
/3~1/	<i>nko</i> ³	<i>nko</i> ¹	'laugh'	5
	<i>nā</i> ³	<i>nā</i> ¹	'be patient'	
	<i>ñe</i> ¹ <i>__kwa</i> ³	<i>ñe</i> ¹ <i>__kwa</i> ¹	'want to eat'	
/35~1/	<i>nde</i> ³⁵	<i>nde</i> ¹	'chew chewing gum/ sugar cane'	4
	<i>hñō</i> ³⁵	<i>hñō</i> ¹	'plaster a wall'	
	<i>ñe</i> ¹ <i>__k-γō</i> ³⁵	<i>ñe</i> ¹ <i>__k-γō</i> ¹	'want to exist'	
/3~12/	<i>(ki</i> ³ <i>-)hndyu</i> ³	<i>(ki</i> ³ <i>-)hndyu</i> ¹²	'appoint oneself'	2
	<i>ñe</i> ¹ <i>__ki</i> ³ <i>_n</i> ³ <i>o</i> ^{a3}	<i>ñe</i> ¹ <i>__ki</i> ³ <i>_n</i> ³ <i>o</i> ^{a12}	'want to heal'	
/34~12/	<i>ntyha</i> ³⁴	<i>ntyha</i> ¹²	'pull'	2
	<i>ntshō</i> ³⁴	<i>ntshō</i> ¹²	'sow again (new grain)'	
/5~1/	<i>kwe</i> ⁵	<i>kwe</i> ¹	'dress'	2
	<i>nkyā</i> ⁵	<i>nkyā</i> ¹	'give', 'hand over'	
/5~12/	<i>ñe</i> ¹ <i>__kwe</i> ⁵	<i>ñe</i> ¹ <i>__kwe</i> ¹²	'want to get dressed'	1
/53~12/	<i>ka</i> ⁵³	<i>k</i> ³ <i>a</i> ¹²	'run over', 'flood'	1

¹⁸ There are two missing verbs that display a tone contrast only but are disyllabic: *ndyo*^{21-γ}*ō*³⁵ 'bring' (SG.OBJ) and *be*⁵-*ntyha*³ 'reach' given in Table 41. The latter comes from S&S and we ignore the status of its glottalization component.

Table 39: Correlation between tone classes and segmental strategies.

Tone class	Tone only	Glottal	Final vowel	Prefix <i>t-</i>	Prefix <i>n-</i>	Formative	Suppletion	Prefix <i>t-</i> & glottal	Formative & suppletive	Total
/1~3/	8			25		1	2			36
/3~1/	5	16		4	1	2		1	1	30
/35~1/	4		5	1	1	1	2			14
/3~12/	2			7						9
/35~12/		1	1	5	2					9
/34~12/	2			3						5
/5~3/			2	3						5
/5~1/	2		1	1			1			5
/53~1/				3						3
/5~12/	1						1			2
/53~12/	1			1						2
/12~34/				1						1
/34~5/							1			1
/35~3/							2			2
/53~34/						1				1
Total	25	17	9	54	4	5	9	1	1	125

All the tone classes in Table 38, except /5~12/ and /53~12/, are also attested in other verbs that present additional segmental means to exhibit the split. The relevant figures are given in Table 39, where we correlate tone class with segmental strategy. The figures reveal certain tendencies. For example, when a verb is specified in the lexicon as being one that changes tone, if its singular stem has tone /1/, its plural stem will have tone /3/. Beyond this the degree of predictability weakens. Similarly, half of the plural stems with a prefix *t-* have a tone /3/, the other half have either tone /12/ or tone /1/. Also, 90% of the verbs that have a glottalization contrast between the singular and the plural stems belong to tone class /3~1/.

When a verb changes tone, and the verb in question has more than one syllable (commonly representing a compound, whether opaque or transparent), the default is for the lexical stem to bear the tone change. This is shown in Table 40 (a and b) with instances of tone class /3~1/. As pointed out in Section 5.6 above, example (b) in Table 40 comes from a complementation structure involving the verb ‘go’. A similar structure may be observed in example (c) where the first morph $\tilde{n}e^{2l}$ is a desiderative.¹⁹

A few disyllabic stems in our sample, given in Table 41, undergo changes in tone involving both syllables of the lexical stem. The first two examples are

¹⁹ In contrast, the modern desiderative verb is the compound active verb $k\tilde{e}^3-ts\tilde{2}\delta^3$ ‘want, desire’.

Table 40: Lexical stem with tone changes in complex verbs.

	Singular	Plural	
a.	{ <i>ndo</i> ³ - <i>nde</i> ³⁴ }	{ <i>ndo</i> ¹ - <i>nde</i> ³⁴ }	‘look up’
b.	<i>ba</i> ⁵ +{ <i>kwa</i> ³ - <i>ngio</i> ³ }	<i>tša</i> ⁵ + <i>ki</i> ³ -{ <i>ma</i> ⁷ - <i>ngio</i> ³ }	‘lie face down’
c.	<i>ñe</i> ⁷ -{ <i>kwa</i> ⁷ }	<i>ñe</i> ⁷ -{ <i>kwa</i> ⁷ }	‘feel like eating’

Table 41: The tonal marking of disyllabic stems.

σ_1	σ_2	Singular	Plural	
/1~3/	/35~1/	<i>ndyo</i> ⁷ - <i>y</i> ⁷ <i>õ</i> ³⁵	<i>ndyo</i> ⁷ - <i>y</i> ⁷ <i>õ</i> ¹	‘bring’ (SG.OBJ) (Lit. ‘come-take’ SG.OBJ)
/1~3/	/5~1/	<i>ndyo</i> ¹ - <i>ch-u</i> ⁵	<i>ndyo</i> ³ - <i>ch-o</i> ¹	‘bring’ (PL.OBJ) (Lit. ‘come-take’ PL.OBJ)
/3~1/	/12~1/	<i>ko</i> ³ - <i>khõ</i> ¹²	<i>t-(k)o</i> ⁷ - <i>hõ</i> ¹	‘send’
/1~3/	/51~3/	<i>ko</i> ¹ - <i>χe</i> ⁵¹	<i>t-(k)o</i> ³ - <i>χe</i> ³	‘weigh, govern’
/5~1/	/3~1/	<i>be</i> ⁵ - <i>ntyha</i> ³	<i>be</i> ¹ - <i>ntyha</i> ¹	‘reach’
		<i>ba</i> ⁵ + <i>be</i> ⁵ - <i>ntyha</i> ³	<i>tša</i> ⁵ + <i>be</i> ¹ - <i>ntyha</i> ¹	‘go and reach’
/5~3/	/3~1/	<i>ba</i> ⁵ + <i>ki</i> ⁵ - <i>ṽmã</i> ³	<i>tša</i> ⁵ + <i>ki</i> ³ - <i>ṽmã</i> ¹	‘go and call’
/5~3/	/5~1/	<i>ñe</i> ⁷ - <i>ki</i> ⁵ - <i>tso</i> ⁵	<i>ñe</i> ⁷ - <i>ki</i> ³ - <i>nda</i> ¹	‘be sleepy’

transparent compounds where each stem is subject to tone changes. In light of these two examples, we also take the other instances as reflecting similar but older compounds.

5.10 Paradigmatic depth of the number split

In the previous sections we have shown that the number split in Amuzgo cleaves the paradigm in two, dividing up the cells associated with singular subjects from those associated with plural subjects. We have also presented the formal mechanisms that make it possible for a verb to have the two distinct stems needed for the number split. Splits involving stems are informative about the way a paradigm is structured. When a consistent type of split applies to the paradigm of many verbs, we witness the occurrence of a characteristic property of inflection in a given system. After observing the behavior of verbs in our sample we conclude that the partition of the paradigm characterizing the number split is also a default way of partitioning a verbal paradigm in Amuzgo. To our knowledge, the verb *ṽna*³⁴ ‘take over’ in Table 42 is the only one in our sample that strays from such partitioning, displaying instead a split in person between the forms for the 1st person and all other persons.

Table 42: Split involving 1st person.

INCPL	'take over'		
1SG	∅-	<i>nha</i> ³	
2SG	∅-	<i>ɔna</i> ⁵¹	-ɔ
3SG	∅-	<i>ɔna</i> ³⁴	
1PL.INCL	∅-	<i>nha</i> ³⁵	⟨ā⟩ ⁵
1PL.EXCL	∅-	<i>nha</i> ⁵¹	
2PL	∅-	<i>ɔna</i> ³⁴	=oɔ ³
3PL	∅-	<i>ɔna</i> ³⁴	=ho ⁵

In contrast, there are a few verbs that have more than two stems, but the third stem occurs in the singular of 1st and 2nd persons, as shown in Table 43, adapted from Buck (2000: 450, 451). This pattern is not found with verbs that lack a stem-based split, so we could argue that the split involving 1st and 2nd person singular is nested within the number split (for nested splits see Corbett 2015). The nested split responds to the same person syncretism we observe for the affix *ma*³- used for the incompletive of the 1st and the 2nd person singular of Class 1 verbs.

Table 43: Nested split involving 1st and 2nd person singular.

INCPL	'drink'			'die'		
1SG	<i>ma</i> ³ -	<i>ɔwa</i> ¹		<i>ma</i> ³ -	<i>bɔi</i> ⟨ɔ⟩ ¹	
2SG	<i>ma</i> ³ -	<i>ɔwa</i> ⁵¹	-ɔ	<i>ma</i> ³ -	<i>bɔio</i> ³	-ɔ
3SG	<i>ɔ</i> ⁵ -	<i>ɔu</i> ⁵		<i>ɔ</i> ⁵ -	<i>be</i> ɔ ³	
1PL.INCL	<i>ko</i> ³ -	<i>b</i> ⟨ε⟩ ¹²	⟨ε⟩ ⁵	<i>ko</i> ³ -	<i>bh</i> ⟨ε⟩ ³⁴	⟨ε⟩ ⁵
1PL.EXCL	<i>ko</i> ³ -	<i>b</i> ⟨ε⟩ ⁵¹		<i>ko</i> ³ -	<i>bh</i> ⟨ε⟩ ⁵¹	
2PL	<i>ko</i> ³ -	<i>be</i> ¹	=oɔ ³	<i>ko</i> ³ -	<i>bhe</i> ³	=oɔ ³
3PL	<i>ko</i> ³ -	<i>be</i> ¹	=ho ⁵	<i>ko</i> ³ -	<i>bhe</i> ³	=ho ⁵

One verb in our sample has the basic number split and, in addition to this, displays three different stems for the persons in the singular. The verb is shown in Table 44.

Finally, there is one verb with a fourth stem, but our expectation is that we would not find others like it in the language because it seems to push irregularity to its extreme. It is the verb *bha*⁵ 'go' in Table 45, which also happens to be a lexical item prone to inflectional idiosyncrasies crosslinguistically. Here the 1st person plural exclusive receives a distinct stem. This verb is also peculiar in having a pattern for the two stems in the singular that differs from the splits in

Table 44: Nested split involving all singular persons.

INCP		'get drunk'	
1SG	∅-	<i>ki⁵ndy<ɔ>³</i>	
2SG	∅-	<i>ki⁵ndyu³</i>	-ʔ
3SG	∅-	<i>ki⁵ndy⁵</i>	
1PL.INCL	∅-	<i>ki⁵ndy<ɛ>⁵</i>	<ɛ> ⁵
1PL.EXCL	∅-	<i>ki⁵ndye⁵¹</i>	
2PL	∅-	<i>ki⁵ndye⁵</i>	=oʔ ³
3PL	∅-	<i>ki⁵ndye⁵</i>	=ho ⁵

Table 45: The four-way split of the verb *bha⁵* 'go'.

INCP		'go'	
1SG	∅-	<i>h<ɔ>¹</i>	
2SG	∅-	<i>bha⁵³</i>	-ʔ
3SG	∅-	<i>bha⁵</i>	
1PL.INCL	<i>ko³</i> -	<i>ʔ<ɔ>⁵</i>	
1PL.EXCL	<i>ko³</i> -	<i>tʂa⁵¹</i>	
2PL	<i>ko³</i> -	<i>ʔo⁵</i>	=oʔ ³
3PL	<i>ko³</i> -	<i>ʔo⁵</i>	=ho ⁵

Tables 43 and 44, but is identical to the singular portion of the paradigm in Table 42. We take this as further evidence for a greater degree of independence of the singular half of the paradigm compared with the plural half.

All these structures point to the existence of a deeply entrenched split in the inflection of Amuzgo verbs. In Table 46 we summarise the distinct types of stem-level splits which we have identified in SP Amuzgo (together with an indication of the number of verbs following each type). Firstly, there are verbs which display no split at stem level; the notation "A" in Table 46 indicates the invariant stem of such verbs. Secondly, in verbs which do display a split, the first default partition is between the singular and plural forms; the notation "B" indicates the occurrence of a second stem that contrasts with the basic A-stem. Although the relation between singular and plural stems could be characterized synchronically as instantiating a form of suppletion, we believe that treating the plural stem as secondary is justified given the nature of the formal relation between them. We call this "split 1a". Once this default split happens, the

Table 46: Different types of stem-level splits in SP Amuzgo verbs.

	No split	Split 1a {Default}	Split 1b {Advanced}	Split 1c {Rare}	Split 1d {Unique}	Split 2	Unattested
1SG	A	A	C	D	D	B	B
2SG	A	A	C	C	A	A	B
3SG	A	A	A	A	A	A	A
1PL.INCL	A	B	B	B	B	B	A
1PL.EXCL	A	B	B	B	C	B	A
2PL	A	B	B	B	B	A	A
3PL	A	B	B	B	B	A	A
Example verb:	'smash'	'lie down'	'drink'	'get drunk'	'go'	'take over'	—
See Table:	10	1	43	44	45	42	—

paradigm can then be further partitioned with nested splits, like split 1b, which is relatively common, where “C” represents a third stem. Split 1c reveals yet another way of splitting the singular portion of the paradigm. Finally, we consider splits 1d and 2 as unique. Other ways of splitting the paradigm are not found; we give an example of an unattested split in the final column of Table 46.

6 Discussion: Making sense of the number split in Amuzgo

We have shown that verbs in Amuzgo split their paradigm in two, partitioning the forms taking a plural subject from those taking a singular subject. We have also seen that, although the split is not found in all verbs, in our sample of 579 verbs it nonetheless affects 393 of them. Of those, at least 91% (360) manifest the split at stem level. How can we make sense of this?

The crucial point is how the relation between grammatical number and grammatical person is encoded. The issue is an old one: for example, in 1902 Beach Bradley published an essay on the treatment of *we* as the plural of *I*. Fortunately, our general knowledge of verbal inflection crosslinguistically has increased dramatically since then, and we know that the phenomenon is intricate from a typological perspective. Daniel (2013), for example, surveys different encoding possibilities regarding personal pronouns in the *World atlas of language structures*. Our interest here is in verbal indexing.

To understand the significance of the system in Amuzgo, we can start with a situation where there is a maximal disassociation between the marking of the person of the subject and the marking of the number of the subject. This situation is known to exist across languages, but it appears to be particularly common in languages of the Americas, especially in Mesoamerica. The canonical example for such a situation is given in Table 47 from Toluca Otomi, an Oto-Manguean language of Mexico distantly related to Amuzgo, where there is only one marker of plurality, the enclitic =hu, adapted from Palancar (2013).

Table 47: Person and number of subject in Toluca Otomi.

INCLP.L.REALIS	'walk'			
	Person	TAM	Stem	Number
1SG	<i>d-</i>	<i>rā=</i>	<i>'yo</i>	–
2SG	<i>g-</i>	<i>rā=</i>	<i>'yo</i>	–
3SG	<i>∅-</i>	<i>ra=</i>	<i>'yo</i>	–
1PL	<i>d-</i>	<i>rā=</i>	<i>'yo</i>	= <u>hu</u>
2PL	<i>g-</i>	<i>rā=</i>	<i>'yo</i>	= <u>hu</u>
3PL(HUM.DEF)	<i>∅-</i>	<i>ra=</i>	<i>'yo</i>	= <u>hu</u>

The crucial thing about the system in Toluca Otomi is that the marking for the person of the subject remains completely independent from the expression of the plurality of the individuals or entities involved (the less marked option is interpreted as singular by default), and as a consequence person markers convey just person, and number markers just number.

In other systems, we witness how specific markers for a given plural person can co-occur with the invariant marking of number. This is observed in the different subparadigms in Table 48 from Tsel'tal, a Mayan language of Mexico, adapted from Polian (2013). The plurality of the person of the subject is indicated in Tsel'tal by the suffix *-ik*. In this, it is the same type of system that we observe in Toluca Otomi in Table 47, but in Tsel'tal, for the marking of 1st person plural, we have the specific marker *-ot* for the completive of an intransitive verb, and the suffix *-t* for the incompletive of a transitive verb.²⁰ This is proof that, at least in this specific case, the language is treating 1st person plural as a class.

²⁰ One referee suggested to us that *-ot/-t* could equally be interpreted as operating within the realm of number. Our position is that this morphology is conflating person (1st person) and number (plural), signalling an emergent class of '1st person plural'.

Table 48: Person and number of subject in Tselal.

CPL	'open' (INTR)			INCPL	'grab' (TR)			
	Stem	Person/S	Number		AUX	Person/A	Stem	Person
1SG	<i>k'ajin</i>	<i>-on</i>	–	1SG>3	<i>ya</i>	<i>j-</i>	<i>tsak</i>	–
2SG	<i>k'ajin</i>	<i>-at</i>	–	2SG>3	<i>ya</i>	<i>a-</i>	<i>tsak</i>	–
3SG	<i>k'ajin</i>	<i>-∅</i>	–	3SG>3	<i>ya</i>	<i>s-</i>	<i>tsak</i>	–
1PL	<i>k'ajin</i>	<i>-ot</i>	<i>-ik</i>	1PL>3	<i>ya</i>	<i>j-</i>	<i>tsak</i>	<i>-t</i>
2PL	<i>k'ajin</i>	<i>-at</i>	<i>-ik</i>	2PL>3	<i>ya</i>	<i>a-</i>	<i>tsak</i>	<i>-ik</i>
3PL	<i>k'ajin</i>	<i>-∅</i>	<i>-ik</i>	3PL>3	<i>ya</i>	<i>s-</i>	<i>tsak</i>	<i>-ik</i>

A related system is also found in Turkish as shown in Table 49, from Göksel & Kerslake (2005). Here we observe a different marker for plural associated to person: *-uz* is for speech-act participants, but *-lar* is for the 3rd person, and is also the marker used for plurals of nouns. Notice that in Turkish, 1st person plural is also treated differently from the singular, this time by elision of the person marker.

Table 49: Person and number of subject in Turkish.

PRS.CONT	'open'			
	Stem	TAM	Person	Number
1SG	<i>aç</i>	<i>-iyor</i>	<i>-um</i>	–
2SG	<i>aç</i>	<i>-iyor</i>	<i>-sun</i>	–
3SG	<i>aç</i>	<i>-iyor</i>	<i>-∅</i>	–
1PL	<i>aç</i>	<i>-iyor</i>	–	<i>-uz</i>
2PL	<i>aç</i>	<i>-iyor</i>	<i>-sun</i>	<i>-uz</i>
3PL	<i>aç</i>	<i>-iyor</i>	<i>-∅</i>	<i>-lar</i>

In systems like these, the marker for number is canonically agglutinative (Plank 1999). Specifically, it has little allomorphy, and where allomorphy does occur it is conditioned by the phonology (i.e., it is normally just one invariant affix or a clitic); its application commonly responds to a transparent morpho-syntactic rule (i.e., when the syntax requires it, speakers use it with all verbs); it has uncomplicated morphotactics; and syntactically it occurs at the outer periphery of the stem.²¹ As for scope, if the language allows for head-marked

²¹ An apparent counterexample to this is found in Piaroa, a South American language of the Saliban family, as shown in (i). Data from Mosonyi (2000), adapted from Rosés Labrada (2015) and Jorge E. Rosés (personal communication). The rarity of this system may be related to its

Table 50: Person and number of object in Tselal.

CPL	'help' (TR)				
	AUX	Person/A	Stem	Person/O	Number
3>1SG	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-on</i>	–
3>2SG	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-at</i>	–
3>3SG	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-∅</i>	–
3>1PL	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-ot</i>	<i>-ik</i>
3>2PL	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-at</i>	<i>-ik</i>
3>3PL	<i>la</i>	<i>s-</i>	<i>koltay</i>	<i>-∅</i>	<i>-ik</i>

indexing of objects, the same number markers may also be used to indicate the plural value of the object, although there may be specific collocation restrictions. This is illustrated in Table 50 in Tselal, from Polian (2013).

Amuzgo is similar to systems like these in also having the encoding of the number of the subject independent from that of the person of subject, but it differs from them in at least four main respects:

- (i) A lexeme may manifest a split in its paradigm principally by having two different stems specified, one for each portion of the split. The stems synchronically hold a suppletive relation to each other, i.e., they are provided by the lexicon and are learned independently.
- (ii) The two stems are used as indexes for the number of the subject, i.e., S/A, and never for the number of the object.
- (iii) Not all verbs have this split paradigm, although the majority do. While this suggests that the default for a verb is to belong to the class that has the split, speakers still need to remember which verbs do not have the split.

origin as Rosés Labrada (2015) shows that the person suffixes have emerged from a grammaticalized inflected copula.

(i)	PRS	'work'		
		Stem	Number	Person
	1SG.M	<i>aditi</i>	–	<i>-sæ</i>
	2SG.M	<i>aditi</i>	–	<i>-hæ</i>
	3SG.M	<i>aditi</i>	–	<i>-∅</i>
	1PL	<i>adit(i)</i>	<i>-æti</i>	<i>-sæ</i>
	2PL	<i>adit(i)</i>	<i>-æti</i>	<i>-hæ</i>
	3PL	<i>adit(i)</i>	<i>-æti</i>	<i>-∅</i>

- (iv) Besides the number split, the verbal inflection of Amuzgo has a fully fledged paradigm of person-number markers, i.e., the information about the person of the subject and the information about the number of such persons is conflated in the grammatical exponence.

Although property (iv) makes the number split seem redundant from a morphosyntactic point of view, stem choice is commonly taken as occurring at a deep level in the building of a given inflectional form in a paradigm (at least in Word and Paradigm models such as Paradigm Function Morphology; Stump 2001). This view has consequences for the way in which we understand the system behind the number split in Amuzgo in relation to the system in languages like Toluca Otomi or Tzeltal. It suggests that when producing a given inflectional form to describe a certain state of affairs, in the derivation one needs to attend first to the number of the subject of the verb to retrieve the right stem of that verbal lexeme. This appears to come at a stage that is necessarily prior to the selection of person marking, which is affixal and for the most part clitic-like.

On the other hand, we have seen in Section 4 that many verbs in Amuzgo manifest the number split at the level of TAM marking in the incomplete. Consider the incomplete subparadigm of the intransitive inactive verb *tyion*^{ʔ35} ‘exert one’s force’ in Table 51, where the grammatical information written in a continuous line applies to the prefixes. The subparadigm in Table 51 is split in two at the affixal level. The affixes in the plural do not convey information about the person of the subject. The person of the subject for this particular section of the paradigm is retrieved from the contrasts present in the pronominal set. The affix is cumulative in that it also conveys information about the incomplete aspect.

Table 51: Person and number of subject plus TAM in SP Amuzgo.

INCP		‘exert one’s force’		
			Stem	PRO
1SG		<i>ma</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>ha</i> ³
2SG	Person (1,2,3) & Number (SG) of S/A [& Aspect (INCP)]	<i>ma</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>ʔu</i> ³
3SG		<i>ʔ</i> ⁵ -	<i>tyion</i> ^{ʔ35}	= <i>e</i> ⁵
1PL.INCL		<i>ko</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>ha</i> ⁵ <i>a</i> ⁵
1PL.EXCL	Number (PL) of S/A [& Aspect (INCP)]	<i>ko</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>ha</i> ⁵ ¹
2PL		<i>ko</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>ʔo</i> ³
3PL		<i>ko</i> ³ -	<i>tyion</i> ^{ʔ35}	= <i>e</i> ³

Table 52: Person and number of subject plus TAM in Standard West Saxon.

	PRO	'bind'		
		Stem		
PRS.IND				
1SG	<i>iċ</i>	<i>bind</i>	<i>-e</i>	
2SG	<i>ðū</i>	<i>bind</i>	<i>-est</i>	Person (1,2,3) & Number (SG) of S/A [& Tense (PRS) & Mood (IND)]
3SG.M	<i>hē</i>	<i>bind</i>	<i>-eð</i>	
<hr/>				
1PL	<i>wē</i>	<i>bind</i>	<i>-að</i>	
2PL	<i>ġē</i>	<i>bind</i>	<i>-að</i>	Number (PL) of S/A [& Tense (PRS) & Mood (IND)]
3PL	<i>hīe</i>	<i>bind</i>	<i>-að</i>	
PST.IND				
1SG	<i>iċ</i>	<i>band</i>	<i>-∅</i>	
2SG	<i>ðū</i>	<i>bund</i>	<i>-e</i>	Person (1,2,3) & Number (SG) of S/A [& Tense (PST) & Mood (IND)]
3SG.M	<i>hē</i>	<i>band</i>	<i>-∅</i>	
<hr/>				
1PL	<i>wē</i>	<i>bund</i>	<i>-on</i>	
2PL	<i>ġē</i>	<i>bund</i>	<i>-on</i>	Number (PL) of S/A [& Tense (PST) & Mood (IND)]
3PL	<i>hīe</i>	<i>bund</i>	<i>-on</i>	
PRS.SUBJV				
1SG	<i>iċ</i>	<i>bind</i>	<i>-e</i>	
2SG	<i>ðū</i>	<i>bind</i>	<i>-e</i>	Number (SG) of S/A [Mood (SUBJV)]
3SG.M	<i>hē</i>	<i>bind</i>	<i>-e</i>	
<hr/>				
1PL	<i>wē</i>	<i>bind</i>	<i>-en</i>	
2PL	<i>ġē</i>	<i>bind</i>	<i>-en</i>	Number (PL) of S/A [Mood (SUBJV)]
3PL	<i>hīe</i>	<i>bind</i>	<i>-en</i>	
PST.SUBJV				
1SG	<i>iċ</i>	<i>bund</i>	<i>-e</i>	
2SG	<i>ðū</i>	<i>bund</i>	<i>-e</i>	Number (SG) of S/A [Mood (SUBJV)]
3SG.M	<i>hē</i>	<i>bund</i>	<i>-e</i>	
<hr/>				
1PL	<i>wē</i>	<i>bund</i>	<i>-en</i>	
2PL	<i>ġē</i>	<i>bund</i>	<i>-en</i>	Number (PL) of S/A [Mood (SUBJV)]
3PL	<i>hīe</i>	<i>bund</i>	<i>-en</i>	

A similar type of TAM-based split is found in Old English in the variety of Standard West Saxon. For this purpose, consider the different subparadigms of the verb *bindan* 'bind' in Table 52, adapted from Sweet (1886: 22). The grammatical information written in a continuous line applies to the suffixes (there are additional stem alternation patterns which are not discussed here). We can see in Table 52 that the pattern is similar to the Amuzgo case in Table 51, with the notable difference that suffixes in West Saxon do not convey information about person in the

Table 53: Person and number of subject plus TAM and stem-split in SP Amuzgo.

INCPL		'lie down'		
			Stem	PRO
1SG		<i>ma</i> ³ -	<i>me</i> ^ɔ ¹	= <i>ha</i> ³
2SG	Person (1,2,3) & Number (SG) of S/A [& Aspect (INCPL)]	<i>ma</i> ³ -	<i>me</i> ^ɔ ¹	= <i>ɔu</i> ³
3SG		<i>ɔ</i> ⁵ -	<i>me</i> ^ɔ ¹	= <i>ɛ</i> ⁵
1PL.INCL		<i>ko</i> ³ -	<i>toe</i> ^ɔ ³	= <i>ha</i> ⁵ <i>a</i> ⁵
1PL.EXCL	Number (PL) of S/A [& Aspect (INCPL)]	<i>ko</i> ³ -	<i>toe</i> ^ɔ ³	= <i>ha</i> ⁵ ¹
2PL		<i>ko</i> ³ -	<i>toe</i> ^ɔ ³	= <i>ɔ</i> ³
3PL		<i>ko</i> ³ -	<i>toe</i> ^ɔ ³	= <i>ɛ</i> ³

subjunctive mood. But despite similarities, what makes Amuzgo typologically remarkable is not only that the TAM-based split can co-occur with a stem-based split, as illustrated in Table 53, but that this co-occurrence is the DEFAULT case.

Finally, it is worth comparing Amuzgo to other Oto-Manguean languages to understand how it relates to other systems in this phylum. We have seen the system of Toluca Otomi in Table 47, but to the best of our knowledge the most similar number split to Amuzgo is found in distantly related Chichimec (Oto-Pamean),²² although there the phenomenon appears to be restricted to only a few verbs.

The verbal inflection of Chichimec abounds in stem alternation patterns. For instance, based on Angulo (1933), Palancar & Avelino (to appear) propose the existence of 14 such patterns. One of the patterns reveals a similar type of stem-based number split to that which we find in Amuzgo. This is shown in Table 54a with the verb *ndü* 'go'. In a sample of 168 verbs, only 18 behave like this (10%). The remaining verbs follow other patterns. One pattern is the absence of stem alternations (Table 54b) instantiated by 30 verbs, while the stem alternation pattern in Table 54c is found in 36 verbs which have a different stem for 3rd person plural. Also, verbs fall into different classes depending on the prefix sets they select. As it happens, the three verbs in Table 54 all belong to different prefix classes. Interestingly, like in Toluca Otomi the prefix set in Table 54(c) does not encode number of the subject. Number of subject – dual and plural (as well as clusivity) – is conveyed by means of suffixes. (Data adapted from Angulo 1933: 165, 174–175.)

²² Within Oto-Manguean, a related system is also found in Tlapanec. According to Suárez (1983), Tlapanec verbs fall into many tone classes according to their inflection, but for many such classes involving intransitive verbs all plural persons receive the same tone (the tone value in question may vary depending on class). However, as this happens with intransitive verbs only, the case at hand involves a plural S, not an A.

Table 54: Person and number of subject in Chichimec.

FUT	a. 'go'			b. 'laugh'			c. 'see'		
1SG	<i>tá-</i>	<i>ndü</i>	–	<i>ta-</i>	<i>tér</i>	–	<i>ga-</i>	<i>nú</i>	–
2SG	<i>kí-</i>	<i>ndü</i>	–	<i>sa-</i>	<i>tér</i>	–	<i>ki-</i>	<i>nú</i>	–
3SG	<i>gá-</i>	<i>ndü</i>	–	<i>ta-</i>	<i>tér</i>	–	<i>ga-</i>	<i>nú</i>	–
1DU.EXCL	<i>tí-</i>	<i>ndü</i>	<i>-m^p</i>	<i>ti-</i>	<i>tér</i>	<i>-m^p</i>	<i>ga-</i>	<i>nú</i>	<i>-m^p</i>
1DU.INCL	<i>tí-</i>	<i>ndü</i>	<i>-s</i>	<i>ti-</i>	<i>tér</i>	<i>-es</i>	<i>ga-</i>	<i>nú</i>	<i>-s</i>
2DU	<i>kí-</i>	<i>ndü</i>	<i>-s</i>	<i>sa-</i>	<i>tér</i>	<i>-es</i>	<i>ki-</i>	<i>nú</i>	<i>-s</i>
3DU	<i>gá-</i>	<i>ndü</i>	<i>-s</i>	<i>ta-</i>	<i>tér</i>	<i>-es</i>	<i>ga-</i>	<i>nú</i>	<i>-s</i>
1PL.EXCL	<i>gú-</i>	<i>ndür</i>	<i>-hũ</i>	<i>ti-</i>	<i>tér</i>	<i>-hũ</i>	<i>ga-</i>	<i>nú</i>	<i>-hũ</i>
1PL.INCL	<i>gú-</i>	<i>ndür</i>	<i>-ín</i>	<i>ti-</i>	<i>tér</i>	<i>-ín</i>	<i>ga-</i>	<i>nú</i>	<i>-n</i>
2PL	<i>kí-</i>	<i>ndür</i>	<i>-ín</i>	<i>sa-</i>	<i>tér</i>	<i>-ín</i>	<i>ki-</i>	<i>nú</i>	<i>-n</i>
3PL	<i>gá-</i>	<i>ndür</i>	–	<i>ta-</i>	<i>tér</i>	–	<i>ga-</i>	<i>nhú</i>	–

7 Conclusion

In this article, we have shown that many verbs in SP Amuzgo have an inflectional split that responds to the number of the subject. While the study has been focused on the language of San Pedro Amuzgos, the existing grammatical materials of Xochistlahuaca Amuzgo (Buck 2012) reveal a very similar inflectional system in all respects, both formally and distributionally.²³ Although a comparison of the two varieties to determine the scope of variation within the family must be left to future research, we can safely conclude that the number split is a trait of Amuzgo.

For the most part, the number split is manifested at stem level. This means that for the building of inflectional forms for the different persons of the subject, one stem is used for the persons of the singular, while another is used for the plural. We have also seen that the formal relation between the singular and plural stems suggests that the singular one is more basic. But the mechanisms deriving plural stems are no longer productive, a fact that, together with the high degree of lexical specificity (i.e., irregularity) that we observe, justifies treating the stems as being in a suppletive relation to each other.

²³ For example, for the cognate transitive verb 'cut' both languages have very similar forms: SP Amuzgo 3SG.INCL *tʰ-[ʔ]waʰ* vs. Xochistlahuaca Amuzgo 3SG.INCL *maʰ-kʰwaʰ* (with the prefix *maʰ-* for all singular persons in the incompletive); and SP Amuzgo 3PL.INCL *koʰ-t-ʔwaʰ* vs. Xochistlahuaca Amuzgo 3PL.INCL *kwiʰ-t-[k]ʔwaʰ*. There are many verbs like this displaying very similar behavior.

Suppletion in turn suggests that for a large portion of the verbal lexicon, speakers have to learn at least two independent stems (there are also verbs with more than two stems) so that the syntax can select either a singular or plural one. In this respect, Amuzgo is typologically remarkable. We know of other languages with verbal number (e.g., Ainu, Yaqui, Hiw, Samoan, etc.), but in those languages the phenomenon is very different: it is not part of the inflection; it involves a reduced number of lexical pairs; and the agreement pattern is with the number of the S and the O arguments. Similarly, while inflectional systems involving a paradigmatic contrast between two (or more) stems are also attested, they often realize aspectual distinctions (e.g., Greek, Farsi, etc.). To the best of our knowledge, a number split as it occurs in Amuzgo has not been reported to occur in any other language, or at least not to the extent that it is manifested in Amuzgo. Because of this, it deserves our attention.

The Oto-Manguean languages of Mexico have challenging morphologies. Amuzgo is no exception. In fact, when other aspects of its inflection are taken into account, such as TAM classes, glottalization classes, and tone classes, we can truly appreciate the extremely complex morphological system of this small language family within Oto-Manguean.

Our final words here should be an acknowledgment. Without the commitment of native linguist Fermín Tapia, it would have taken linguistics much longer to discover a unique language to expand our understanding of the treatment of number in the verbal domain. This is good news, because for a threatened language like San Pedro Amuzgo, it comes at the right time.

Acknowledgments: We are immensely grateful to Fermín Tapia for letting us use his database on San Pedro Amuzgo and to Yuni Kim for her generosity in sharing her revisions with us. We presented a first version of this article in Budapest in 2014 at the International Morphology Meeting and we are thankful for the comments received. We also want to thank Frans Plank as editor of *LT* for his help and to the three anonymous referees for their valuable time and for all their useful comments. All errors and deficiencies remain our own responsibility.

Abbreviations: 1/3 1st/3rd person; COMP = complementizer; CONT = continuous; CPL = completive; DEM = demonstrative; DET = determiner; DU = dual; EXCL = exclusive; FUT = future; HON = honorific; HUM = human; INCL = inclusive; INCLP = incomplete; IND = indicative; INTR = intransitive; IRR = irrealis; M = masculine; OBJ = object; PL = plural; PRF = perfect; PRG = progressive; PRO = pronoun; PRS = present; PST = past; RED = reduplication; SG = singular; SUBJ = subject; SUBJV = subjunctive; TR = transitive.

References

- Angulo, Jaime de. 1933. The Chichimeco language (central Mexico). *International Journal of American Linguistics* 7. 152–194.
- Apóstol Polanco, Jair. 2014. *Clases flexivas verbales en el amuzgo de Xochistlahuaca, Guerrero*. México, DF: Centro de Investigaciones y Estudios Superiores en Antropología Social MA thesis.
- Beach Bradley, Cornelius. 1902. “We” the plural of “I”? *Modern Language Notes* 17(1). 1–3.
- Buck, Marjorie J. 2000. Gramática amuzga de San Pedro Amuzgos, Oaxaca. In Stewart & Stewart 2000, 361–480.
- Buck, Marjorie J. 2012. Gramática del amuzgo de Xochistlahuaca. Manuscript.
- Cabredo Hofherr, Patricia. 2010. Verbal plurality and event plurality. Manuscript for the DGFS-CNRS Summer School on Linguistic Typology, Leipzig. https://www.eva.mpg.de/lingua/conference/2010_summerschool/pdf/course_materials/Cabredo2010plV.pdf (accessed 1 July 2015)
- Cabredo Hofherr, Patricia & Brenda Laca. 2012. Introduction: Event plurality, verbal plurality and distributivity. In Patricia Cabredo Hofherr & Brenda Laca (eds.), *Verbal plurality and distributivity*, 1–24. Berlin: De Gruyter.
- Caron, Bernard. 2012. *Hausa grammatical sketch*. Version 3. <https://halshs.archives-ouvertes.fr/halshs-00647533> (accessed 1 July 2015)
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2015. Morphosyntactic complexity: A typology of lexical splits. *Language* 91. 145–193.
- Daniel, Michael. 2013. Plurality in independent personal pronouns. In Matthew Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max-Planck-Institut für evolutionäre Anthropologie. <http://wals.info/chapter/35> (accessed 1 July 2015)
- Donohue, Mark & Søren Wichmann (eds.). 2008. *The typology of semantic alignment*. Oxford: Oxford University Press.
- Eulenberg, John B. 1971. Conjunction reduction and reduplication in African language. In Chin-Wu Kim & Herbert Stahlke (eds.), *Papers in African linguistics* (Current Inquiry into Language and Linguistics 1), 71–80. Edmonton: Linguistic Research.
- François, Alexandre. 2009. Verbal number and suppletion in Hiw. Paper read at the Eleventh International Conference on Austronesian Linguistics, Aussois, France.
- Gillon, Carrie & Robert Mailhammer. 2015. Reanalyzing the morphology and semantics of verbal plural marking in Quechan. *International Journal of American Linguistics* 81. 573–593.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. London: Routledge.
- Henderson, Robert. 2012. *Ways of pluralizing events*. Santa Cruz, CA: University of California at Santa Cruz doctoral dissertation.
- Jamieson, Carole Ann. 1982. Conflated subsystems marking person and aspect in Chiquihuitlán Mazatec verbs. *International Journal of American Linguistics* 48. 139–167.
- Kim, Yuni. In press. Tonal overwriting and inflectional exponence in Amuzgo. In Enrique L. Palancar & Jean Léo Léonard (eds.), *Tone and inflection: New facts and new perspectives*. Berlin: De Gruyter Mouton.
- Kim, Yuni. To appear. A morphophonological parameter hierarchy for Amuzgo glottalization classes. In Palancar et al. (eds.) (to appear).

- Maiden, Martin. 2005. Morphological autonomy and diachrony. In Geert E. Booij & Jaap van Marle (eds.), *Yearbook of morphology 2004*, 137–175. Dordrecht: Springer.
- Mithun, Marianne. 1991. Active/Agentive case marking and its motivations. *Language* 67. 510–546.
- Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.
- Mosonyi, Esteban Emilio. 2000. Elementos gramaticales del idioma piaroa. In María Stella González de Pérez & María Luisa Rodríguez (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 657–668. Bogotá: Instituto Caro y Cuervo.
- Newman, Paul. 1980. *The classification of Chadic within Afroasiatic*. Leiden: Leiden University Press.
- Palancar, Enrique L. 2013. The evolution of number in Otomi: The many faces of the dual. *Studies in Language* 37. 94–142.
- Palancar, Enrique L. & Heriberto Avelino. To appear. Inflectional complexity and verb classes in Chichimec. In Palancar et al. (eds.) (to appear).
- Palancar, Enrique L., Matthew Baerman & Timothy Feist (eds.). To appear. *Inflectional complexity and verb classes: A view from the Oto-Manguean languages of Mexico*.
- Plank, Frans. 1999. Split morphology: How agglutination and flexion mix. *Linguistic Typology* 3. 279–340.
- Polian, Gilles. 2013. *Gramática del tselta de Oxchuk*. México, DF: Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Rosés Labrada, Jorge E. 2015. *The Mako language: Vitality, grammar and classification*. London, Ont.: University of Western Ontario doctoral dissertation; Lyon: Université Lumière-Lyon 2 doctoral dissertation.
- Rupp, Jaime & Nadine de Rupp. 1996. *Diccionario chinanteco de San Juan Lealao, Oaxaca* (Serie de Vocabularios y Diccionarios Indígenas “Mario Silva y Aceves” 35). México, DF: Instituto Lingüístico de Verano.
- Smith-Stark, Thomas C. & Fermín Tapia García. 1986. The active-static nature of Amuzgo (Otomanguean). Paper presented at the annual meeting of the Linguistic Society of America, New York.
- Smith-Stark, Thomas C. & Fermín Tapia García. 2002. El amuzgo como lengua activa. In Paulette Levy (ed.), *Del cora al maya yucateco: Estudios lingüísticos sobre algunas lenguas mexicanas*, 81–129. México, DF: Universidad Nacional Autónoma de México.
- Stewart, Cloyd & Ruth Stewart. 2000. *Diccionario amuzgo de San Pedro Amuzgos, Oaxaca*. México, DF: Instituto Lingüístico de Verano.
- Stump, Gregory T. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Suárez, Jorge A. 1983. *La lengua tlapaneca de Malinaltepec*. México, DF: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.
- Sweet, Henry. 1886. *An Anglo-Saxon primer, with grammar, notes and glossary*, 3rd edition. Oxford: Clarendon.
- Tamura, Suzuko. 2000. *The Ainu language*. Tokyo: Sanseido.
- Valenzuela, Pilar M. 1997. *Basic verb types and argument structures in Shipibo-Conibo*. Eugene, OR: University of Oregon doctoral dissertation.

THE ORGANIZATION OF CHINANTEC TONE PARADIGMS

Matthew Baerman

Surrey Morphology Group, University of Surrey

Enrique L. Palancar

SeDyL-CELIA (UMR8202), CNRS

Surrey Morphology Group, University of Surrey

Abstract

The tonal paradigm of verbs in Tlapeuzco Chinantec counts as one of the most complex and opaque ever described, with five tone values distributed over twelve cells (distinguishing person/number and aspect) to yield c. 70 distinct paradigm types, with no consistent mapping between morphological form and morphosyntactic function. We suggest that useful generalizations will emerge only when we consider units of analysis larger than the individual inflected form, which we dub *inflectional series*. For Tlapeuzco Chinantec this means concatenating the three aspectual forms for each person/number value. The resulting units allow us to see structural relationships between the elements of the paradigm which were previously inaccessible.

1 Introduction¹

Chinantec (or Chinantecan) is a language family spoken in Mexico which is a branch of Oto-Manguean, which in turn is one of the most diverse linguistic phyla in the Americas. Chinantec verb paradigms have served as the poster child for complex inflectional systems (*e.g.* Finkel & Stump 2009), mainly because there is no clear form-function mapping and little predictability of inflection classes. In this paper, we attempt a step in the right direction, looking at the person marking of the verbs of Tlapeuzco Chinantec (henceforth TlCh), a Central Chinantec language. Our analysis is based on

¹ This paper was written under the auspices of the ESRC/AHRC grant RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguean languages” and the ERC grant ERC-2008-AdG-230268 MORPHOLOGY. We thank Olivier Bonami for his invaluable feedback and an anonymous reviewer for his/her very insightful comments.

the study of the data and the grammatical description in the large dictionary by Merrifield & Anderson (2007).

1.1 Some preliminaries about Tlatempuzco Chinantec.

Like the other Chinantec languages, TlCh uses tone as an exponent of inflection. Tonal inflection consists of three level tones (1, 2, 3 = high, mid, low) and two contour tones (31, 32 = high rising, mid rising). Verbs mark subject person/number and aspect. While the pronominal system distinguishes six person/number values, verbs distinguish just four, namely first person singular, first person plural, second person and third person. This can be seen in (1).

(1)	‘I bent it’?	ka ³ -hú ³	hni ²
	‘we (INCL) bent it’	ka ³ -hú ³ 31	hniaŋ ¹
	‘we (EXCL) bent it’	ka ³ -hú ³ 31	hnieŋ ¹
	‘you (SG) bent it’	ka ³ -hú ³	ʔniŋ ²
	‘you (PL) bent it’	ka ³ -hú ³	ʔniaŋ ³²
	‘she, he, they bent it’	ka ³ -hú ²	dsa ²

There are three aspectual stems (PRS, FUT, CPL),² from which additional TA values can be derived through simple prefixation, as shown in (1). For example, the PRS stem on its own expresses incompletive and habitual aspectual values, while the addition of the prefixes *mi*¹- and *ma*²- yield the so-called “Imperfect tense” and the “Perfect tense”. For the purposes of this paper, it is sufficient to focus just on the bare stems in (3).

(1)	TA	1SG	1PL	2	3	Stems
	Present	hú ³²	hú ³²	hú ³²	hú ²	PRS
	Imperfect	mi ¹ -hú ³²	mi ¹ -hú ³²	mi ¹ -hú ³²	mi ¹ -hú ²	
	Perfect	ma ² -hú ³²	ma ² -hú ³²	ma ² -hú ³²	ma ² -hú ²	
	Future	hú ³¹	hú ³¹	hú ³¹	hú ²	FUT
	Potential	mi ¹ -hú ³¹	mi ¹ -hú ³¹	mi ¹ -hú ³¹	mi ¹ -hú ²	
	Past	ka ³ -hú ³	ka ³ -hú ³¹	ka ³ -hú ³	ka ³ -hú ²	CPL
	Hodiernal	na ² -hú ³	na ² -hú ³¹	na ² -hú ³	na ² -hú ²	

² Following the practical system employed by Merrifield and Anderson (2007) we employ tense labels ‘present’ and ‘future’, for the first two stems, when in reality they could be alternatively characterized as ‘incompletive’ and ‘irrealis’.

(3)

	1SG	1PL	2	3
PRS	húʔ ³²	húʔ ³²	húʔ ³²	húʔ ²
FUT	húʔ ³¹	húʔ ³¹	húʔ ³¹	húʔ ²
CPL	húʔ ³	húʔ ³¹	húʔ ³	húʔ ²

1.2 The nature of the problem.

In a maximally transparent inflectional system, we would expect a one-to-one mapping between form and function. That TlCh is not like this is already patently obvious from the repertoire of inflectional formatives: given that there are five tonal formatives distributed across twelve functions (four person/number values x three aspects), a one-to-many mapping (syncretism) is inevitable. And in fact, it works in the other direction as well, with the multiple tonal formatives mapping onto the same function (allomorphy), depending on the lexeme. For a concise illustration, compare the tonal paradigms of *ʔnai*³² ‘sell something’ and *kieg*² ‘lie down’ in (4a) and (4b): both employ the same set of tones, but in an almost completely different distribution.

(4)

	1SG	1PL	2	3
PRS	32	2	2	32
FUT	31	1	1	3
CPL	3	1	1	3

a. [*ʔnai*³² ‘sell something’, Class 13b]

	1SG	1PL	2	3
PRS	2	2	32	2
FUT	1	31	31	3
CPL	2	31	3	2

b. [*kieg*² ‘lie down’, Class 51f]

While such many-to-many mappings are a familiar feature of inflectional systems, TlCh – and indeed all Chinantec languages – takes this to an extreme. An impression of the full extent can be gleaned from Table 1, which charts the tone formatives found in the inflected forms of 613 verbs of TlCh for each person/number value and each TA stem.³ There is an almost complete lack of form-function mapping, *e.g.* all tone units are used for all possible values.

³ Verbs from Merrifield & Anderson’s (2007) dictionary that have complete paradigms, *i.e.* omitting intransitive inanimates that lack first and second person subject forms, as well as a few items with incomplete or contradictory information.

Tone	1SG			1PL			2			3		
	PRS	FUT	CPL	PRS	FUT	CPL	PRS	FUT	CPL	PRS	FUT	CPL
1	149	198	149	146	265	264	151	289	225	10	7	10
2	96	21	92	172	20	43	184	20	85	242	200	240
3	11	67	287	11	68	56	11	67	176	31	335	300
31	70	315	71	74	248	237	70	226	86	12	22	12
32	287	12	14	210	12	13	197	11	41	318	49	51
Total	613	613	613	613	613	613	613	613	613	613	613	613

Table 1: Tone formatives in the inflected forms of 613 verbs of TlCh

At first glance one might believe this is an almost random distribution of forms, but the fact that they fall into ‘only’ around 70 distinct paradigm types suggests that there is some kind of order to them.⁴ But what? As a first step, consider the entropy measures in Table 2, generated using Raphael Finkel and Gregory Stump’s *Analyzing Principal Parts* software.⁵ The table illustrates the predictability of the forms shown in columns, given the forms shown in the rows (measured in bits x 100), factoring in type frequency. The higher the number, the less predictable the form; thus the correct form of the 3rd person present for a given verb can relatively easily be predicted from the 3rd person completive (value of 19), while the correct 2nd person completive is very difficult to predict if one knows just the 1PL future.

	1sgP	1sgF	1sgC	1plP	1plF	1plC	2P	2F	2C	3P	3F	3C
1sgP	---	72	70	121	139	140	40	98	95	53	103	65
1sgF	106	---	34	136	127	142	133	38	131	135	70	134
1sgC	95	25	---	128	140	131	123	63	110	123	94	113
1plP	177	158	158	---	75	64	177	156	190	160	153	170
1plF	186	140	161	66	---	54	188	139	194	178	139	173
1plC	186	155	152	54	54	---	186	153	185	174	152	165
2P	40	99	97	121	142	140	---	66	90	43	98	56
2F	134	40	73	135	128	142	101	---	126	131	58	129
2C	87	89	76	126	139	130	81	83	---	90	119	84
3P	65	113	109	116	144	139	55	107	110	---	63	21
3F	139	72	105	133	128	142	134	59	163	87	---	84
3C	75	110	98	125	137	129	66	104	103	19	58	---

Table 2: Conditional entropy of the tone formatives

What is striking about these figures is that, overall, the entropy measures divide the paradigm into cross-cutting distinctions of person and aspect, as broken down in Table 3.

⁴ As pointed out by an anonymous referee, five allomorphs over twelve cells yields 248,832 possible paradigm types, so we would expect closer to 613 types.

⁵ <http://www.cs.uky.edu/~raphael/linguistics/analyze.html>

3 rd person:	55
1PL:	61
1SG:	67
2 nd person:	91
Present:	94
Future:	108
Completive:	123
Differing in both person and aspect:	129

Table 3: Average entropy within restricted domains of the paradigm

That is, within a given value for person, it is easier to predict the different aspectual forms for a lexeme than it is to predict the different person forms within a given aspect, and both are easier than predicting between forms that differ in both person and aspect. In this way person and, to a lesser extent, aspect, are organizing parameters of the paradigm, in that they define domains of relative interpredictability. This suggests that, in spite of the apparent amorphousness of the tonal paradigm, it is in some sense built on conventional featural distinctions, with person playing a dominant role. We propose here to pursue this idea further by providing an account of the TlCh tonal paradigm as a person-marking system, in which the properties of and relationships between morphological exponents can be described in relatively simple terms. This allows us to demystify, at least somewhat, the otherwise puzzling complexity of the system.

The crucial point to be gleaned from comparing the figures in Tables 1 and 3 is that the discernible role of person in the organization of the paradigm only begins to emerge when we look at its exponence across multiple aspect values. We propose therefore to treat as our basic unit of form not the individual exponents of person and aspect, but rather the combination of all aspect values for a given person. But before we attempt to characterize the paradigm in such terms, we need to dispense with other possible explanations for the tone patterns and the resulting wealth of inflection classes.

2 Rejected explanations

We need to discard first a set of possible alternative explanations for the variation we observe.

2.1 Semantic explanations

Verbs fall into a few different syntactic/semantic classes based on transitivity and animacy. These are relevant for valence increasing mechanisms, for the syntax of agreement and for the syntax of voice (inverse, passive), but have no observable relationship to the tonal formatives. For example, the verbal lexicon is largely split according to the animacy of the ABS argument (object

of transitive or subject of intransitive). The members of animate/inanimate pairs differ along various morphological parameters, including tone, but there are no dedicated animate or inanimate tone patterns. Consider for this purpose the tone of the CPL stem of three gender pairs in Table 4.

CPL	inanimate				animate			
	1SG	1PL	2	3	1SG	1PL	2	3
'buy'	lá ²	lá ³¹	lá ³	lá ³	lán ²	lan ³¹	lán ³	lan ²
'grab'	gúu ²	guu ³¹	guu ¹	guu ³	guén ²	guen ³¹	guén ³	guen ³
'hit'	bá ²	bá ³¹	bá ³	bá ²	bán ²	bán ³¹	bán ³	bán ²

Table 4: Completive tonal paradigm of three animate/inanimate pairs

The inanimate verb 'buy (something inanimate)' has the same tonal inflection as the animate verb 'grab (somebody)', but its animate counterpart 'buy (something animate)' has the same tone pattern as the pair of verbs meaning 'hit', which do not mark animacy through tone. This lack of correlation between tone patterns (or indeed any individual tone value) and animacy is typical of the verbal lexicon as a whole.⁶

2.2 Resegmentation

Besides tone, there is an additional prosodic feature, involving the contrast between the so-called 'ballistic stress' ("characterized by an initial surge and rapid decay of intensity, with a resultant fortis articulation of the consonantal onset" (Foris 2000: 16) and 'controlled stress' (characterized by a longer duration and medium intensity) (see Foris 2000 or Pace 1990; for an alternative phonological oriented account as a glottis phonation feature, see Silverman 1994). If we combine this with tone, it doubles the number of morphological exponents, but it does not bring any noticeable clarity.

2.3 Stem phonology

Tonal alternations do not obviously correlate with any other phonological properties of stems, e.g. nasalization, nucleus quality, etc.

2.4 Morphological correlations

Around a fifth of the verbs (137/613) display stem alternations. While these show some interesting interactions with prosodic alternations there is again no clear predictability in either direction (cf. Baerman, forthcoming).

⁶ The relation between animacy and inflection class is a conceivable one which is for example observed in the distribution of the subject suffixes of another Chinantec language such as Lealao Chinantec (for more details, see Palancar, submitted).

3 The proposal: redefinition of the morphological unit

Following Palancar (forthcoming), and the reasoning outlined in §1.2, we suggest dividing the paradigm up not into twelve cells marking person and aspect, as in (5a), but into four cells marking person, as in (5b). Palancar calls the resulting units ‘inflectional series’.

(5)	<p>a. The 12-cell paradigm</p> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="padding-right: 5px;"></td> <td style="padding-right: 10px;">1SG</td> <td style="padding-right: 10px;">1PL</td> <td style="padding-right: 10px;">2</td> <td style="padding-right: 10px;">3</td> </tr> <tr> <td style="padding-right: 5px;">PRS</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> <tr> <td style="padding-right: 5px;">FUT</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> <tr> <td style="padding-right: 5px;">CPL</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> </table>		1SG	1PL	2	3	PRS					FUT					CPL					<p>b. The 4-cell paradigm</p> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="padding-right: 5px;"></td> <td style="padding-right: 10px;">1SG</td> <td style="padding-right: 10px;">1PL</td> <td style="padding-right: 10px;">2</td> <td style="padding-right: 10px;">3</td> </tr> <tr> <td style="padding-right: 5px;">PRS</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> <tr> <td style="padding-right: 5px;">FUT</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> <tr> <td style="padding-right: 5px;">CPL</td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> <td style="border: 1px solid black; width: 40px; height: 20px;"></td> </tr> </table>		1SG	1PL	2	3	PRS					FUT					CPL				
	1SG	1PL	2	3																																						
PRS																																										
FUT																																										
CPL																																										
	1SG	1PL	2	3																																						
PRS																																										
FUT																																										
CPL																																										

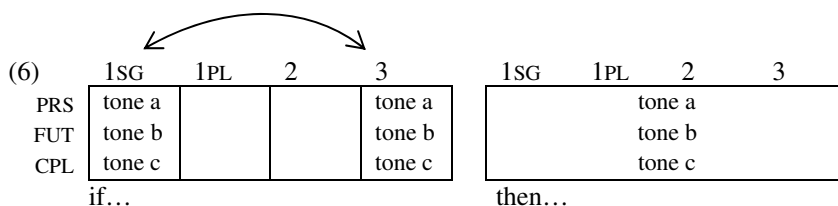
This is the concept of morphological unit we will adopt in this paper, and it will have the shape [# # #], consisting of *present-future-completive tone*. In the following sections, we will show that this new segmentation of the cell reveals a certain degree of regularity both in inflection class assignment and in form-function mapping.

This notion of inflectional series should be distinguished from other types of subparadigmatic unit. For example, Finkel and Stump (2007) employ the concept of *segregated inflection classes*. As an example they take the Sanskrit verb paradigm, which can be divided into the present system and the aorist system. Each of these constitutes a self-contained system, with its own stem and its own set of inflection classes. By contrast, the mapping between the two systems is arbitrary and unsystematic. The Chinantec inflectional series is in effect the mirror image of this. That is, while the individual inflectional series do not themselves constitute morphologically coherent units, once this grouping is made, there is a certain degree of predictability between the resulting units. By analogy, we can think of the inflectional series as phonemes, and the individual tone values as constitutive, distinctive features.

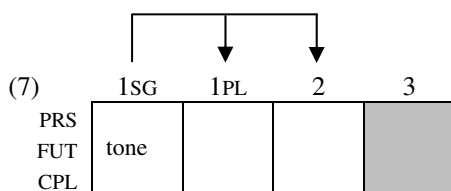
3.1 Inflection class assignment

Applying the concept of an inflectional series to inflectional allomorphy makes it possible to obtain two interesting generalizations. We treat inflectional series as principal parts, employing a maximum of two: one for the 3rd and another for the 1SG (the 1PL and the 2nd would serve equally well, with very minor changes in the numbers):

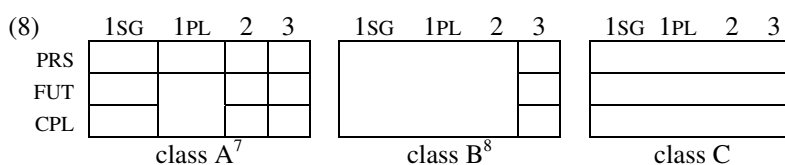
- As shown in (6), if the 1SG form is the same as the 3rd person, then the forms for the 1PL and 2nd person are the same too. This accounts for 123 lexemes, with only one exception.



- As shown in (7), if the 1SG form consists of three identical tones, the 1PL and 2nd person are identical to it. This accounts for 198 lexemes, with five only exceptions.



Overall these two generalizations cover $321/613 = 52\%$ of the lexicon, mapping onto what Merrifield & Anderson (2007) classify as types B & C (there are five exceptional lexemes). Everything else belongs to their class A. This is shown in (8).



The lines in (8) represent the maximal permitted distinctions, but often there will be conflation within that, e.g. the following represents what would count as a class B paradigm (treated as B1f in Merrifield & Anderson, 2007):

⁷ Future/Completive syncretism in the 1PL is almost exceptionless.

⁸ Merrifield & Anderson (2007) characterize class B simply in terms of person syncretism, but since 97% (198/205) of the lexemes also display aspectual syncretism, we treat both as characteristics of this (revised) class.

(9)	1SG	1PL	2	3
PRS				32
FUT	3			
CPL				
	class B			

The remainder of the lexicon (class A) still needs another account. One option is to construe one of the inflectional series as the default exponent for each person value. We have chosen the most frequent ones, and they account for 55% (129/285) of the remaining data-points.

(10)	1SG	1PL	2	3
PRS	32	32	32	32
FUT	31	31	31	3
CPL	3	31	3	3
	233	156	105	135
	(82%)	(54%)	(37%)	(48%)

Note that the figures in (9) represent the frequency for each form in isolation, not a default paradigm as such; only 16 lexemes (class A11a) actually show all four at once.

3.2 Form-function mapping

3.2.1 The marking of person

The individual inflectional series can be seen as exponents of person, at least to a certain extent. This is shown in Table 5 below (in the table, everything with a token count under 6 (1%) is not highlighted). If we set aside the forms which do not distinguish person at all (C paradigms), the following two generalizations hold:

- (i) The 3rd person form is distinct from 1st (SG/PL) and 2nd person.
- (ii) While 1SG is always distinct from 1PL, 2nd person is usually syncretic with 1SG or 1PL.⁹

Considering now the C paradigms, we can make a third observation:

- (iii) C paradigm forms typically coincide with the distinct 3rd person forms found in the A and B paradigms (88/117).

⁹ 2nd person is [2 1 1] or [32 31 31] only if 1PL is too (65/67 and 15/15); 2nd person is [32 31 3] only if 1SG is too (102/105).

The broad generalization that emerges is that the system potentially distinguishes 1SG, 1PL, 2 and 3, but that (i) 2nd person often takes the form of either 1SG or 1PL, (ii) all 1st and 2nd person values may conflate in a dedicated non-3rd person form (B paradigms), and (iii) forms that do not distinguish person at all often coincide with 3rd person form.

3.2.2 The marking of aspect

The inflectional series are made up of tonal alternations for aspect. By looking now at the system from the perspective of aspectual marking, we gain some insight into how the series are constituted. Consider again the figures from Table 1 above. If we now divide them by paradigm type (A, B, C), we see there are some striking regularities in the distribution of the tonal components. Take first the A paradigm verbs illustrated in Table 6 (in the tables, figures under 12 lexemes (2%) are not highlighted).

The patterns emerge most clearly if we group the individual tonal exponents into three tone types:

- HIGH, made up of high tone (1) and ascending to high (31)
- MID, made up of mid tone (2) and ascending to mid (32)
- MID/LOW, made up of mid (2) and low (3)

Series	A				B		C
	1SG	2	1PL	3	1/2	3	1/2/3
[1 1 1]	4	6	1	1	137		6
[1 3 1]							3
[2 1 1]		67 ←	118				
[2 1 2]	52	46	4				
[2 1 3]		20					
[2 1 32]		10					
[2 2 2]	1			172		3	14
[2 2 3]				1			4
[2 3 2]				16		1	17
[2 3 3]				1		2	7
[2 31 3]							1
[2 31 31]			7				
[3 3 3]					1	21	10
[31 31 31]	1		4	1	60	1	10
[32 2 3]				1			2
[32 3 3]			1			76	
[32 3 32]							2
[32 31 1]		7					
[32 31 3]	233	→105	→2				9
[32 31 31]	1 ←	15 ←	155				
[32 3 2]				16			1
[32 3 3]				211			26
[32 31 32]		16					
[32 32 32]				15		24	11

Table 5: Inflectional series as exponents of person

Tone	PRS				FUT				CPL			
	1SG	1PL	2	3	1SG	1PL	2	3	1SG	1PL	2	3
1	3	0	5	1	53	120	144	1	3	118	79	1
2	51	127	139	120	1	0	0	104	51	2	44	131
3	0	0	0	0	0	1	0	168	234	3	123	141
31	0	4	0	1	235	168	146	1	1	167	16	1
32	236	159	146	163	1	1	0	11	1	0	28	11

Table 6: Tones and aspect in A paradigm verbs

The tone/aspect mappings that result are the following:

- (i) The PRS stem is MID tone.
- (ii) The FUT of non-3rd person is HIGH tone and MID/LOW for 3rd, keeping the marking of 1/2 vs. 3rd distinct.
- (iii) The marking of the CPL stem is complex: 1PL is syncretic with FUT; 2nd person can be realized by almost anything; and 1SG and 3rd person have MID/LOW tone. In the 3rd person, this means that the same tone profile as in the FUT, but the forms are not necessarily syncretic.

All this can give us a perspective on the syncretism of person values seen above in Table 5. For example, the distinct 3rd person FUT tone profile ensures that the 3rd person form will never be syncretic with the 1st and 2nd person forms. Other person values have distinct tone profiles only in the CPL. The fact that 2nd person CPL can have any tone means that sometimes it will be distinct from the others, sometimes coincide with them, accounting for the variable syncretic behaviour of the 2nd person seen in Table 5. Now let us see what happens in B and C paradigms.

Tone	B				C			
	1/2	3			(TA inflecting)			(uninflecting)
		PRS	FUT	CPL	PRS	FUT	CPL	
1	137	0	0	0	3	0	3	6
2	0	76	73	74	29	6	18	14
3	1	21	100	99	0	56	49	10
31	60	1	1	1	0	10	0	10
32	0	100	24	24	40	0	2	11

Table 7: Tones and aspect in B and C paradigm verbs

The aspectless 1/2 form of the B paradigms has the same tone profile (MID) as the 1/2 FUT of the A paradigms. Whereas C paradigm verbs that inflect for TA behave more or less like 3rd person forms, the tone of uninflecting verbs does not show any particular tendency, suggesting that tone assignment here is given in the lexical entry.

4 Conclusion

In this paper, we have argued that the notoriously opaque tonal inflection of TlCh can be at least partially tamed by a judicious choice of analytic units. First, we showed that person can be taken as the dominant organizing principle, in that person values define the most reliable domains of interpredictability within the paradigm. From this observation we extracted a

morphological unit that we have labelled inflectional series, made up of the three aspectual forms for each person value. By thus construing the tonal paradigm as a four-cell paradigm marking person, we were able to describe the distribution of forms in terms of a few patterns of syncretism, and concisely characterize the inflection classes in terms of implicational relationships between these forms. We then examined the nature of the inflectional series more closely by looking at aspectual marking in isolation, consolidating tonal marking into three major types. This showed that, given the person/number value and aspect, tonal inflection is highly constrained, the complexity of the system emerging at certain discrete points (*e.g.* the completive of A class verbs).

For obvious reasons, our analysis in terms of inflectional series cross-cuts an analysis in terms of optimal principal parts. Because of the relative interpredictability of individual tones within each inflectional series, optimal principal parts will tend to be spread across the series. For example, as Olivier Bonami (personal communication) has shown, if we limit ourselves to the four cells that are the best predictors of the entire inflectional paradigm, they will be drawn from each of the four inflectional series, as shown in (11). (These reduce the average conditional entropy of the remaining cells to 0.138 bits, or 138 in terms of the scale used in Table 2.)¹⁰ Significantly, these optimal principle parts are drawn from each of the three aspects as well; recall from Table 3 that aspect also defines a zone of interpredictability (though less reliable than person), so this is what we would expect. Thus the selection of principle parts is complementary to the parameters of paradigmatic organization that we have claimed operate in TlCh, namely person and aspect.

(11)

	1SG	1PL	2	3
PRS				
FUT				
CPL				

References

Baerman, M. Forthcoming. "Inflectional class interactions in Oto-Manguean", in J-L. Léonard and A. Kihm (eds.), *Issues in Meso-American morphology*. Paris: Michel Houdiard.

¹⁰ Though note that this measure does not factor in type frequency; if it did, the figure would be lower.

- Finkel, R and G Stump. 2009. "Principal parts and degrees of paradigmatic transparency", in J. P. Blevins and J. Blevins (eds), *Analogy in Grammar: Form and Acquisition*. Oxford: Oxford University Press, 13-53.
- Finkel, R and G. Stump. 2007. "Principal parts and morphological typology". *Morphology* 17: 39-75.
- Foris, D. P. 2000. *A grammar of Sochiapam Chinantec* (Studies in Chinantec Languages 6). Dallas, TX: SIL International and the University of Texas at Arlington.
- Merrifield, W. R. and A. E. Anderson. 2007. *Diccionario Chinanteco de la diáspora del pueblo antiguo de San Pedro Tlatepuzco, Oaxaca*. [2nd Edition] (Serie de vocabularios y diccionarios indígenas "Mariano Silva y Aceves" 39). Mexico DF: Summer Linguistic Institute.
- Pace, W. J. 1990. "Comaltepec Chinantec verb inflection", in W. R. Merrifield and C. R. Rensch (eds.), *Syllables, tones and verb paradigms* (Studies in Chinantec Languages 4) Arlington, Texas: The Summer Linguistic Institute and the University of Texas at Arlington, 21-62.
- Palancar, E. L. Forthcoming "Revisiting the complexity of the Chinantec verb conjugation classes", in J-L. Léonard and A. Kihm (eds.), *Issues in Meso-American morphology*. Paris: Michel Houdiard.
- Palancar, E. L. Submitted "When two agreement systems compete: The suffix classes of Lealao Chinantec" MS submitted to *Morphology*.
- Silverman, D. 1994. "A Case Study in Acoustic Transparency: [spread glottis] and Tone in Chinantec". *Proceedings of NELS* (Amherst: University of Massachusetts) 24, 559-572.

In the paper, I introduce the novel concept of an 'inflectional series', which I take to be an invariant set of prosodic exponents involved in the making of the principal stems of a Chinantecan verb. I take inflectional series to play an important role in the overall design of the verbal inflection of TCh. Using this unit as an organizing principle, I propose that the verbs of TCh fall into four different conjugation types, which I call 'conjugation patterns', which represent inflectional templates of verbs attending to the way grammatical person is realized.

Verbs of conjugation pattern 1 are uninflected, but all other verbs fall into different conjugation classes. The conjugation classes of TCh are typologically interesting because they are structured in complex ways. As I see it, much of this complexity lies in the way third person is encoded with respect to other grammatical persons. The other grammatical persons relevant for the grammar of TCh are: a non-third person, on the one hand, and the first person singular, the second person and the first person plural, on the other.

For TCh, I propose that the ways the third person is realized are independent from the ways involved in the expression of other grammatical persons. For example, it is the case that the forms that realize a third person value for verbs of patterns 3 and 4 are also found in verbs of patterns 1 and 2, but they do not realize person. I take this to mean that the forms that realize a third person value for some verbs acquire this function only by virtue of the existence of other forms in the paradigm of the same verbs which genuinely realize other grammatical persons.

One of the crucial points about the conjugation classes of TCh is the following: verbs that belong to patterns 3 or 4 belong simultaneously to patterns 1 or 2 for the purpose of realizing the third person. This simultaneous membership has consequences for the way conjugation classes are structured in the language. As an illustration, consider the following scenario.

Verbs of pattern 2 belong to ten conjugation classes, as presented in §5.2. In this view, when a verb of patterns 3 or 4 belongs to pattern 2 for the encoding of the third person, such a verb simultaneously belongs to one of the ten conjugations of pattern 2. In reality, such verbs only belong to six of those ten classes. Verbs of pattern 3 fall into six other classes themselves presented in §5.3, and those of pattern 4 into 29, given in §5.4. Consequently, a verb can belong simultaneously to different inflectional classes.

On the other hand, verbal lexemes in TCh need to store information about this simultaneous membership, i.e. lexeme *x* belongs to class *A* for the encoding of the third person but to class *a* for the encoding of other persons; but lexeme *y* also belongs to class *A* for the third person but to class *β* for other persons; while lexeme *z* belongs to class *B* for the third person, but to class *β* for other persons, and so on and so forth. These different combinations need to be stored as inflectional properties of lexemes. But as some lexemes have the same type of membership, i.e. a given lexeme *xx* may select the same inflectional exponents like lexeme *y*, the lexemes in question also gather in classes, but at a more abstract level. I call such classes 'superordinate inflectional classes' and constitute a unique morphological property of TCh.

The possible permutations resulting from combining the classes for the third person with those of other persons are incredibly numerous. The numbers are increased further by adding verbs which select an uninflected form for the third person, that is, verbs which simultaneously belong to pattern 1 for the encoding of the third person, but to other classes for the rest of the persons. Such a scenario is responsible for the high degree of apparent irregularity of the inflectional system of a Chinantecan language. For example, for verbs of pattern 3 we could obtain up to 66 different such combinations (6 classes of pattern 3 x 11 possibilities for the encoding of the third person (10 classes of pattern 2, plus the uninflected form proper of pattern 1), whereas for the verbs of pattern 4, the combinations would amount to 319, resulting into 385 hypothetically possible superordinate inflectional classes!

In reality not all combinations are allowed, suggesting there is some internal structure that constrains the system. In a sample of 790 verbs, we find that only 58 such permutations are allowed (i.e. there are only 58 superordinate inflectional classes), and while 25 such pairings are instantiated by one lexeme in each case, there are also combinations which are instantiated by dozens of lexemes, suggesting the existence of defaults in the choice of inflections.

To show all this, the paper is structured as follows. In the next section, I present an introduction to the verb types that exist in Chinantecan languages as well as to their most representative features of their inflection. In section 3, I introduce the traditional way the conjugation classes of Chinantecan languages have been characterized and I show how Merrifield and Anderson (2007) (henceforth M&A) argue that there are 89 inflectional classes in TCh. An alternative view of such classes is needed to allow for a better perspective on the generalizations of the system. This is provided in section 4, and it is implemented in section 5. The paper concludes in section 6 with a summary and a review of the proposal.

2. VERB TYPES IN THE CHINANTECAN LANGUAGES

In the Chinantecan languages, there are two main types of verbs: stative verbs and dynamic verbs (also called 'active verbs'). While the inflection of stative verbs remains poorly understood to date and for that reason they will be left out in this paper, we know more of the inflectional properties of dynamic verbs. Dynamic verbs fall into two types: (i) simple verbs as in (1a), which consist of a mono-morphemic and (mostly) monosyllabic root, and (ii) bi-partite verbs (also called 'binomial') given in (1b), which originated from old V+V compounds, and are thus disyllabic and consist of two morphs. Additionally, bi-partite verbs have a number of inflectional idiosyncrasies of their own (see appendix 1 for more details).²

- (1) a. *jih^hʔ* (tr) 'bend'; *hiɬim^hʔ* (tr an) 'push'; *gøʔ* (intr) 'break down', etc.
- b. *ɪ^h+hɪɪ^hʔ* (intr an) 'go out'; *gum^hʔ+dkən^hʔ* (tr an) 'ride'; *mɪ^h+leɪ^hʔ* (tr) 'show', etc.

The inflection of dynamic verbs can be complex. Maximally inflected verbs have a number of stems or bases, as in (2).

- (2) - Present stem (or progressive stem)
- Future stem (or tentative stem)
- Completive stem (or past stem)
- Directive stem (or ambulative stem)
- Round-trip stem (or andative-venitive stem)
- Imperative stem
- Other stems: hodiernal, evidential, injunctive, and prohibitive

Such stems are built on a base to which tone and pharyngealization are added as inflectional exponents.³ Such exponents also commonly realize grammatical person; the maximal person distinction involves four persons: 1sg, 1pl, 2 and 3. An example of the paradigm of a maximally inflected verb appears in (3) with the paradigm of the verb *jih^hʔ* (tr) 'bend' from TCh (Merrifield and Anderson 2007) (henceforth M&A). The cell in shading {3, present stem} is used as the citation form for the lexeme. In the presentation of the paradigm, I include the cells for the third person first because it provides a convenient way to reflect important properties of the structure of verbal inflection, as will become clear.

(3)	TCh <i>jih^hʔ</i> (tr) 'bend'	3	1sg	2	1pl				
Present stem	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr> <td style="padding: 2px;"><i>jih^hʔ</i></td> <td style="padding: 2px;"><i>jih^hʔ</i></td> <td style="padding: 2px;"><i>jih^hʔ</i></td> <td style="padding: 2px;"><i>jih^hʔ</i></td> </tr> </table>	<i>jih^hʔ</i>	<i>jih^hʔ</i>	<i>jih^hʔ</i>	<i>jih^hʔ</i>				
<i>jih^hʔ</i>	<i>jih^hʔ</i>	<i>jih^hʔ</i>	<i>jih^hʔ</i>						

² Abbreviations: C completive stem; CLF nominal classifier; *dir-an* ditransitive with animate (theme) object; *dir* ditransitive with inanimate (theme) object; F future stem; *intr an* intransitive with animate subject; *intr* intransitive with inanimate subject; P past stem; PST past; *tr an* transitive with animate object; *tr* transitive with inanimate object.

³ Stems may undergo palatalization, apophony and consonant mutation and may also realize gender agreement, inversion, passive, and applicative morphology.

Future stem	<i>júh²</i>	<i>júh¹³</i>	<i>júh¹⁵</i>	<i>júh¹³</i>
Completive stem	<i>júh²</i>	<i>júh¹</i>	<i>júh¹</i>	<i>júh¹³</i>
Directive stem	<i>júh²</i>	<i>júh¹</i>	<i>júh¹</i>	<i>júh¹³</i>
Round-trip stem	<i>júh¹³</i>	<i>júh¹³</i>	<i>júh¹³</i>	<i>júh¹³</i>
Imperative stem		<i>júh¹</i>		

The different grammatical tenses are built by adding TAM prefixes (at times also in the form of prefix series and proclitics) to the different stems. An illustration is given in (4). It may be noted that present and future are tense values realized by the bare stem. The examples in (4) are given for the first person singular of the verb *júh²* (tr) 'bend', the enclitic =*jni* is an obligatory subject agreement marker for the first person singular.

(4) Tense:	Prefix:	Stem:	Example:
<i>present</i>	-	Present stem	<i>júh² = jni</i> 'I'm bending'
<i>imperfect</i>	<i>mí²-</i>	Present stem	<i>mí²-júh¹³ = jni</i> 'I was bending'
<i>future</i>	-	Future stem	<i>júh¹³ = jni</i> 'I'll bend'
<i>potential</i>	<i>mí²-</i>	Future stem	<i>mí²-júh¹³ = jni</i> 'I'd bend'
<i>past</i>	<i>ca¹-</i>	Completive stem	<i>ca¹-júh¹ = jni</i> 'I bent'
<i>hodiernal</i>	<i>na²-</i>	Completive stem	<i>na²-júh¹ = jni</i> 'I've bent (today)'

Starting from Merrifield (1968), the existing descriptive tradition of Chinantecan languages from the SIL suggests that the entire paradigm of a verb is retrievable from the inflectional information provided by only 12 cells, given in shading in (5). In this light, other cells in the paradigm are predictable from these 12 cells by rules of referral (except the stems for the third person, of which little, if nothing at all, is added in the sources).

(5)	3	1sg	2	1pl
Present stem	X	X	X	X
Future stem	X	X	X	X
Completive stem	X	X	X	Use future stem
Directive stem	X?	Use completive stem	Use completive stem	Use future stem
Round-trip stem	X?	Use future stem	Use future stem	Use future stem
Imperative stem		Use completive stem	Use completive stem	

The 12 principal cells are made up of a combination of prosodic exponents as illustrated in (6) for the principal parts of the verb *júh²* (tr) 'bend'. The language has six tones, but only five of them have a morphological function: three of them are level tones (3 is high, 2 is middle and 1 is low) and two are ascending tones (12 is low to middle and 13 is low to high). Additionally, notice that in the principal cells in (6) there is a high degree of syncretism.

(6) <i>júh²</i> (tr) 'bend'	3	1sg	2	1pl
Present stem	2	12	12	12
Future stem	2	13	13	13
Completive stem	2	1	1	13

Apart from tone, there is another phenomenon that plays an important role in the making of inflectional paradigms in Chinantecan languages. M&A and other SIL linguists posit that syllables in Chinantecan languages have two stress types: "controlled stress" vs. "ballistic stress". A characterization of these stress types is found in Foris (2000: 16, but also in other authors such as Rensch 1978, among others): "Ballistic syllables are characterized by an initial

surge and rapid decay of intensity, with a resulting fortis articulation of the consonantal onset. [...] Controlled syllables generally display a more gradual surge and decay of stress, as well as a longer duration of the maximum stress."

In contrast to this received view, ballistic stress has also been characterized as a reflection of laryngeal phonation either by an increased sub-glottal pressure in Mugsle (1982) or by a laryngeal abduction in Silverman (1994).⁴ M&A indicate ballistic stress by means of an acute stress vowel. In the paradigms, two phenomena may occur involving ballistic stress. One is its implementation in a lexical stem which otherwise would not have it in the lexical representation of the lexeme. Another is its removal in stems that would naturally have it in the lexical representation. The implementation is indicated here by the plus sign (+), its removal with the subtraction sign (-).

With this in mind, and following M&A's analysis, there are 11 prosodic exponence rules that play a role in the making of a verbal paradigm in TCh. These are spelled out in (7).

- (7) 1 [add low tone]
- + 1 [add ballistic stress (if pertinent) & add low tone]
- 2 [add middle tone]
- + 2 [add ballistic stress (if pertinent) & add middle tone]
- 2 [remove ballistic stress (if pertinent) & add middle tone]
- 3 [add high tone]
- + 3 [add ballistic stress (if pertinent) & add high tone]
- 3 [remove ballistic stress (if pertinent) & add high tone]
- 12 [add ascending tone to middle tone]
- 13 [remove ballistic stress (if pertinent) & add ascending tone to middle tone]
- 13 [add ascending tone to high tone]

In (8), I give the overall distribution of the prosodic exponents in (7) from a sample of the 182 verbs of M&A that have contrastive (non-syncretic) person marking for all grammatical persons, (P stands for 'present stem', F for 'future stem', and C for 'completive stem'). The total is of 2,184 inflected forms, i.e. 546 forms inflected for each grammatical person per verb.

(8)	3			1sg			2			1pl		
	P	F	C	P	F	C	P	F	C	P	F	C
1	135	121	256		3				132			
+1				18			3		132			
2	4	3	7	119	42	161	1	1	3	114		114
+2	32	32	45	109	18		18	48	48	96		
-2	4		4	8								
3												
+3							137	137	48		92	184
-3				2	76	2	80				15	
12	141	11	11	163	43	104	133	1	1	135	66	1
-12							10					
13	1	1	1	3			15			132	1	73
	182	182	182	546	182	182	546	182	182	546	182	182
												546

As (8) shows, the prosodic exponents {1}, {+2} and {12} appear to be defaults in the realization of third person. However, these are also involved in the realization of the second person in addition to {+3} and {13}. In general, although there are tendencies across the board, most prosodic exponents are involved in the realization of more than one grammatical person, with the exception of {3}, which just realizes first person plural.

⁴ I want to thank Daniel Silverman (p.c.) for pointing this to me.

The distribution of the prosodic exponents in the realization of the three principal stems is given in (9). Here we observe that there is a tendency to keep the present and the future stems as maximally distinct formally. The figures in (9) are calculated from a total of 728 inflected forms for each stem (e.g. 182 present stems inflected for the four persons = 728 forms).

(9)	P	F	C
12	383	31	55
2	238	46	1
+2	98	32	93
13	221	75	
1	138	255	
3	92	92	
+3	63	137	
-3	77	18	
+1	18	-	
-1	10	-	
-2	4	4	
	728	728	728

In this section, I have shown that there are 11 different prosodic formatives involved in the realization of the principal 12 cells of the paradigm of a maximally inflected verb in TCh. Some exponents are used more than others in the realization of person, and there is a clear contrast or preference for those involved in the realization of TAM values in the three stems, i.e., ascending {12}, and middle tone are mainly involved in the realization of the present stem, while ascending {13}, low tone {1} and high tone are mainly involved in the realization of either the future stem or the completive stem.

3. CONJUGATION CLASSES IN TCH: THE TRADITIONAL VIEW

As they select different inflectional exponents for their inflection, dynamic verbs (simple and bi-partite) fall into different conjugation classes.⁵ This is illustrated if we compare the exponents involved in the principal cells for the paradigms of the two verbs *jih²* (tr) 'bend' and *tiu²* (tr) 'cut' in (10).

(10)		3	1sg	2	1pl
<i>jih²</i> (tr) 'bend'	Present stem	<i>jih²</i>	<i>jih²</i>	<i>jih²</i>	<i>jih²</i>
	Future stem	<i>jih²</i>	<i>jih¹³</i>	<i>jih¹³</i>	<i>jih¹³</i>
	Completive stem	<i>jih²</i>	<i>jih¹</i>	<i>jih¹</i>	<i>jih¹³</i>
<i>tiu²</i> (tr) 'cut'	Present stem	<i>tiu²</i>	<i>tiu²</i>	<i>tiu²</i>	<i>tiu²</i>
	Future stem	<i>tiu¹</i>	<i>tiu¹³</i>	<i>tiu¹³</i>	<i>tiu¹³</i>
	Completive stem	<i>tiu¹</i>	<i>tiu¹</i>	<i>tiu¹³</i>	<i>tiu¹³</i>

The inflectional exponents involved in the paradigm of these two verbs are given in (11). The class to which the verb *jih²* (tr) 'bend' belongs requires a middle tone and ballistic stress for the third person. However, this is not obvious in (10) above because the verb in question has an inherent ballistic stress in its lexical representation. Similarly, the class to which the verb *tiu²* (tr) 'cut' belongs requires removal of ballistic stress for the completive stem and the second person, but as this verb is not taken to have ballistic stress in its lexical representation.

⁵ Aronoff (1994: 64) defines an inflectional class as "a set of lexemes whose members each select the same set of inflectional realizations".

(11)

The *jih²* class

	3	1sg	2	1pl
Present stem	+2	12	12	12
Future stem	+2	13	13	13
Completive stem	+2	1	1	13

The *tiu²* class

	3	1sg	2	1pl
Present stem	12	12	2	2
Future stem	1	13	+3	3
Completive stem	1	1	-3	3

The class to which *jih²* (tr) 'bend' belongs contains 94 verbs in M&A's dictionary; the one to which *tiu²* (tr) 'cut' belongs has 66 verbs. There are many other classes. Fundamentally, M&A's (2007) analysis of the conjugation classes of TCh rests on the concept of a 'triplet', which can be defined as in (12).

(12) **Triples.** An invariant set of prosodic exponents involved in the making of the three principal stems for a specific grammatical person.

There are nine different triplets which are involved in the realization of the third person, given in (13).

(13)

Examples of triplets of the 3rd person

	1	2	3	4	5	6	7	8	9
Present stem	12	12	12	+2	-2	2	13	2	-3
Future stem	1	1	12	+2	1	2	13	1	-3
Completive stem	1	+2	12	+2	-2	1	13	1	-3

M&A classify verbs in four different groups according to the quantity of triplet types they require: **group A** verbs require four triplets (each triplet is used for the four grammatical persons: 1sg, 1pl, 2 and 3); **group B** verbs require two triplets (one is used to encode the third person; the other one for all other persons (i.e. a non-third person)); and **group C** verbs require only one triplet for all persons (i.e., there is no person distinction).⁶ Next, I will be presenting the inflection of such groups separately.

3.1. **GROUP A VERBS.** M&A (2007: 692) propose that group A verbs inflect for grammatical person by selecting one among a set of nine different triplets for each person.

(14)

	1	2	3	4	5	6	7	8	9
Present stem	12	12	12	+2	-2	2	13	2	-3
Future stem	1	1	12	+2	1	2	13	1	-3
Completive stem	1	+2	12	+2	-2	1	13	1	-3

⁶ M&A also identify a group of verbs they call 'group D' but the verbs in such a group are all intransitive verbs that belong to group C but have an inanimate argument, and for this reason they

(15)

		2 nd person for group A								
		1	2	3	4	5	6	7	8	9
Present stem		I2	2	2	I2	I2	+2	2	I2	2
Future stem		I3	+3	+3	I3	I3	+3	+3	I3	-3
Completive stem		I	2	-3	I3	I2	+I	-I2	-3	-3

(16)

		1 st person (sg & pl) for group A														
		a			b			c			d			e		
		Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	
Present stem		I2	I2	2	2	I2	I2	+2	2	2	2	2	2	+2	2	
Future stem		I3	I3	I3	3	+3	I3	+3	I3	+3	+3	+3	+3	+3	I3	
Completive stem		I	I3	I	3	+2	I3	+2	I3	+2	+3	+2	+3	+2	I3	

		f											
		g			h			i			j		
		Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl	Isg	Ipl
Present stem		-3	I3	2	-3	I2	I3	-3	I2	-3	I2	-3	I2
Future stem		-3	I3	2	-3	I2	I3	-3	I3	-3	I3	-3	I3
Completive stem		-3	I3	2	-3	I2	I3	-3	I3	-3	I3	-3	I3

The conjugation class to which a given verb belongs to is indicated in M&A's dictionary by a combination of the alphanumeric characters given to the triplets in (14-16), as shown in (17).

3rd person, 2nd person, 1st person

(17) {A11a}, {A22b}, {A45d}, etc.

For example, the classes to which the verbs *jilil²* (tr) 'bend' and *tilil²* (tr) 'cut' in (10) belong are described as being {A41a} and {A13b}, respectively. This notation attends to the different triplets they require for their inflection. This should be clear if the principal cells in (11) are checked out against the charts in (14-16). The total number of possible permutations in (11) are 721, but in reality, the grammar of the language only allows 55 such possibilities (i.e. 55 conjugation classes). The combinations are listed in (18), from M&A (692-693).

(18)

{A11a}	{A14a}	{A16f}	{A21a}	{A32f}	{A42e}	{A46e}	{A56a}
{A12a}	{A15a}	{A17a}	{A22b}	{A33h}	{A42f}	{A47a}	{A61a}
{A12b}	{A15d}	{A17e}	{A23b}	{A34a}	{A43b}	{A48}	{A65a}
{A12d}	{A15f}	{A18a}	{A27a}	{A35a}	{A44a}	{A49b}	{A79i}
{A13a}	{A16b}	{A18d}	{A31a}	{A41a}	{A45a}	{A51b}	{A82b}
{A13b}	{A16d}	{A19g}	{A32d}	{A42b}	{A45d}	{A51e}	{A99g}
{A13f}	{A16e}	{A19j}	{A32e}	{A42d}	{A46d}	{A51f}	

These 55 classes are not equally populated by the same number of verbs. M&A report that 75% of the verbs of group A belong to 15 such classes. This implies that a quarter of group A verbs is distributed across 40 minor classes.

3.2. GROUPS B AND C VERBS: The inflection of verbs of groups B and C are described in a similar way, (M&A: 695, 696).

(19)

		3 rd person for group B								
		1	2	3	4	5	6	7	8	9
Present stem		I2	I2	I2	+2	-2	I2	I3	2	I
Future stem		I	I	I2	+2	I	2	I3	I	I
Completive stem		I	+2	I2	+2	-2	I	I3	I	I

(20)

		non-3 rd person for group B							
		a	b	c	f	g	h		
Present stem		I2	2	I2	I	I3	-3		
Future stem		I3	+3	I3	I	I3	-3		
Completive stem		I	2	I3	I	I3	-3		

Again, the possible permutations for the verbs of group B sum up to 54, but in reality only the 19 classes in (21) are allowed in the language, (M&A 2007: 696).

(21)

{B1f}	{B6a}
{B3g}	{B7h}
{B3h}	{B8h}
{B1h}	{B4b}
{B2h}	{B9g}
{B4g}	{B9h}
{B3a}	{B4h}
{B3b}	{B5b}
{B3c}	{B5h}

Finally, verbs of group C have the 15 conjugation classes in (22).

(22)

		All persons for group C														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Present stem		I2	I2	I2	+2	-2	I2	I3	2	I	I2	2	2	-3	-3	I2
Future stem		I	I	I2	+2	I	2	I3	I	I	I3	I3	2	-3	I	I
Completive stem		I	+2	I2	+2	-2	I	I3	I	I	I	I	-3	-3	I	I

All conjugation classes put together sum up to a total of 89 (55 of group A, 19 of group B and 15 of group C). M&A (2007: 691) actually propose 103 because they consider the 14 different classes of verbs of group D, but in reality these are just verbs of group C, although with an inanimate subject.⁷

The notation used by M&A for conjugation classes has pedagogical advantages, but it has serious shortcomings for the linguist interested in perceiving grammatical generalizations because it obscures the perception of encoding similarities across the classes. For example, the notation veils the way similar exponents are employed in the realization of the different persons. As an illustration, consider for example 'triplet 1' in (15) for the second person, and notice that it has the same exponents as 'triplet *a'*' in (16) for the first person singular. This points to the possible existence of syncretism between these two grammatical persons, but interestingly, the same exponents also occur in 'triplet *a'*' in (20), and this time they realize a 'non-third person', and even in 'triplet 10' in (22) when they do not realize person at all. All this is terribly confusing and an alternative view is helpful to shed light on how the inflectional classes are constructed. In the next section, I propose one such alternative view using a morphological ontology I call 'inflectional series'.

⁷ Triplet 11 in (22) is apparently the only one not employed in the inflection of an intransitive inanimate verb, but this may be due to chance.

4. INFLECTIONAL SERIES AS AN ALTERNATIVE VIEW TO THE CONJUGATION CLASSES OF TCh.

I will follow M&A's tenet that the triplets have an important bearing in the shaping of the inflectional paradigms in TCh and not just the prosodic exponents by themselves. However, I hold an alternative view of their data and propose treating each different triplet as a different and mutually exclusive unit, which I call 'inflectional series', defined as in (23).

(23) **Inflectional series.** An invariant set of prosodic exponents involved in the making of the three principal stems, e.g. {12, 1, 1}, {-2, 1, -2}, {12, 2, 1}, etc.

There are 24 such inflectional series, given in (24) arranged according to their distribution in the encoding of grammatical person. Five of these inflectional series involve the same prosodic exponent for all stems, i.e. there is no stem distinction, e.g. {+2}.

(24)	P	F	C	No person	3	non-3	1sg	1sg+2	2	1pl
I		+2		17.6%	35.7%		0.6%			
II		I2		7.9%	7.3%		0.6%			
III		I		6.5%	4.6%	0.4%				
IV		I3		4.3%	0.4%					1.3%
V		-3		2.3%	0.2%	96.3%	3.9%	1.4%		0.5%
1	I2	I	I	28.3%	43%					
2	-2	I	-2	15.2%	3.4%					
3	I2	I	+2	1.6%	3.6%					
4	2	I	I	4.6%	0.6%					
5	I2	I3	I	4.3%		0.9%	72.5%	96.1%	1.6%	
6	2	2	I	3.2%	0.8%					
7	-3	I	-3	1.6%						
8	I2	2	I	1.3%	0.2%					
9	I2	I	I2	1.0%						
10	2	I3	I	0.3%						
A	2	+3	-3						37.0%	5.2%
B	2	+3	2			2%			23%	
C	+2	+3	+I						9.8%	
D	I2	I3	I3			0.4%			8.2%	54.7%
E	I2	I3	I2						9.8%	
F	2	+3	-I2						5.4%	
G	I2	I3	-3						3.8%	
a	+2	+3	+2						26.3%	35.9%
b	2	+3	+3							2.4%
c	2	I3	I3							

The inflectional series in (24) appear in four different alphanumeric notations. The different notations are meant as mnemonic indexes to the following criteria:

- Numbers, whether Roman (I-V) or Arabic (1-10), indicate series which realize the third person.
- Roman numbers indicate series with no TAM distinction.
- Arabic numbers indicate a series that has a TAM distinction in at least two stems.
- Letters, whether capital (A-G) or lower case (a-c), indicate that the series are NOT used for the third person.
- Capitals indicate series used 'mainly' for the second person.

- Lower case letters indicate series used for the first person only.

With this in mind, I present next the results of an analysis of the conjugation classes of TCh based on the implementation of the inflectional series in (24). The study in question is based on a sample of 760 verbs from M&A, but first a word about the sample is in order.

M&A (2007: 684) report that the TCh dictionary contains 912 entries of dynamic verbs with enough inflectional information to discern their class membership. From my experience, this is not entirely true. Of these 912 verbs, I have only been able to find 906: 748 simple verbs plus 158 bi-partite verbs. The sample on which I base this study consists of 790 dynamic verbs. This means that I have not included a total of 122 other verbs found in the dictionary. This is due to the following reasons: (i) 14 verbs are treated as irregular and no further class information is given about them; (ii) 13 verbs are ambivalent about class membership, i.e., they can select the inflectional realizations of other classes for certain cells but not for all;⁸ and (iii) 89 verb entries in M&A are not entirely independent lexemes. Although the verbs are treated by M&A (p. 690) as 'objectless transitive verbs', Foris (2000) advances a superior analysis of their counterparts in Sochiapam Chinantec as being in inverse forms. In this way, for example, the verbal form *dsej¹²* 'pull out' in (25b) is the inverse form of the verb *DSEY¹²* 'pull out', of which one direct form is given in (25a).

- (25) a. *ca'-dsej¹²=jni* *chi¹² chieh¹²*
 PST-pull.out.ANIM/OBJ.DIRECT.PST.1SG/DIRECT=1SG CLF chicken
 'I pulled out the hen (from the box).' (M&A: 52)
- b. *ca'-dsej¹²=dsq²* *mi¹² quianh¹²=jni jni¹²*
 PST-pull.out.PST.3>1/INVERSE=3 HUMAN CLF.PL mate=1POSS I
 'My mates pulled me out (of work).' (M&A: 52)

I have taken all inverse forms, like the one in (25b), to realize the same lexemes whose direct entry is given elsewhere in the dictionary. For this paper, I have only analysed direct entries. A future study should aim to include inverse forms to understand their role in the making of conjugation classes. Removing thus the 122 verbs, the sample contains 790 verbs, as specified in (26).

(26) Simple	660
Bi-partite	130
Total	790

On the other hand, among these 790 verbs, there are many lexical pairs that could in principle be taken to represent the same lexeme, but realizing a different gender agreement. Verbs agree with their S or the O argument according to their animacy. Transitive verbs typically come in pairs: one entry which is used when the O argument is animate (e.g. *ciq^h¹²* 'eat an animate thing (a chicken, etc.)' and another when O is inanimate (e.g. *ceq^h¹²* 'eat an inanimate thing (beans, etc.)'. Often, these pairs display a similar inflectional behaviour, but not always. In this paper, I have treated them as representing two different lexemes for the purposes of inflection. A subtler and more sophisticated analysis of such pairs is still needed. For the present purposes, Table 1 shows that for inflectional purposes there is no substantial difference between the forms involved in the encoding of a third person subject, may it be inanimate or animate. Transitive verbs commonly have an animate subject.

⁸ Given the great amount of inflectional variation, the fact that only 1.4% of the verbs in this sample display ambivalent behavior regarding class membership is in itself a remarkable phenomenon.

Series for the 3 rd person	# of tr. verbs, mostly with an animate subject	# of intr. verbs with an animate subject	# of intr. verbs with an inanimate subject
1	205 42%	38 27%	53 33.3%
2	18 3.6%	17 12%	27 16.9%
3	16 3.2%	3 2.1%	4 2.5%
4	5 1%	5 3.5%	7 4.4%
5	3 0.7%	6 4.2%	4 2.5%
6	5 1%	3 2.1%	6 3.7%
7	-	3 2.1%	2 1.2%
8	1 0.2%	2 1.4%	2 1.2%
9	-	2 1.4%	1 0.9%
10	-	1 1%	-
I	166 34%	28 20%	33 20.7%
II	33 6.7%	17 12%	9 5.6%
III	25 5.1%	9 6.3%	8 5%
IV	9 1.8%	4 2.8%	2 1.2%
V	4 0.8%	3 2.1%	1 0.9%
	490 100%	141 100%	159 100%

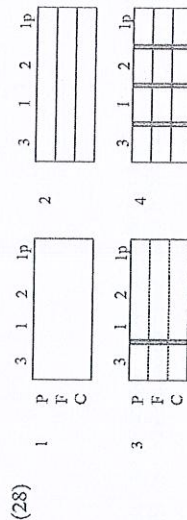
Table 1. Encoding of third person and type of subject (animate vs. inanimate).

5. CONJUGATION PATTERNS AND CONJUGATION CLASSES

We have seen that M&A already treat verbs in groups A, B and C attending to the number of grammatical persons we observe in their inflection. Following this tradition, but modifying it a little to accommodate further contrasts, I propose that the dynamic verbs of TCh inflect according to four conjugation patterns. A conjugation pattern is defined in (27).

(27) **Conjugation pattern.** A schematic template for the inflection of a dynamic verb which reveals the treatment of grammatical person.

The conjugation patterns represent in themselves schematic verbal classes in that verbal lexemes in TCh are categorized according to the type of pattern they require for their inflection. If a verb belongs to a certain pattern, this gives rise to a number of expectations about its inflectional behavior. The four conjugation patterns are given in (28).

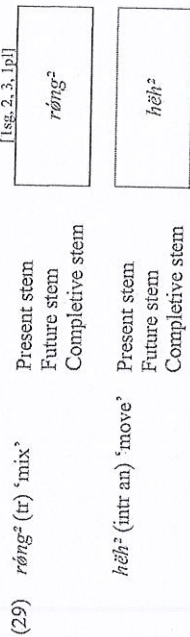


The conjugation patterns in (28) are templates that instantiate three possible stages of inflectional complexity with respect to the realization of person for TCh, and possibly for all other Chinantecan languages. Pattern 1 and 2 are 'simplex' because they are insensitive to person. Pattern 1 encapsulates uninflected verbs in the language. Pattern 2 has TAM

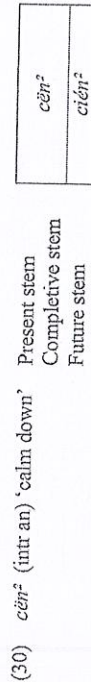
distinctions. In pattern 3, there is a first split in the paradigm between the third person and all other persons. There are various subdivisions of this pattern according to whether there are additional TAM distinctions. Finally, pattern 4 displays a maximal person distinction which is often accompanied by TAM distinctions as well, but even here there are further splits, which I will discuss briefly in §6. Except for pattern 1, all other patterns display inflectional variance in the form of conjugation classes. I discuss these patterns in the following sections together with the classes we find.

5.1. CONJUGATION PATTERN 1: UNINFLECTED VERBS

Verbs of conjugation pattern 1 have only one stem and are not inflected for person. Two examples are given in (29).



Although most of the verbal lexemes requiring conjugation pattern 1 for their inflection could be said to be outside the morphology for inflectional purposes, some of them like the verb *chh²* (intr an) 'calm down' in the example in (30) may have a distinctive future stem, produced by palatalization. The existence of such stems is an indication that at some deeper level, verbs of this type are sensitive to further TAM splits.



The distribution of the verbs with conjugation pattern 1 in the sample is given in (31). While the simple verbs in the sample having this pattern only amount to ten percent, the pattern appears to be specially favored by bi-partite verbs, as almost 30% of such verbs display the invariant conjugation pattern 1.

(31)

	Total	Pattern 1	%
Simple verbs	660	71	10.7%
Bi-partite verbs	130	38	29.2%
Total	790	109	13.8%

M&A treat verbs of this type as constituting the conjugation classes C3, D3, C4, D4, C7, D7, C9, D9, C13 and D13. Contrary to such a view, I take the only stem of these verbs as the lexical representation of the lexemes (in the sense by Aronoff 1978). In this new light, the prosodic features in the stem (e.g. whether tone and/or stress) should be seen as representing the segmental phonology of the lexeme. Nonetheless, the lexical representation (LR) of a verb of this conjugation pattern shows certain tendencies in its phonological structure, as shown in (32), where it can be seen that many of them have a middle tone plus ballistic stress.

third person. Verbs that inflect according to pattern 3 are simultaneously classified in the lexicon as belonging to patterns 1 or 2 for the purpose of encoding the third person. This implies that for a given verbal lexeme, speakers would have to remember the exponents for the third person independently from the exponents for other persons. When the verb belongs to different conjugation classes for this purpose, the lexeme needs to store this simultaneous membership for the information to be recoverable for inflectional purposes. Verbs can in turn cluster in groups attending to their membership. I call such groupings 'superordinate inflectional classes', and will be presented in the following sections.

5.3.1. *Verbs of sub-pattern 3a*. The verbs of this sub-pattern have the simplest structure as they only have two forms inflected for person: one is used for the third person; the other form is used to cross-reference all other grammatical persons. In other words, the alternative form cross-references a non-third person. Two examples appear in (38).

	3	non-3
(38)	giŋg ²	giŋg ³
giŋg ² (tr) 'dig'	Present stem Future stem Completive stem	Present stem Future stem Completive stem
quidh ² (tr an) 'choose'	quidh ²	quidh ³
	Present stem Future stem Completive stem	Present stem Future stem Completive stem

The form used for the third person is the default lexical representation of the verb. This means that the verbs following this sub-pattern are basically pattern 1 verbs, where the form of the lexical representation acquires the function of exponent of a third person value by virtue of the existence of the other form. The distribution of the verbs of this pattern in the sample is given in (39). Notice that more than half of the bi-partite verbs require conjugation pattern 3a.

(39)	Total	Pattern 3a	%
Simple verbs	660	53	8%
Bi-partite verbs	130	69	53%
Sample	790	122	15.4%

The verbs of this pattern fall into two classes emerging from the combination of the selection they make for third person and the class encoding non-third person. To differentiate these superordinate classes (S-class) from the other types, I note them using roman numbers in small caps (i.e., I, II, III, etc.).

(40)	S-class	3	non-3	Total	M&A	bi-part.			
			Class						
I	LR	a	B3h, B4h	95	77.7%	46	86.8%	49	71%
II	LR	b	B3g, B4g, B9g	27	22.3%	7	13.2%	20	29%
				122	100%	53	100%	69	100%

As expected, the phonological form of the lexical representation of these 122 verbs follows the same tendencies of the uninflected verbs of pattern 1. This is shown in (41), where the verbs of pattern 3a are compared with those of pattern 1.

(41)	Patr. 3a	%	Patr. 1
1	23	18.8%	11.0%
+2	74	60.6%	42.2%
12	24	19.7%	16.5%
13	1	0.9%	
Total	122	100%	

If a verb has a given /+2/ structure in their lexical default form (i.e., their LR), the language user has to remember whether it further belongs to conjugation pattern 1 or 3a, or to other patterns for that matter, as we will soon see. Once the verb is recognized as belonging to a given pattern, NOTHING in the phonology of the LR will inform the language user about what other forms the verb will take. For example, if a verb is stored as inflecting according to pattern 3a, nothing in the phonology of the default form, used for the third person, will inform one what forms the non-third person will take.

5.3.2. *Verbs of sub-pattern 3b*. A very few verbs have this pattern. Such verbs have one stem for the third person but three different stems for the non-third person. As with verbs of pattern 3a, I assume here that the stem for the third person is the LR of the verb. This sub-pattern illustrated in (42).

(42)	jénh ² (intr an) 'have hiccups'	3	non-3
		jénh ²	jénh ² jénh ³ jénh ²
		Present stem Future stem Completive stem	Present stem Future stem Completive stem
	quioŋ ² (intr an sg) 'lie'	quioŋ ²	quioŋ ² quioŋ ³ quioŋ ²
		Present stem Future stem Completive stem	Present stem Future stem Completive stem

In the sample, there are only 5 verbs that display pattern 3b.

(43)	Simple verbs	2
	Bi-partite verbs	3
	Total	5

However, despite being a few verbs, the pattern has three conjugation classes, as shown in (44), two of them populated by only one member each.

(44)	S-class	3	non-3	3	non-3	Total
	III	LR	d	P	F	C
	IV (L2)	LR	e	2	+3	2
	V (L3)	LR	f	12	13	13
				12	13	1
						1
						5

Again, the lexical representation of these five verbs conforms to the general trend shown by patterns 1 and 3a: two verbs have /+2/ which occurs in the 42.2% of verbs of pattern 1 and in 60.6% of pattern 3, and the three remaining verbs have /12/, which occurs in 16.5% of verbs of pattern 1 and 19.7% of pattern 3a.

5.3.3. *Verbs of sub-pattern 3c*. This inflectional pattern is the mirror image of pattern 3b. It consists of verbs with three stems for the third person and only one for the non-third person, as shown in (45). 88 verbs in the sample inflect according to this pattern.

(45)

<i>juen¹²</i> (tr) 'break'	Present stem	3 rd	non-3 rd
	Future stem	<i>juen¹²</i>	<i>juen¹³</i>
	Completive stem	<i>juen¹</i>	
<i>iaj¹²</i> (tr an) 'feed'	Present stem	<i>iaj¹²</i>	<i>iaj¹³</i>
	Future stem	<i>iaj¹</i>	
	Completive stem	<i>iaj¹</i>	

The distribution of pattern 3c is given in (46).

(46)

	Total	Pattern 3c	%
Simple verbs	660	80	12.1%
Bi-partite verbs	130	1	0.7%
Sample	790	81	10.2%

Verbs following pattern 3c fall in six major conjugation classes, which emerge when we put together how the two grammatical persons are encoded. In reality, only two classes contain most verbs.

(47)

S-class	3	non-3	Class	P	F	C	M&A	Total	%
VI	1	a	1/2	1	1	1	B1h	42	51.9%
VII	1	b	1/2	1	1	1/3	B1g	33	40.1%
VIII (L4)	1	c	1/2	1	1	1	B1f	1	1.2%
IX (L5)	2	a	-2	1	-2	-3	B5h	1	1.2%
X	3	a	1/2	1	+2	-3	B2h	2	2.5%
XI	4	a	2	1	1	-3	B8h	2	2.5%
								81	100%

5.3.4. *Verbs of sub-pattern 3d*. The last sub-pattern is 3d and it is only found in two verbs of the sample. Such verbs have three stems for the third person and three other for the non-third person. The two verbs in question are given in (48), and by themselves they instantiate the two different classes in (49).

(48)

L6	<i>huan²</i> (intr an pl) 'exit'	Present stem	3	non-3
		Future stem	<i>huan²</i>	<i>huan²</i>
		Completive stem	<i>huan²</i>	
L7	<i>jog¹²</i> (intr an sg) 'come again'	Present stem	<i>jog¹²</i>	<i>jog¹²</i>
		Future stem	<i>jog²</i>	<i>jog¹³</i>
		Completive stem	<i>jog¹</i>	<i>jog¹</i>

(49)

S-class	3	non-3	Class	P	F	C	P	F	C	M&A	
XII (L6)	2	d	-2	1	-2	2	3	2	3	2	B5b
XIII (L7)	8	f	12	2	1	12	13	1	13	1	B6a

5.4. CONJUGATION PATTERN 4: MAXIMALLY INFLECTED VERBS

Verbs following this conjugation pattern can be maximally inflected. They commonly have three stems for each grammatical person: third person (3), first person singular (1sg), second person (2), and first person plural (1pl), as illustrated by the two verbs in (10), repeated here as (50). There is a great degree of syncretism among the stems involved.

(50)

<i>jih²</i> (tr) 'bend'	Present stem	3	1sg	2	1pl
	Future stem	<i>jih²</i>	<i>jih¹²</i>	<i>jih¹²</i>	<i>jih¹²</i>
	Completive stem	<i>jih²</i>	<i>jih¹³</i>	<i>jih¹³</i>	<i>jih¹³</i>
<i>tih¹²</i> (tr) 'cut'	Present stem	<i>tih¹²</i>	<i>tih¹²</i>	<i>tih²</i>	<i>tih²</i>
	Future stem	<i>tih¹</i>	<i>tih¹³</i>	<i>tih¹³</i>	<i>tih¹³</i>
	Completive stem	<i>tih¹</i>	<i>tih¹</i>	<i>tih³²</i>	<i>tih³</i>

There are 287 verbs of this conjugation pattern in the sample of 790 verbs (36.4%). These verbs display a great amount of inflectional variance, which could be organized in 29 different classes only attending to the allomorphy for the encoding of the first singular, the second person, and the first person plural. Such classes appear in (51).

(51)

Class	Isg	2	Ipl	3	Isg	2	Ipl	M&A
	Ser.	Ser.	Ser.	Ser.	P	F	C	
i	5	A	b	1	12	3	1	A13b
ii	5	B	b	1	12	3	2	A42b
iii*	5	5	b	1	12	3	1	A51b
iv*	5	5	b	1	12	3	1	A31a, A41a
v	5	D	D	1	12	13	13	A34a
vi	5	E	D	1	12	13	12	A35a, A45a
vii	5	F	D	1	2	3	-12	A47a
viii	5	G	D	1	12	13	3	A18a
ix	a	C	D	+2	+3	+2	2	A46d
x	a	E	D	+2	+3	+2	2	A45d
xi	a	B	D	+2	+3	+2	2	A32d, A42d
xii	a	B	c	+2	+3	+2	2	A32f
xiii	a	B	A	+2	+3	+2	2	A32e, A42c
xiv	a	C	A	+2	+3	+2	2	A16c
xv	a	5	A	+2	+3	+2	2	A51c
xvi (L8)	5	B	D	12	13	1	2	12a
xvii (L9)	5	C	D	12	13	1	+2	12a
xviii (L10)	5	C	b	12	13	1	+2	16b
xix (L11)	5	V	b	12	13	1	-3	49b
xx (L12)	a	S	c	+2	+3	+2	12	13
xxi (L13)	a	A	c	+2	+3	+2	2	13
xxii (L14)	a	C	c	+2	+3	+2	2	13
xxiii (L15)	a	E	c	+2	+3	+2	12	13

xv (L16)	a	F	A	+2	+3	+2	2	+3	-2	2	+3	-3	17c
xvi (L17)	a	G	D	+2	+3	+2	12	13	-3	12	13	13	18d
xvii (L18)	V	V	D	-3	-3	-3	-3	-3	-3	12	13	13	19j
xviii (L19)	II	V	IV	-3	-3	-3	-3	-3	-3	13	13	13	99g, 19g
xvix (L20)	I	A	V	+2	+2	+2	2	+3	-3	13	13	13	79i
													33h

*In these two classes, the default exponents for the first person singular have been leveled up to realize the second person.

These conjugation classes are not populated evenly. Some have many members, like classes *i* and *iv*, while most have only a handful. This is shown in (52).

(52)

	Total	simple	bi-part.	Total	simple	bi-part.
<i>i</i>	65	63	2	1	1	1
<i>ii</i>	25	25	0	1	1	1
<i>iii</i>	11	11	0	1	1	1
<i>iv</i>	90	83	7	1	1	1
<i>v</i>	15	15	0	1	1	1
<i>vi</i>	8	8	0	1	1	1
<i>vii</i>	9	9	0	1	1	1
<i>viii</i>	6	6	0	1	1	1
<i>ix</i>	11	10	1	1	1	1
<i>x</i>	9	9	0	1	1	1
<i>xi</i>	5	5	0	1	1	1
<i>xii</i>	3	3	0	1	1	1
<i>xiii</i>	8	8	0	3	2	1
<i>xiv</i>	4	4	0	1	1	1
<i>xv</i>	2	2	0	287	273	14

Like verbs of pattern 3, verbs of pattern 4 belong simultaneously to either pattern 1 or 2 to realize the third person. Of the 287 verbs of pattern 4, 115 (40%) behave like pattern 1 and 172 (60%) like pattern 2. The 172 verbs that are like pattern 2 belong in turn to class 1 or class 2. In (53), I compare the behavior of such verbs with those of the TAM inflected verbs of pattern 2.

(53)

Class	Series	Pattern 2 verbs			Pattern 4 verbs with 3 rd person like pattern 2			
		P	F	C	Total	%	Total	%
1	1	12	1	1	84	45.6%	136	79.0%
2	2	-2	1	-2	46	25.0%	15	8.7%
3	3	12	1	+2	10	5.4%	16	9.3%
4	4	2	1	1	5	2.7%	1	1.3%
5	5	12	13	1	14	7.6%	1	1.7%
6	6	2	2	1	13	7.0%	3	1.7%
7	7	-3	1	-3	5	2.7%	1	1.6%
8	8	12	2	1	3	1.6%	1	1.3%
9	9	12	1	12	3	1.6%	1	1.3%
10	10	2	13	1	1	0.8%	1	0.8%
					184	100%	172	100%

In principle, for the encoding of the third person, knowing whether a verb of pattern 4 is of class 1 or it just requires the LR already accounts for 80% of all 287 verbs that follow this pattern.

This suggests that there is some degree of predictability in the system. The actual crisscrossing in the combinations for the encoding of all persons appear in (54), resulting in 45 other conjugation classes, which together with the 13 other displayed by pattern 3 verbs in §5.3 amount to a gross total of 58 superordinate inflectional classes.

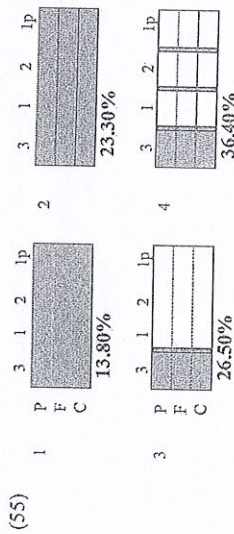
(54)

S-Classes	3	1sg, 2, 1pl	#	M&A
XIV	<i>i</i>	<i>i</i>	64	A13b
XV (L21)	LR	<i>i</i>	1	A43b
XVI	LR	<i>ii</i>	5	A42b
XVII	3	<i>ii</i>	12	A22b
XVIII	<i>i</i>	<i>ii</i>	7	A12b
XIX (L22)	3	<i>ii</i>	1	A82b
XX	2	<i>iii</i>	11	A51b
XXI	LR	<i>iv</i>	70	A31a, 41a
XXII	<i>i</i>	<i>iv</i>	16	A11a
XXIII	3	<i>iv</i>	3	A21a
XXIV (L23)	6	<i>iv</i>	1	A61a
XXV	LR	<i>v</i>	3	A34a
XXVI	<i>i</i>	<i>v</i>	12	A14a
XXVII	LR	<i>w</i>	3	A35a, 45a
XXVIII	<i>i</i>	<i>w</i>	2	A15a
XXIX	6	<i>w</i>	3	A63a
XXX	LR	<i>wi</i>	2	A47a
XXXI	<i>i</i>	<i>wi</i>	6	A17a
XXXII (L24)	3	<i>wi</i>	1	A27a
XXXIII	<i>i</i>	<i>wii</i>	6	A18a
XXXIV	LR	<i>ix</i>	7	A46d
XXXV	<i>i</i>	<i>ix</i>	4	A16d
XXXVI	LR	<i>x</i>	2	A45d
XXXVII	<i>i</i>	<i>x</i>	7	A15d
XXXVIII	LR	<i>w</i>	5	A32d, 42d
XXXIX	LR	<i>wi</i>	3	A32f
XI.	LR	<i>wii</i>	8	A32c, 42e
XII	<i>i</i>	<i>xv</i>	2	A16c
XIII	LR	<i>xv</i>	2	A46c
XIIII	2	<i>xv</i>	2	A51c
XLV (L8)	<i>i</i>	<i>xv</i> (L8)	1	L2a
XLV (L9)	2	<i>xvii</i> (L9)	1	56a
XLVI (L10)	<i>i</i>	<i>xviii</i> (L10)	1	16b
XLVII (L11)	LR	<i>xix</i> (L11)	1	49b
XLVIII (L12)	2	<i>xx</i> (L12)	1	51f
XLIX (L13)	<i>i</i>	<i>xxi</i> (L13)	1	13f
L (L14)	<i>i</i>	<i>xxii</i> (L14)	1	16f
LI (L15)	<i>i</i>	<i>xxiii</i> (L15)	1	15f
LII (L16)	<i>i</i>	<i>xiv</i> (L16)	1	17c
LIII (L17)	<i>i</i>	<i>xv</i> (L17)	1	18d
LIV (L18)	<i>i</i>	<i>xvi</i> (L18)	1	19j
LV	<i>i</i>	<i>xvii</i>	2	A19g
LVI (L25)	LR	<i>xviii</i>	1	99g
LVII (L19)	LR	<i>xviii</i> (L19)	1	79i
LVIII (L20)	LR	<i>xvix</i> (L20)	1	33h

6. SUMMARY AND CONCLUSIONS

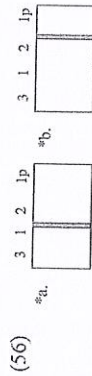
I have proposed that dynamic verbs in Tch fall into five different conjugation patterns. A number of conclusions can be advanced regarding the different ways grammatical person is encoded in the inflection of dynamic verbs in a Chinantecan language. The different available splits presented are repeated again in (55), together with their overall frequency. The areas in

shading in patterns 3 and 4 correspond to the zones in the paradigms where we find the defaults of patterns 1 and 2, here serving as exponents of third person.

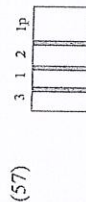


Patterns 1 and 2, with no person distinction, amount to 37.1% of all verbs in the sample. Pattern 3 represents a primitive person split, creating a distinction between a third person and a non-third person, and accounts for more than a quarter of the verbs in the sample (26.5%). The rest, with almost equal weight as the verbs of patterns 1 and 2 together, are the verbs of pattern 4 with the more personal distinctions.

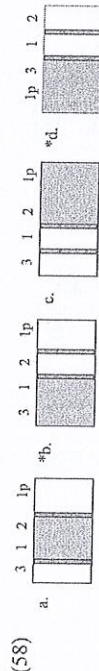
On the other hand, the type of person split in pattern 3 involves a distinction between a third person and all other persons. For example, the potential splits in (56), which involve the merging of third person with other persons, do NOT exist in Tch. This is in support of the analysis we have advanced that the encoding of the third person in conjugation patterns 3 and 4 verbs responds to independent criteria different from those applying to other persons.



Furthermore, when by virtue of further splits other persons emerge beyond the two primitive persons in pattern 3, it is common to maintain a full contrast among all persons like in (57). This happens in 58.2% of all verbs of the maximally inflected verbs of pattern 5 (167/287).



Another possibility is to have a person split according to the four options in (58). Collapse in person marking is indicated by the shading.



Of the possibilities in (58), only (a) and (c) occur, not (b) or (d). Pattern (a) is common, as it occurs in 36.5% of verbs of pattern 5 (105/287). The 15 remaining verbs show the split in (c). These sub-patterns could be interpreted as the result of morphological levelling: the inflectional series realizing the second person is commonly used for the first person elsewhere, so the direction of the levelling is treating the second person as a first person singular. For pattern (c), the second person is also levelled out to the first plural. In contrast, the prosodic exponents involved in the realization of the non-third person are specifically devoted to this person, and do not emerge from the levelling of other persons (see 24 above).

I have proposed that the encoding for the third person combines with the encoding of other persons to create the different conjugation classes at a higher level of the inflectional system. I take the sample of 790 verbs as representative of the verbal lexicon of the language involving dynamic verbs. With this in mind, there are about 14% of verbs that do not have inflected stems for neither TAM nor person (pattern 1) and another 23% that only inflect for TAM, and such verbs fall into at least 10 different classes, two of them with the largest number of members. The rest of the verbs in the lexicon are sensitive to realize a person value (patterns 3 and 4). To encode a third person value, all such verbs are either characterized as belonging to pattern 1, using a lexical default, or to pattern 2, and they thus need to belong to at least six of the ten classes available for verbs of pattern 2 (class 1 is the default class). The very same verbs will inflect for other persons: a non-third person if they inflect according to pattern 3 and all other persons if they inflect like pattern 4. To realize a non-third person, such verbs fall into 6 different classes (a-f). Those that inflect according to pattern 4 are distributed in 29 classes (i-xvz). For inflectional purposes, a given verb of pattern 4, for example, may pertain to class 1 to encode the third person and to class iv to encode other persons. However, speakers not only remember these combinations, but they have clustered verbs in groups according to the different classes they belong to. For example, there is no verb in the sample that belongs to class 4 and to class i. Only a number of combinations are possible. I have treated such possible combinations as superordinate conjugation classes, and they amount to 58 (I-LVIII).

Graph 1 in the appendix shows the four most frequent combinations resulting in classes such as i, ii, xiv and xxi, which have more than 30 members each. At the other extreme, we have combinations which are hardly entrenched at all. They are only instantiated by one lexeme in the sample, and they amount to 25 of them. These are given in Graph 2. Graph 3 shows the possible combinations of classes whose class membership ranks from two members to less than 30.

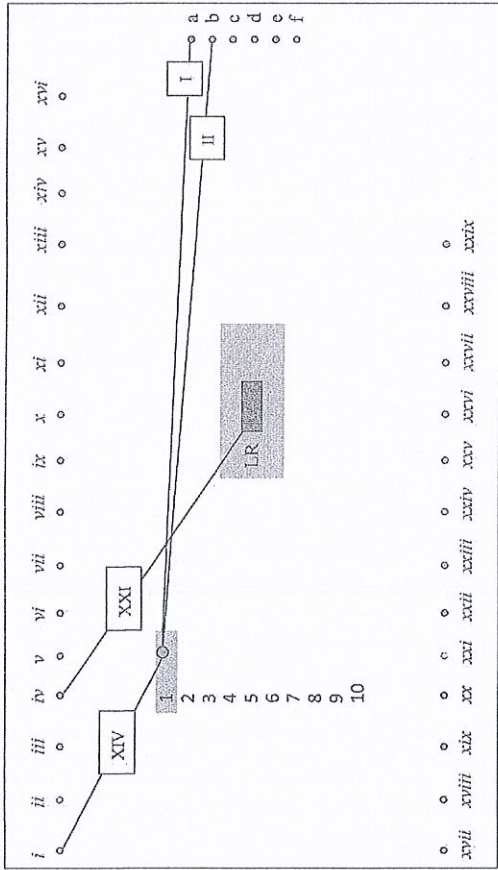
I hope to have shown that this new approach to the complexities of the verbal inflection of a Chinantecan language reveals a number of novel patterns of organization of the system which have not previously been detected. Such patterns are likely to be common to all the languages of the family.

REFERENCES

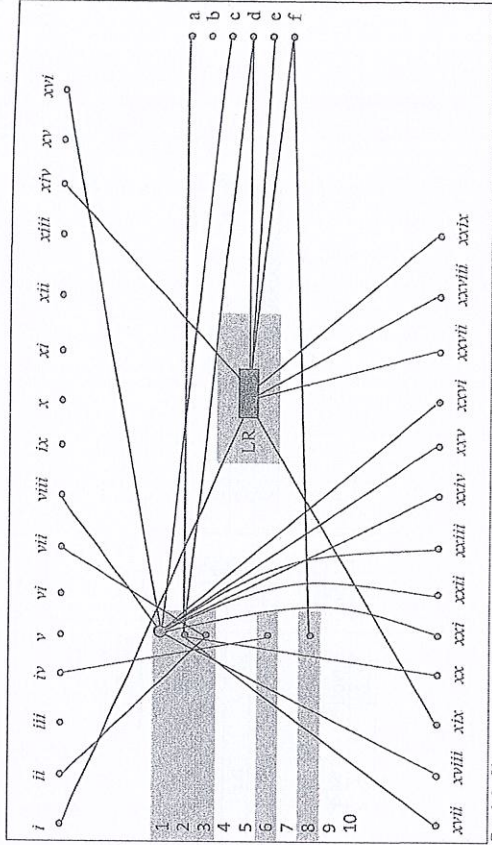
- Aronoff, Mark. 1978. Lexical representations, in D. Farkas, W. M. Jacobsen, and K. W. Todrys (eds.), *Papers on the Parasession on the Lexicon*, 12-25. Chicago: Chicago Linguistics Society.
- 1994. *Morphology by Itself: Stems and Inflectional Classes* (Linguistic Inquiry Monograph 22). Cambridge, MA: MIT Press.
- Finkel, Raphael and Gregory Stump. 2009. Principal parts and degrees of paradigmatic transparency. *Analogy in Grammar: Form and Acquisition*, ed. by James P. Blevins and Juliette Blevins, pp. 13-53. Oxford: Oxford University Press.
- Foris, David Paul. 2000. *A grammar of Sochiapan Chinantec*. Studies in Chinantec languages 6. Dallas, TX: SIL International and The University of Texas at Arlington.
- Merrifield, William R. 1968. *Palantla Chinantec grammar*. *Papeles de la Chinantla* 5, Serie Científica 9. México: Museo Nacional de Antropología.
- Merrifield, William R. and Alfred E. Anderson. 2007. *Diccionario Chinanteco de la diáspora del pueblo antiguo de San Pedro Tlatepuzco, Oaxaca*. [2nd Edition]. Serie de vocabularios y diccionarios indígenas "Mariano Silva y Aceves" 39. México DF: Summer Linguistics Institute. Accessible at: <<http://www.sil.org/mexico/chinanteca/tlatepuzco/S039a-DiccChinTlatepuzco-cpa.htm>>.
- Merrifield, William and Calvin R. Rensch. 1990. *Syllables, Tone, and Verb Paradigms* [Studies in Chinantec Languages 4]. Summer Institute of Linguistics and The University of Texas at Arlington.
- Mugele, R. L. 1982. *Tone and Ballistic Syllables in Lalana Chinantec*. Ph.D. dissertation.

Austin: University of Texas.
 Rensch, Calvin. 1978. Ballistic and controlled syllables in Otomanguean Languages, in Alan Bell and Joan B. Hooper (eds.), *Syllables and Segments*, pp. 85-92. Amsterdam: North Holland Publishing Company.
 Silverman, Daniel. 1994. A Case Study in Acoustic Transparency: [spread glottis] and Tone in Chinantec. *Proceedings of NELS (Amherst: University of Massachusetts)* 24.
 Stump, Gregory and Raphael Finkel. 2007. Principal parts and morphological typology, *Morphology* 17, 39-75.

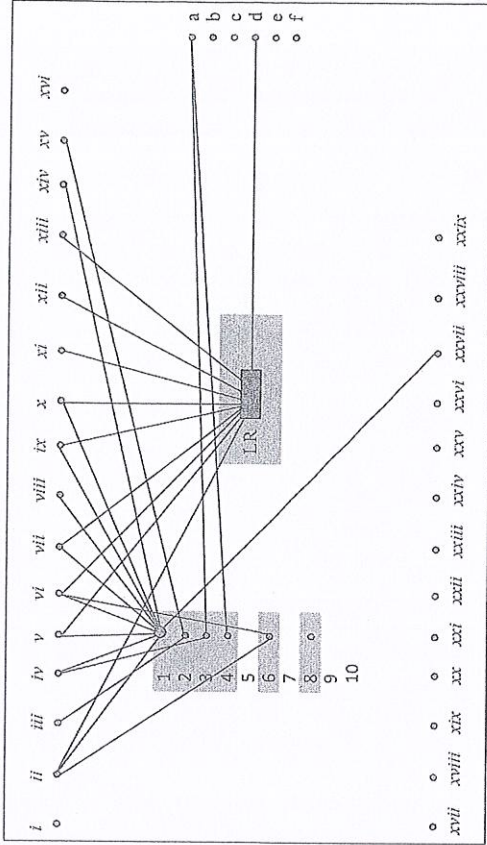
Appendix 1: Superordinate conjugation classes.



Graph 1. Classes with more than 30 members.



Graph 2. Classes with just one member.



Graph 3. Classes with more than one member and less than 30.

Appendix 2: Inflection of bi-partite verbs.

In the inflection of a bi-partite verb like (1), both morphs have inflectional properties. The first morph inflects according to the fixed pattern in (1a), while the second, the morphological root, may have inflectional forms similar to simple verbs, as illustrated in (1b).

(1) $h\theta^2+ts\acute{o}h^{12}$ (tr) 'stump with foot'

$h\theta^2+ts\acute{o}h^{12}$	$h\theta^2+ts\acute{o}h^{12}$	$h\theta^2+ts\acute{o}h^{12}$	$h\theta^2+ts\acute{o}h^{12}$
$h\theta^2+ts\acute{o}h^1$	$h\theta^2+ts\acute{o}h^{13}$	$h\theta^2+ts\acute{o}h^3$	$h\theta^2+ts\acute{o}h^3$
$h\theta^2+ts\acute{o}h^1$	$h\theta^2+ts\acute{o}h^1$	$h\theta^2+ts\acute{o}h^1$	$h\theta^2+ts\acute{o}h^1$

Present stem
Future stem
Completive stem

	3	1sg	2	1pl	3	1sg	2	1pl
Present stem	2	2	2	2	$h\theta^2+$	$h\theta^2+$	$h\theta^2+$	$h\theta^2+$
Future stem	1	3	3	3	$h\theta^1+$	$h\theta^1+$	$h\theta^1+$	$h\theta^1+$
Completive stem	1	3	3	3	$h\theta^1+$	$h\theta^1+$	$h\theta^1+$	$h\theta^1+$
Present stem	12	12	2	2	$ts\acute{o}h^{12}$	$ts\acute{o}h^{12}$	$ts\acute{o}h^{12}$	$ts\acute{o}h^{12}$
Future stem	1	13	+3	+3	$ts\acute{o}h^1$	$ts\acute{o}h^{13}$	$ts\acute{o}h^3$	$ts\acute{o}h^3$
Completive stem	1	1	-3	+3	$ts\acute{o}h^1$	$ts\acute{o}h^1$	$ts\acute{o}h^3$	$ts\acute{o}h^3$

a.

b.

The conjugation classes of Tilapa Otomi: An approach from canonical typology

ENRIQUE L. PALANCAR

*A Thomas C. Smith-Stark[†]
en honor al noble vikingo en él
y a su pasión por Mesoamérica*

Abstract

From the seminal work by Aronoff (1994), the study of inflectional classes has become of interest in morphological theory. Most of our current knowledge of verb inflectional classes, also known as “conjugation classes”, remains circumscribed to European languages. On the other hand, the Oto-Manguean languages of Mexico have recently awoken interest because of their puzzling internal diversity and morphological abundance, but much of their morphological complexities remain poorly understood because of the lack of comprehensive materials. To increase our understanding of such classes, in this article I present an analysis of the three conjugation classes of Tilapa Otomi, an Otomian language close to extinction which belongs to the Oto-Pamean branch of Oto-Manguean. In the article, I also provide an example of how the canonicity of conjugation classes can be ranked following the approach advocated by Corbett (2009). The results of such an evaluation show that while two of the three classes of Tilapa Otomi are canonical to a certain extent, a third one is a very poor example of an inflectional class for the purposes of typological comparison.

1. Introduction¹

It is well known that languages vary in how they express the same inflectional information in different words. While in English, the phrases *she sings*, *she wants* or *she opens* all involve the addition of the same ending *-s* to the verb to express that there is a third person singular subject, in Spanish there are two different endings in the corresponding forms, depending on which verb it is: *cant-a* ‘sings’ but *quier-e* ‘wants’ and *abr-e* ‘opens’. In some instances there are even three different endings; compare *cant-amos* ‘we sing’, *quer-emos* ‘we want’ and *abr-imos* ‘we open’. As Spanish verbs differ in the way they encode grammatical categories of tense and the person of the subject, these simple data

illustrate the common knowledge that Spanish verbs fall into at least three different *inflectional classes*, defined in Aronoff (1994: 64) as “a set of lexemes whose members each select the same set of inflectional realizations”.

Although it is generally agreed that the function of inflectional morphology is to convey grammatical meaning (Anderson 1992), the existence of inflectional classes presents a challenge to this view, as they do not have any identifiable morphosyntactic function, that is, they are *morphomic* (Aronoff 1994). And while some languages do without them completely, inflectional classes are far from rare, and demonstrate a remarkable resilience over time (see Maiden 1992, 2005 for examples in Romance languages).

However, our knowledge of inflectional classes to date is largely based on European languages, and is consequently limited by their typological characteristics. A comprehensive theory of inflectional classes must expand its horizons beyond this, but to do so, there is a need for high quality linguistic materials on lesser known languages. The conjugation classes of the Oto-Manguean languages from Mexico have recently received some attention in morphology because of their puzzling internal diversity and morphological abundance (see for example Finkel and Stump [2009] and Stump and Finkel [2007] for Comaltepec Chinantec, which are based on Pace [1990]), but all in all there are still very few studies that treat inflectional classes comprehensively, notable in this respect are the work by Angulo (1932) on Chichimeco and by Smith-Stark (2002) on Chichicapan Zapotec. With the goal in mind of expanding our knowledge and understanding of verb inflectional classes in Oto-Manguean, in this article I present a thorough analysis of the conjugation classes of Tilapa Otomi, an Otomian language of the Oto-Pamean branch which is on the verge of extinction.

The present article has three main goals: One goal is to present the inflectional properties of the three conjugation classes of Tilapa Otomi in detail because the classes are based on formal contrasts which are of interest for a theory of morphology. The interest lies in the fact that the contrasts go beyond the complexities posited by classes realized by inflectional endings (or thematic vowels) as in the well-known case of Indo-European languages (mainly Italic), also observed in Algic Yurok (Blevins 2005), while they keep some behavioral similarities with the *binyanim* of Semitic languages (Aronoff 1994: Chapter 5). A second goal is to advance generalizations concerning how class membership is assigned in the lexicon of Tilapa Otomi verbs. As I will show, the assignment displays unique characteristics not to be expected in inflectional classes. The third goal of the article is to introduce how the profile of a conjugation class can be evaluated following the canonical approach advocated by Corbett (2005, 2006, 2007). Such an approach serves as a useful *tertium comparationis* to understand the typological relevance of inflectional classes across languages against a theoretical space of structural possibilities. In this respect,

I advance a simple way in which the qualitative criteria used in canonical typology can be measured in quantitative terms.

To achieve these goals, I have organized the article as follows. In the next section, I briefly introduce the basic notions of canonical typology together with the criteria introduced in Corbett (2009) to define canonical inflectional classes. This is intended as a way to introduce the theoretical background from which to achieve an evaluation of the canonicity of the conjugation classes of Tilapa Otomi. An introduction to the language and its different grammatical tenses is given in Section 3. Subsequently, in Section 4 I present the basic inflectional contrasts characterizing the three conjugation classes, and in Section 5 I discuss the full paradigms in greater detail. In Section 6, I study the various factors that contribute to the assignment of class membership, especially for two of the minor conjugation classes. The evaluation of the degree of canonicity of the three conjugation classes is given in Section 7. Final remarks follow in Section 8, along with a brief summary of the proposal.

2. Canonical inflectional classes

In this article, I evaluate the conjugation classes of Tilapa Otomi against the reference background provided by Corbett's (2009) approach to inflectional classes from a canonical typology perspective. Canonical typology has been previously applied to the syntax of agreement in Corbett (2006), Comrie (2003), Evans (2003), and Polinsky (2003), and in morphological studies in Baerman et al. (2005: 27–35), Spencer (2005), Stump (2005), and Thornton (2010, 2011).

A canonical approach to language phenomena creates a space of theoretical possibilities defined by principles and criteria. This space of possibilities forms an ideal category or an ideal process, which may be called the “canon”, against which the actual constructs of a language instantiating that particular phenomenon can be evaluated. The canon is a logical space, and as Corbett (2007: 9) puts it: “[t]he canonical instances, that is, the best, clearest, indisputable (the ones closely matching the canon), are unlikely to be frequent. Rather, they are likely to be rare or even nonexistent.” In this light, the canon is not something good or bad, or a prototype, for that matter, or something that one language lacks but the next has, it is just a logical space. Accordingly, if one structure is more canonical than another, this should be taken to mean just that, that is to say, it does not mean it is more common, more frequent, or more normal in any sense.

The logical space of the canon is defined by common sense principles that are based on our general understanding of what the best scenario for such a situation should be. In Corbett's (2009: 1) own words: “[t]his means that we extrapolate from what there is to what there might be, in order to define the

theoretical space. Within that scheme of theoretical possibilities we can situate the real instances we find. An effect of this approach is to separate out coincidental overlaps in the examples that exist; we may then start to ask which characteristics *happen* to be the way they are and which *have* to be the way they are”. In this way, the canon becomes a useful *tertium comparationis* to understand the degree of divergence in natural languages, establishing a potential sound base for typological classification.

Corbett (2009) is a first approach to canonical inflectional classes. In this proposal, the canon is ruled by two principles: one maximizes its formal distinctiveness with respect to other classes; the other the independence of its members from the members of other classes. A number of criteria with different scope are postulated as manifestations of these principles. However, real linguistic instances are often far from the canon. Such instances can be accounted for as representing a weakening of the criteria in question.

The formal principle states that the “classes are fully comparable and are distinguished as clearly as is possible” (Corbett 2009: 3) and is manifested by the following four possible criteria in (1), (Corbett 2009: 3–4).

- (1) Criterion 1: “In the canonical situation, forms differ as consistently as possible across inflectional classes, cell by cell.”
 Criterion 2: “Canonical inflectional classes realize the same morphosyntactic or morphosemantic distinctions.”
 Criterion 3: “Within a canonical inflectional class each member behaves identically.”
 Criterion 4: “Within a canonical inflectional class each paradigm cell is of equal status.”

If we attend to criteria 1 and 4, an example of canonical inflectional classes is shown in Table 1 from Burmeso, a Papuan language of the East Bird’s Head-Sentani subgroup. Burmeso verbs fall into two classes according to what paradigm of gender prefixes they select to cross-reference its subject (there are six gender values), (Donohue 2001: 100, 102; in addition to Corbett 2009: 9).

Table 1. *The two verbal classes of Burmeso*

Gender values		Class I		Class II	
		SG	PL	SG	PL
I	male	<i>j-</i>	<i>s-</i>	<i>b-</i>	<i>t-</i>
II	female, animate	<i>g-</i>	<i>s-</i>	<i>n-</i>	<i>t-</i>
III	miscellaneous	<i>g-</i>	<i>j-</i>	<i>n-</i>	<i>b-</i>
IV	mass nouns	<i>j-</i>	<i>j-</i>	<i>b-</i>	<i>b-</i>
V	banana, sago tree	<i>j-</i>	<i>g-</i>	<i>b-</i>	<i>n-</i>
VI	arrows, coconuts	<i>g-</i>	<i>g-</i>	<i>n-</i>	<i>n-</i>

As for criterion 1, all the prefixes in Table 1 differ in form consistently. Weakening of the criterion occurs when there are shared or default forms in some of the cells, but here the markers are all different, that is to say, no marker of class I is used for class II, and vice versa. Besides, although there is some degree of syncretism (for example, the prefix for gender classes I and II is the same in the plural), this syncretism follows a pattern that is mirrored in both classes. As for criterion 4, each paradigm cell in Table 1 appears to be of equal status too, and predictions can be made based on formal implications: e.g., if *g-* is used to cross-reference the singular value of a noun of gender class II, then *g-* will be also used for the singular of class III and VI, and the plural of V and VI. In exactly the same contexts, *n-* will also be used, but for class II verbs.

In the canonical situation, according to criterion 2, the classes should realize the same grammatical distinctions. This creates the expectation that the system would make it possible for all members of an inflectional class to have identical paradigm shapes. In reality, a number of members in a given class commonly end up having a reduced paradigm because of semantic restrictions or lexical defectivity. If the restrictions are not motivated because of their class membership, they should not be an ascribable property to the class. Criterion 3 dictates that within a canonical inflectional class each member behaves identically, that is there are no morphophonological alternations, and no stem alternations. If such alternations exist, but are predictable by other rules (i.e., regular), the instance is more canonical than when they should be listed. In reality, the actual instances of natural languages are far from being canonical in this respect, but this mere fact highlights precisely how interesting real types can be.

The independence principle rules class membership and provides that “the distribution of lexical items over canonical inflectional classes is synchronically unmotivated” (Corbett 2009: 5). This is justified because canonical inflectional classes are morphology internal, that is, they are not motivated by other principles, semantic or otherwise. If that were the case, there would be no motivation to treat them as inflectional classes. There are five criteria realizing this principle, which are given in (2), (Corbett 2009: 6–7).

- (2) Criterion 5: “The larger the number of members of an inflectional class (up to an equal ‘share’ of the available items) the more canonical that class.”
- Criterion 6: “In the canonical situation, the distribution of lexical items over inflectional classes is not phonologically motivated.”
- Criterion 7: “In the canonical situation, the distribution of lexical items over inflectional classes is not syntactically motivated.”

- Criterion 8: “In the canonical situation, the distribution of lexical items over inflectional classes is not motivated by Part of Speech”.
- Criterion 9: “In the canonical situation, the distribution of lexical items over inflectional classes is not motivated by pragmatics (including information structure).”

Criterion 5 provides the canonical profile of the lexical distribution of canonical classes. To be canonical, all classes should have the same number of members. There is no reason why this should not be the case in the logical definition. In reality, languages organize their lexicons far from the canon. The common case is that there is one major class that serves as the lexical default. On the other end, in a system with many inflectional classes, it is also common to find thinly populated ones.

In the canonical situation, the assignment of a particular lexeme to a particular class is unmotivated because, according to the canon, an inflectional class is a genuine morphological ontology. However, weakening of the criteria applies precisely when a given property associated with the lexeme (phonetic, syntactic, word class or pragmatic) serves as a token signaling its class membership. To Corbett’s criteria in (2), I add the criterion in (3).

- (3) Criterion 10: In the canonical situation, the distribution of lexical items over inflectional classes is not semantically motivated.

In the canonical situation, the morphology involved in the making of inflectional classes is not assigned any semantic function. In other words, a canonical inflectional class is realized by means of morphological exponence that does not encode semantic categories such as agentivity, stativity, animacy, size, volume, shape, semantic gender, etc. The real instances of inflectional classes across the world’s languages vary in their degree of canonicity with respect to criterion 10. Most commonly this degree varies according to the feature in question and to the type of word class. Typologically, noun classes are less canonical than verbs. In nouns, for example, an inflectional class with many members of a particular semantic gender would rank low in canonicity, as would one that includes only inanimate or only animate nouns. On the other hand, it is more common to find canonical inflectional classes of verbs, as it is often the case that there is no readily identifiable semantic property that all or most members share that can readily be used as a token of class membership. The typical example here is the conjugation classes of Latin and Romance languages.

Having introduced the general profile of a canonical inflectional class, I pass now to introduce the Otomi language whose complex conjugation I study in this article.

3. Tilapa Otomi

Otomi is a small language family spoken in Central Mexico which belongs to the Oto-Pamean branch of Oto-Manguean. While the internal linguistic diversity of Otomi is uncontroversial (summarized in Lastra [2001]), there is no established consensus as to which of the different linguistic subsystems identified so far as Otomi should be treated as independent languages and which as dialects of those languages. This is why Otomi has traditionally been talked about as forming a dialectal continuum rather than a family and why authors in Mexico avoid using terms such as “language” and “dialect” and instead use “variety” as a convenient cover term. In contrast to this tradition, the current official stance by the National Institute of Indigenous Languages (INALI) in the *CLIN* (2008) is to treat Otomi innovatively as forming a “linguistic group” with nine different “linguistic varieties”.² INALI’s linguistic varieties are treated as genuine languages for official purposes, that is, for educational, administrative, judicial, social and informative purposes. This is also the stance I take in the present article.

One such language is Tilapa Otomi (henceforth T-Oto), which is spoken in the village of Santiago Tilapa, located within the municipality of the large town of Santiago Tianguistenco in the State of Mexico.³ The language has not been previously described and is currently on the verge of extinction as it is spoken by approximately a score of old people in their eighties, most of whom no longer use the language in daily life. Otomi has been spoken in this village from late prehispanic times. The old language developed there in complete isolation from other nearby variants, namely Atlapulco and Acapulco, and has emerged as a linguistic island with respect to other Otomi languages. It not only displays innovative sound changes but also a high degree of morphological conservatism, having preserved a great number of the inflectional exponents and categories found in Old Otomi; the common language spoken in the XVI century (see Palancar [2011] for a discussion).

The verbs of T-Oto fall into three conjugation classes. In principle, this morphological property is not entirely surprising because two other Otomi languages have been reported to have conjugation classes, for example Eastern Highlands Otomi (Voigtlander and Echegoyen 1979) and Ixtenco Otomi (Lastra 1997). Nevertheless, it constitutes an intriguing phenomenon when yet other Otomi languages are taken into consideration, such as Mezquital Otomi (Hess 1968; Bartholomew 2004), Santiago Mexquititlan Otomi (Hekking 1995), San Ildefonso Tultepec (Palancar 2009a), and Southwestern Otomi (Andrews 1993), which are spoken in the high plateau in the State of Mexico, Querétaro and Hidalgo. In such varieties, there is one major inflectional class to which the vast majority of the verbs in the lexicon belong, and another small class containing no more than forty intransitive verbs, called “activity verbs”

in Palancar (2009a), which could in principle be treated as a second conjugation class, but with reservations. While the comparison of the conjugation classes in Otomi is a promising enterprise that may shed light onto the different ways verb conjugation classes can emerge and decay in a linguistic family, such an enterprise falls out of the scope of this article. In the next sections, I will present basic information about verb inflection in T-Oto, which is useful to understand the morphological structure of its conjugation classes, presented in Section 5.

3.1. *Verb inflection of Tilapa Otomi*

T-Oto is a VOS head-marking language (i.e., no case marking on nouns) with accusative alignment in monotransitive verbs, and an agentive/patientive split in intransitive verbs (see Section 6 for some examples). Like other Otomi languages, T-Oto is a tonal language with three tonal contrasts in lexical roots (only functional morphemes have a high tone *vs.* a low tone contrast). Like other Oto-Pamean languages, Otomi tones lack the important functional load they have in other branches of Oto-Manguan.

T-Oto verbs inflect for features of tense, aspect and mood (TAM) by means of a set of inflectional formatives. Many such formatives commonly cross-reference the subject NP in a cumulative fashion, and for this reason, they will be treated here as “P&TAM” formatives (P stands for person of subject). I use the term “inflectional formative” in the sense of Bickel and Nichols (2007), who use it as a cover term for any marker of inflectional information, regardless of its morphological type, that is, regardless of whether it is an affix, a clitic or a functional word.

The examples in (4) illustrate the use of three such P&TAM formatives for the inflection of the transitive verb *jü^hti* ‘pay someone’ realizing grammatical person, past tense, and realis mood.⁴ In (4), it may be further noted that the object is cross-referenced by means of a suffix for the first and the second persons; the bare stem is instead used to encode reference to a third person object.

- (4) a. **tü**=*jü^hti-k’i*
 1.PST.R=pay-2OBJ
 ‘I paid you.’
- b. **gü**=*jü^hti*
 2.PST.R=pay[3OBJ]
 ‘You paid him/her/them.’
- c. **bi**=*kü^hti-ki*
 3.PST.R=SS/pay-1OBJ
 ‘He/she/they paid me.’

The P&TAM formatives in (4) always precede the verbal stem, that is, they cannot occur in other positions in the syntax of the clause. On phonological grounds, they could be characterized as simple clitics (Zwicky 1977) because of their prosodic characteristics. For example, while they are commonly phonologically hosted on the stem, like in (5), they may be uttered before a pause like independent prosodic units in erratic or doubtful speech, as shown in (5), (suspension points indicate pause). Additionally, please notice that a plural value of person is encoded by means of enclitics. This marking is only available for the first and the second person. In (5), the number enclitic is *'mbé* of first person plural exclusive.

- (5) *ja^hkamba tũ... ko='mbé*
 then 1.PST.R return=PL.EXCL
 'We came back then.' (Txt)

Furthermore, these clitics can be associated with other clitics preceding them, as for example in (6), where the inflectional formative in (5) is now hosted as an enclitic on the form *hín*, itself a dependent alternant of the negative functional word *híngi*, resulting in the prosodically free word *híndũ* /'híntĩ/ ['híndĩ]. Notice that the lenis consonant /t/ surfaces as a [d] in (6), instead of a [t].

- (6) *hín=dũ ndé=gá='mbé*
 NEG=1.PST.R want[3OBJ]=1(EMPH)=PL.EXCL
 'We didn't want it.' (Txt)

Inflected verb forms in T-Oto may involve more than one inflectional formative. This happens, for example, when a grammatical tense is formed on another (more basic) tense by adding a formative extra to the ones already realizing the basic tense. This is shown in (7a) with the imperfect continuous realis, realized by associating the clitic *má* to the exponents of present continuous realis; or in (7b), where the building of the past irrealis is carried out by associating the clitic *gi* to the markers of present irrealis. The tense value of the base markers is cancelled out by the co-occurrence of *má* or *gi*, this is why it appears in parentheses.

- (7) a. *syákáso ñũ=gá='mbé*
 just_in_case DEF=1=PL.EXCL
txá=má='mbũ='mbé ndõxi
 1(.PRES).CONT.R=IMPF=be.at.place=PL.EXCL Monday
 'We were there on Monday just in case.' (Txt)
- b. *tí=gi=jwǎ=i jǎ'ni=gwá*
 3(.PRES).IRR=PST.IRR=kill[3OBJ].B=PL person=here
 '(So that) he was going to kill the people here.' (Txt)

These inflectional formatives do not behave uniformly when hosted on the stem. An instance is given in (8), where the exponent for the first person *tú* can be either hosted as an enclitic on the preceding clitic cluster, like in (8a), or it may be hosted as a proclitic on the stem, like in (8b), (square brackets indicate junctures of phonological phrases; the main stress in the inflectional form falls on the first syllable of the verbal stem).

- (8) *para ga ta tú^hpét'u^hú*
 a. *para* [[*ga=ta*]=*tú*]=[^h*pét'u=hú*]
 b. *para* [*ga=ta*]=[*tú*=^h*pét'u=hú*]
 PURP 1.AMBU.IRR=AMBU=1=make.tortillas.AS=PL.INCL
 'So that we (you and me) go and make tortillas.' (Txt)

The structure in examples like (8) call for the existence of subtle morphophonological rules that restrain the order in which the formatives may cluster together in the shaping of an inflectional form, which at times involve different outcomes in vowel harmonization processes (see Palancar 2009b for a discussion). Nevertheless, such rules are poorly understood, and because of this, the formatives will be represented as independent morphemes orthographically. Such a treatment is immaterial for the purpose of the present article.

3.2. *The tenses of Tilapa Otomi*

The grammatical tenses of T-Oto are organized around a realis/irrealis mood distinction. The paradigm of the transitive verb *peni* 'wash (clothes)' given in Table 2 instantiates the 15 tenses that I have been able to identify in the language, with approximate translations in English. For convenience, the inflected forms exemplifying each tense are given in the second person.

In the realis mood, there is an aspectual distinction between continuous and habitual for the present and the imperfect. The present and imperfect habitual are also used in nominal predication. In addition, in Table 2 there are some tenses with local semantics that provide motion information about the subject: the *ambulative* (an alternative term for "perlocative") portrays the subject as moving about and the *andative* depicts the subject as moving away from the speech act situation in order to perform the action.

In other tenses, local values may be realized in the inflected form by means of additional morphemes associated with a specific tense, as shown in Table 3. For example, an andative value can be encoded in the ambulative realis by means of the element *-r*. Similarly, a *cislocative* value that portrays the subject moving towards the speech act situation can be encoded in the ambulative realis by means of a labial affix, which surfaces as an infix, and which is at times complemented by the formative *tí*. In other tenses, the same labial affix is used

Table 2. *The grammatical tenses of T-Oto*

Realis			
Present	continuous	<i>grá^hpeni</i>	'you're washing it now'
	habitual	<i>grú^hpeni</i>	'you commonly wash it'
Ambulative Imperfect		<i>gá^hpeni</i>	'you wash it away (here and there)'
	continuous	<i>grá má^hpeni</i>	'you were washing it'
	habitual	<i>grú mú^hpeni</i>	'you used to wash it'
	ambulative	<i>gá má tí^hpeni</i>	'you were washing it away/long ago'
Past		<i>gú^hpeni</i>	'you washed it'
Perfect		<i>xkú^hpeni</i>	'you've already washed it'
Pluperfect		<i>xkí^hpeni</i>	'you'd already washed it'
Irrealis			
Present		<i>gi^hpeni</i>	'you'll wash it'
Immediative		<i>xta gi^hpeni</i>	'you're about to wash it'
Ambulative		<i>gi tí^hpeni</i>	'you'll wash it away (here and there)'
Andative		<i>gri^hpeni</i>	'you'll go wash it'
Past		<i>gi gi^hpeni</i>	'you'd wash it'
Perfect		<i>xki gi^hpeni</i>	'you'd have washed it'

Table 3. *Local values*

Realis			
Ambulative	andative	<i>gá-r peni</i>	'you wash it away (here and there)'
	cisloc.	<i>g^{ow}á tí^hpeni</i>	'you're washing it as you come'
Past	transloc.	<i>g^{ow}ú^hpeni</i>	'you washed it somewhere else'
Perfect	transloc.	<i>xk^{ow}ú^hpeni</i>	'you've already washed it somewhere else'
Pluperfect	transloc.	<i>xk^{ow}ú^hpeni</i>	'you'd already washed it somewhere else'
Irr			
Present	transloc.	<i>g^{ow}ú^hpeni</i>	'you'll (go and) wash it somewhere else'
Past	transloc.	<i>g^{ow}ú g^{ow}ú^hpeni</i>	'you'd (go and) wash it somewhere else'

to express a *translocative* value. This value portrays the subject as either doing the action in a different place than the speech act situation or as moving away to such a place (i.e., it may convey andative semantics). This is shown in Table 3, where the labial affix appears as an infix.

Conveying local information morphologically in the verb is an old grammatical property that is manifested in all Otomi languages in some way or another, but it is only in morphologically conservative languages like T-Oto and Eastern Highlands Otomi (Voigtlander and Echegoyen 1979) where it has kept the more inflectional contrasts, although with slight differences in usage. In general, while other languages have more limited inventories, the abundance of grammatical tenses in Otomi is typologically remarkable within a larger Oto-Manguan context where the average system has inflectional contrasts

typically resolving around three main aspectual dimensions inherited from the proto-language, i.e., completive, incomplete and potential (see Kaufmann 1990).

4. The conjugation classes of Tilapa Otomi

In the previous section, I have briefly introduced a number of grammatical properties of verbs in T-Oto and have shown the different grammatical tenses that exist in the language. In this section, I show that the verbs of T-Oto fall into three classes for inflectional purposes. This can be observed in a number of facts. Consider for this purpose the inflectional behavior of the three intransitive verbs *tági* ‘fall’, *kúni* ‘grind corn’, and *fet’i* ‘shiver’ when they are inflected for the first person of present irrealis. The inflected forms are shown in (9).⁵

- (9) 1st person, present irrealis
- a. *gu*^h*tági* ‘I’ll fall’
 - b. *gu tú*^h*kúni* ‘I’ll grind (corn)’
 - c. *gi tí**fet’i* ‘I’ll shiver’

The marking contrasts in (9) point to the existence of three inflectional classes. The verbs *tági* ‘fall’ and *kúni* ‘grind corn’ are inflected by means of the P&TAM clitic *gu*, while *fet’i* selects *gi*. The morphological marking of the verb *tági* ‘fall’ is the least marked of the three as it involves only one formative. In contrast, the inflected forms of both *kúni* ‘grind corn’ and *fet’i* ‘shiver’ receive other formatives: the formative *tú* of first person, and the formative *tí*, which as we will soon see, does not realize any specific morphosyntactic or morphosemantic feature, and does not have a derivational function.

Nevertheless, the type of formal contrasts in (9), where the inflectional realizations of the three classes is different in each case, is not at all common in T-Oto. What one commonly finds instead is a situation where a given marker is used in the paradigm of more than one class at a time. For example, the verbs *tági* ‘fall’ and *kúni* ‘grind corn’ are both inflected for the first person present continuous realis by means of the same exponent, the P&TAM clitic *txá*, as shown in (10a) and (10b). The verb *fet’i* ‘shiver’ selects the same clitic plus the formative *tí* in (9c) above. This is illustrated in (10c).

- (10) 1st person, present continuous realis
- a. *txá*^h*tági* ‘I’m falling’
 - b. *txá*^h*kúni* ‘I’m grinding (corn)’
 - c. *txá tí**fet’i* ‘I’m shivering’

Similarly, for the encoding of the past realis, the verbs *tági* ‘fall’ and *fet’i* ‘shiver’ are now the ones treated alike for inflectional purposes, as shown in

(11a) and (11c), while *kúni* ‘grind corn’ selects the same clitic plus the formative *tú* of first person in (9b) above. This is illustrated in (11b), where the person formative *tú* /t̥/ surfaces as [d̥i].⁶

- (11) 1st person, past realis
- a. *tú* ^h*túgi* ‘I fell’
 - b. *tú* *dú* ^h*kúni* ‘I ground (corn)’
 - c. *tú* *fet’i* ‘I shivered’

The class to which verbs like *túgi* ‘fall’ belong is a default lexical class with the greatest number of members, and I will treat it as conjugation “class I”. The inflectional class to which the verb *kúni* ‘grind corn’ belongs has a number of formal properties that relate it to class I, as we will see in Section 6.1 I will label this class conjugation “class II”. Finally, verbs that behave like *fet’i* ‘shiver’ will be called conjugation “class III”.

In Examples (10) and (11), classes II and III have the same P&TAM clitics as class I. This is a common pattern, as we will see in the next section. Nevertheless, there are also cells in the paradigms of classes II and III which are realized by the same markers, but such markers contrast with the ones that realize the equivalent feature value for the verbs of class I. The case is often observed in the cells of the third person, although not always, and is illustrated in the present irrealis by the contrast between (12a) with respect to (12b) and (12c).⁷

- (12) 3rd person, present irrealis
- a. *ta* *túgi* (I) [3.PRES.IRR(I) ss/fall] ‘He/she/they’ll fall’
 - b. *ti* *kúni* (II) [3.PRES.IRR(II/III) ss/walk] ‘He/she/they’ll grind (corn)’
 - c. *ti* *fet’i* (III) [3.PRES.IRR(II/III) shiver] ‘He/she/they’ll shiver’

A summary of the possible contrasts are given in Table 4.

My analysis of the inflection of T-Oto verbs is based on the study of a small sample of 630 verbs. This sample was collected by me in collaboration with Néstor H. Green and Selene Hernández. The lexemes in the sample were collected from both their occurrence in natural texts and from elicitation sessions with native speakers.⁸ The distribution of the verbs in the sample attending to the different inflectional classes to which they belong is shown in Table 5,

Table 4. *Marking contrasts among the conjugation classes*

Contrasts	Ex.	Person and TAM
I ≠ II and ≠ III	(9)	1 st person, present irrealis
I = II but ≠ III	(10)	1 st person, present continuous realis
I = III but ≠ II	(11)	1 st person, past realis
II = III but ≠ I	(12)	3 rd person, present irrealis

Table 5. *Distribution of verbs per conjugation in the sample*

	verbs	%
I	395	62.5%
II	72	11.6%
III	163	25.8%
Total	630	100%

Table 6. *Ratio of transitive and intransitive verbs per class*

class	transitive	intransitive	ratio
I	204	191	1:1
II	0	72	–
III	87	76	1:1.1
Total	291	339	1:1.6

where it is further shown that class I is the class with the largest number of members.

If we also attend to the transitivity of the verbs, the sample is balanced because the ratio between transitive and intransitive verbs is almost 1:1 (actually 1:1.1, calculated by dividing the 339 intransitive verbs by the 291 transitive verbs).⁹ In principle, this ratio could be taken to be an artifice of this specific sample, and thus not representative of the lexicon, but two larger samples of other distant Otomi languages happen to reveal a similar ratio: 1:1 in Eastern Highlands Otomi from 2,000 verbs from Echegoyen and Voigtlander (2007) (967 transitive/1,033 intransitive) and 1:1.3 in Mezquital Otomi from Hernández Cruz et al. (2004) (942 transitive/1,239 intransitive). Whether this coincidence should be taken to be an indicator of how the Otomi verbal lexicon is organized according to the transitivity of a verb remains an open question. The relevant figures for T-Oto are given in Table 6.

More importantly, Table 6 shows that conjugation class II includes intransitive verbs only, revealing significant tendencies of how transitivity is mapped as a property into the conjugation classes. 70 percent of transitive verbs are found in conjugation class I, the rest in class III. Intransitive verbs are found in all classes, half of them in class I (56%) and the rest distributed evenly between II and III (21.5% and 22.5%, respectively).

In this section, I have shown the basic formal contrasts on which the three conjugation classes are based. In the next section, I first give an overview of the paradigms of each of the classes and then comment on each of the relevant subparadigms in greater detail.

5. The paradigms of the three conjugation classes

In this section, I study the morphology involved in the inflectional paradigms of each of the three conjugation classes in T-Oto. The entire paradigms for the tenses of the realis mood are given for convenience in Table 7; while those of the tenses of the irrealis appear in Table 8. In following sections, I comment on each subparadigm in further detail. In the tables, shading is used to indicate that the inflectional exponents of a given class are also used in another class. I refer to such a phenomenon as “paradigmatic neutralization”, which I interpret to be the outcome of a morphological leveling affecting classes II and III in the direction of class I, but not in all cases (see Section 4.3 for a discussion). Similarly, cells in boldface indicate the few spaces in the paradigm where the marking contrasts across the classes are maximally differentiated.¹⁰

We may see in Tables 7 and 8 that there is a high degree of neutralization across the paradigms of each class. The highest degree, with complete neutralization across the three classes, happens in many of the cells realizing the translocative value, but also in the present andative irrealis. However, it is more commonly the case that a given form is either shared by verbs of class I and II or by verbs of class I and III. In general, neutralization between class I and II happens in most of the present and the imperfect tenses of the realis mood, while between classes I and III, it occurs in most of the past tenses. There is also a certain tendency in the inflectional system of T-Oto to treat the first and the second persons in a block, independently of the morphology realizing the third person. For the third person, we also have special patterns of neutralization. For instance, while the three classes select the same marker in complete neutralization in the past realis, partial neutralization only involves either class I and II (e.g., the imperfect ambulative or the pluperfect realis) or class II and III (e.g., the perfect realis, the present irrealis, the immediative irrealis). In this respect, it should be noted that classes II and III show neutralization when the third person is involved, whereas classes I and III never do.

In the tables, there are also a great number of shaded cells throughout the paradigms. In other words, there is a high degree of paradigmatic neutralization. This makes the relatively few cells left unshaded stand out significantly as morphologically distinctive forms of a specific class. Similarly, while the cells in boldface are even fewer, they point to very specific areas in the paradigms where the marking of the three classes remains maximally distinctive. In reality, this only occurs in two tenses of the irrealis mood. Apart from encoding morphosyntactic and morphosemantic information, all such inflected forms, the ones left unshaded and the ones in boldface, also inform us about the class membership of the inflected lexeme, i.e., they serve as indexes for an inflectional class feature, which is purely morphological, in the sense of Corbett and Baerman (2006).

Table 7. Paradigms of realis

Realis mood		I	II	III	
Present continuous	1st	<i>txá^htági</i>	<i>txá^hkúni</i>	<i>txá tí fet'i</i>	
	2nd	<i>grá^htági</i>	<i>grá^hkúni</i>	<i>grá tí fet'i</i>	
	3rd	<i>ra^htági</i>	<i>ra^hkúni</i>	<i>ra tí fet'i</i>	
Present habitual	1st	<i>txú^htági</i>	<i>txú^hkúni</i>	<i>txú tú fet'i</i>	
	2nd	<i>grú^htági</i>	<i>grú^hkúni</i>	<i>grú tú fet'i</i>	
	3rd	<i>(ru)^htági</i>	<i>(ru)^hkúni</i>	<i>ru tú fet'i</i>	
	Transloc. 3rd	<i>brú^htági</i>	<i>brú^hkúni</i>	<i>brú tú fet'i</i>	
Ambulative	1st	<i>tá^htági</i>	<i>tá^hkúni</i>	<i>tá tí fet'i</i>	
	2nd	<i>gá^htági</i>	<i>gá^hkúni</i>	<i>gá tí fet'i</i>	
	3rd	<i>a^htági</i>	<i>a^hkúni</i>	<i>a tí fet'i</i>	
	Cisloc.	1st	<i>tá^htági</i>	<i>tá^hkúni</i>	<i>tá tí fet'i</i>
			<i>tá tí^htági</i>		
		2nd	<i>g^wá^htági</i>	<i>g^wá^hkúni</i>	
		3rd	<i>g^wá tí^htági</i>		<i>g^wá tí fet'i</i>
		3rd	<i>bá tági</i>	<i>bá kúni</i>	<i>bá tí fet'i</i>
	Andative	1st	<i>tár tági</i>	<i>tár kúni</i>	<i>tár tí fet'i</i>
		2nd	<i>gár tági</i>	<i>gár kúni</i>	<i>gár tí fet'i</i>
3rd		<i>ár tági</i>	<i>ár kúni</i>	<i>ár tí fet'i</i>	
Imperfect continuous	1st	<i>txá má^htági</i>	<i>txá má^hkúni</i>	<i>txá má tí fet'i</i>	
	2nd	<i>grá má^htági</i>	<i>grá má^hkúni</i>	<i>grá má tí fet'i</i>	
	3rd	<i>má^htági</i>	<i>má^hkúni</i>	<i>ma tí fet'i</i>	
Imperfect habitual	1st	<i>txú mú^htági</i>	<i>txú mú^hkúni</i>	<i>txú mú tú fet'i</i>	
	2nd	<i>grú mú^htági</i>	<i>grú mú^hkúni</i>	<i>grú mú tú fet'i</i>	
	3rd	<i>mú rú^htági</i>	<i>mú rú^hkúni</i>	<i>mú rú tú fet'i</i>	
Imperfect ambulative	1st	<i>tá má tí^htági</i>	<i>tá má dí^hkúni</i>	<i>tá má tí fet'i</i>	
	2nd	<i>gá má tí^htági</i>	<i>gá má dí^hkúni</i>	<i>gá má tí fet'i</i>	
	3rd	<i>má dí tági</i>	<i>má dí kúni</i>	<i>má tí fet'i</i>	
Past	1st	<i>tú^htági</i>	<i>tú dú^hkúni</i>	<i>tú fet'i</i>	
	2nd	<i>gú^htági</i>	<i>gú gú^hkúni</i>	<i>gú fet'i</i>	
	3rd	<i>bi tági</i>	<i>bi kúni</i>	<i>bi fet'i</i>	
	Transloc.	2nd	<i>g^wú^htági</i>	<i>g^wú^hkúni</i>	<i>g^wú fet'i</i>
		3rd	<i>b^wú^htági</i>	<i>b^wú^hkúni</i>	<i>b^wú fet'i</i>
	Perfect	1st	<i>x^tú^htági</i>	<i>x^tú dú^hkúni</i>	<i>x^tú fet'i</i>
2nd		<i>x^kú^htági</i>	<i>x^kú gú^hkúni</i>	<i>x^kú fet'i</i>	
3rd		<i>x^y ntági</i>	<i>x^pi nkúni</i>	<i>x^pi fet'i</i>	
Transloc.		2nd	<i>x^k^wú^htági</i>	<i>x^k^wú^hkúni</i>	<i>x^k^wú fet'i</i>
		3rd	<i>x^pá ntági</i>	<i>x^p^wú nkúni</i>	<i>x^p^wú fet'i</i>

Table 7. (Continued)

Realis mood		I	II	III
Pluperfect	1st	<i>x tá h'tági</i>	<i>x tá tú h'kúni</i>	<i>x tá fet'i</i>
	2nd	<i>x ki h'tági</i>	<i>x ká gú h'kúni</i>	<i>x ki fet'i</i>
	3rd	<i>x ka (n) tági</i>	<i>x ka (n) kúni</i>	<i>x ki fet'i</i>
Transloc.	2nd	<i>x k^wú h'tági</i>	<i>x k^wá gú h'kúni</i>	<i>x k^wú fet'i</i>
	3rd	<i>x k^wa tági</i>	<i>x k^wa kúni</i>	<i>x k^wú fet'i</i>

Table 8. Paradigms of irrealis

Irrealis mood		I	II	III
Present	1st	<i>gú h'tági</i>	<i>gú tú h'kúni</i>	<i>gi tí fet'i</i>
	2nd	<i>gi h'tági</i>	<i>gú gú h'kúni</i>	<i>gi tí fet'i</i>
	3rd	<i>ta tági</i>	<i>ti kúni</i>	<i>ti fet'i</i>
Transloc.	1st	<i>g^wú h'tági</i>	<i>g^wú h'kúni</i>	<i>g^wú fet'i</i>
	2nd	<i>g^wú h'tági</i>	<i>g^wú h'kúni</i>	<i>g^wú fet'i</i>
	3rd	<i>t^(w)ú h'tági</i>	<i>t^(w)ú h'kúni</i>	<i>t^(w)ú fet'i</i>
Ambulative	1st	<i>ga tá h'tági</i>	<i>ga tá tú h'kúni</i>	<i>ga tá fet'i</i>
	2nd	<i>gi tí h'tági</i>	<i>ga gú h'kúni</i>	<i>gi tí fet'i</i>
	3rd	<i>ta ga ntági</i>	<i>ta ga nkúni</i>	<i>ti gi fet'i</i>
Immediative	1st	<i>x ta gú h'tági</i>	<i>x ta gú tú h'kúni</i>	<i>x ta gú fet'i</i>
	2nd	<i>x ta gi h'tági</i>	<i>x ta gi h'kúni</i>	<i>x ta gi fet'i</i>
	3rd	<i>x ta tági</i>	<i>x ta gi h'kúni</i>	<i>x ta gi fet'i</i>
Andative	1st	<i>gri h'tági</i>	<i>gri h'kúni</i>	<i>gri fet'i</i>
	2nd	<i>gri h'tági</i>	<i>gri h'kúni</i>	<i>gri fet'i</i>
	3rd	<i>ti tági</i>	<i>ti kúni</i>	<i>ti fet'i</i>
Past	1st	<i>gú gú h'tági</i>	<i>gú gú h'kúni</i> <i>gú tú h'kúni</i>	<i>gú gú fet'i</i>
	2nd	<i>gi gi h'tági</i>	<i>gi gi h'kúni</i> <i>gú gú h'kúni</i>	<i>gi gi fet'i</i>
	3rd	<i>ti gi tági</i>	<i>ti gi kúni</i>	<i>ti gi fet'i</i>
Transloc.	1st	<i>g^wú g^wú h'tági</i>	<i>g^wú g^wú h'kúni</i>	<i>g^wú g^wú fet'i</i>
	2nd	<i>g^wú g^wú h'tági</i>	<i>g^wú g^wú h'kúni</i>	<i>g^wú g^wú fet'i</i>
	3rd	<i>tú g^wú h'tági</i>	<i>tú g^wú h'kúni</i>	<i>tú g^wú fet'i</i>
Perfect	1st	<i>x kú gú h'tági</i>	<i>x ti gú tú h'kúni</i>	<i>x kú gú fet'i</i>
	2nd	<i>x ki gi h'tági</i>	<i>x ti gú gú h'kúni</i>	<i>x ki gi fet'i</i>
	3rd	<i>x ti gi tági</i>	<i>x ti gi kúni</i>	<i>x ti gi fet'i</i>

Table 9. *The present tenses of the realis*

Realis mood		
Present continuous	Person	TAM
1st	<i>t-</i>	<i>rá</i>
2nd	<i>g-</i>	<i>rá</i>
3rd	\emptyset	<i>ra</i>
Present habitual	Person	TAM
1st	<i>t-</i>	<i>rú</i>
2nd	<i>g-</i>	<i>rú</i>
3rd	\emptyset	(<i>rú</i>)

5.1. *Inflectional contrasts in the realis mood*

In this section, I study the forms involved in the inflection of the tenses in the realis mood in more detail. There is complete paradigmatic neutralization between class I and class II in the present and imperfect tenses of the realis mood.¹¹ The cells of class III verbs in these subparadigms are all characterized, in turn, by the presence of the formative *tí*, which occurs in first position, closest to the stem.¹² In the paradigms, the P&TAM clitic in these tenses consists of a series of affixes. Consider for example the subparadigms of the present continuous and the present habitual illustrated in Table 9.

The cluster /tʀ/ resulting from the sequences *t-rá* and *t-rú* for the first person is realized as a lenis palatal affricate *tx* [tʃ̠], i.e., *txá* and *txú*. The first and the second person are further realized by an additional high tone that lands on the segmental marker of the TAM marker *ra* and *rú*. Under this analysis, the value of the third person in forms such as *rú hʔági* ‘he/she/they fall’ is realized by the bare stem, or by means of a zero morph. The marker *rú* of present habitual is also omissible, so the value of this tense can again be realized just by the bare stem.

The imperfect continuous is realized by the TAM clitic *má*; the habitual by *mú*. These clitics occur in two different positions, which are conditioned by person, as shown in Table 10.

The TAM clitic *mú* occurs closer to the stem for the first and the second persons, but precedes the P&TAM in the third person, i.e., the form for the third person is *mú rú hʔági* ‘he/she/they used to fall’, and not **rú mú hʔági*. This suggests that the same happens to the clitic *má*, although this is not obvious because the marker *rá* does not co-occur with *má* in T-Oto. In contrast, the

Table 10. *The imperfect tenses of the realis*

Realis mood				
Imperfect continuous				
	IMP.	Person	TAM	IMP.
1st		<i>t-</i>	<i>rá</i>	<i>má</i>
2nd		<i>g-</i>	<i>rá</i>	<i>má</i>
3rd	<i>má</i>	\emptyset	–	
Imperfect habitual				
1st		<i>t-</i>	<i>rú</i>	<i>mú</i>
2nd		<i>g-</i>	<i>rú</i>	<i>mú</i>
3rd	<i>mú</i>	\emptyset	<i>rú</i>	

IMP. stands for “imperfect”.

marker *ry*, which can be elided in the present habitual, is obligatory in the imperfect. Notice that *ry* further assimilates to the high tone of the clitic *mú*.

The morphology of the subparadigm of the ambulative appears in Table 11. Associated to this tense, one may have the suffix *-r* to encode an additional andative value. A cislocative sense is expressed in more complex ways. The labial prefix expressing a local value (indicated here as *B*) is associated to the forms of the second and the third persons.¹³ This labial prefix further undergoes metathesis with the onset consonant of the P&TAM clitic, and it surfaces as a labialization feature, i.e., $/_B\text{-CV}/ > /_C\text{-B-V}/ > /_C^w\text{V}/$.¹⁴ For verbs of class I, there are two encoding options for the first and the second persons, which are given as A and B, respectively. These two options are in free variation (see Section 5.2, for a discussion about this coding phenomenon). In option A, the subparadigms of class I and II are neutralized, while in option B, which involves the use of the formative *tí* of class III, there is neutralization between I and III. Neutralization is again indicated by shading. For convenience, in Table 11 I have treated the formative *tí*, characteristic of the inflection of class III, as being a flag of this inflectional class.¹⁵

The pattern of option B for class I is also found in the imperfect ambulative, as shown in Table 12. This tense is mainly used as an imperfective distant past, and is partly built on the ambulative cislocative. There is some variation in the forms, but the subparadigm in this table appears to be the most stable one across speakers. Notice that this is the only realis tense which has inflectional contrasts across the three classes, but oddly distributed by the occurrence of the formatives *tí* and *dí*.

In the subparadigm of the past realis, there is a complete neutralization between class I and III, as shown in Table 13. In contrast to the present tenses,

Table 11. *The ambulative realis*

Realis mood		I				II				III					
Ambulative		L.	P.	TAM	A.	L.	L.	P.	TAM	A.	L.	P.	TAM	A.	III
	1st	t-	á			t-	á				t-	á		tí	
	2nd	g-	á			g-	á				g-	á		tí	
	3rd	Ø	a			Ø	a				Ø	a		tí	
	Andative														
	1st	t-	á	-r		t-	á	-r			t-	á	-r	tí	
	2nd	g-	á	-r		g-	á	-r			g-	á	-r	tí	
	3rd	Ø	á	-r		Ø	á	-r			Ø	á	-r	tí	
	Cislocative														
	A 1st	t-	á			t-	á								
	B	t-	á		tí						t-	á		tí	
	A 2nd	⟨B⟩	g-	á		⟨B⟩	g-	á							
	B	⟨B⟩	g-	á	tí						⟨B⟩	g-	á	tí	
	3rd	B-	Ø	á		B-	Ø	á			B-	Ø	á	tí	

L. stands for a “local” value, which has a cislocative meaning in this tense, P. for “person”, and A. for “andative”.

Table 12. *The imperfect ambulative realis*

Realis mood		I				II				III				
Impf. ambulative		IMP.	P.	TAM	IMP.	IMP.	P.	TAM	IMP.	IMP.	P.	TAM	IMP.	III
	1st	t-	á	má	tí	t-	á	má	dí		t-	á	má	tí
	2nd	g-	á	má	tí	g-	á	má	dí		g-	á	má	tí
	3rd	má	Ø	-	dí	má	Ø	-	dí		má	Ø	-	tí

I treat the markers in past paradigms as cumulative exponents of both TAM values and person values. Class II has the person formatives *tú* and *gú*. On the other hand, with a local value, which has a translocative sense with this tense, the paradigm shows complete neutralization across the classes.¹⁶

The inflection of the perfect tenses is more complex because there are more marking contrasts involved. This is shown in Table 14. In the perfect, verbs of

Table 13. *The past realis*

Realis mood		I		II			III	
Past		L.	P&TAM	L.	P&TAM	P.	L.	P&TAM
Past	1st		tú		tú	tú ¹⁷		tú
	2nd		gú		gú	gú		gú
	3rd		bi		bi			bi
Translocative	2nd		gú		gú			gú
	3rd		bi		bi			bi

Table 14. *The perfect tenses of the realis*

Realis mood		I			II				III		
Perfect		PF.	L.	P&TAM	PF.	L.	P&TAM	P.	PF.	L.	P&TAM
Perfect	1st	x		tú	x		tú	tú	x		tú
	2nd	x		gú	x		gú	gú	x		gú
	3rd			xú	x		bi		x		bi
Translocative	2nd	x		gú	x		gú		x		gú
	3rd	x	[B-	Ø-á]	x		bi		x		bi
Pluperfect	1st	x		tá	x		tá	tú	x		tá
	2nd	x		gí	x		gá	gú	x		gí
	3rd	x		ga	x		ga		x		gí
Translocative	2nd	x		gí	x		gá	gú	x		gí
	3rd	x		ga	x		ga		x		gí

PF. stands for “perfect”.

class I and class III are inflected identically for the first and the second person. For the third person, classes II and III select the same marker in the perfect, while in the pluperfect, it is classes I and III which do.

The perfect tenses in Table 14 appear to be of late emergence. Here the perfect marker *x* results from the grammaticalization of the adverb *xá [ʃp]

‘already’, written as *xo* in Cárceres (1907 [1580]), a colonial grammar that is the earliest attestation of the common language.¹⁸ The resulting perfect tenses are hybrids but managed to maintain paradigmatic contrasts. For example, in the perfect, the forms are based on the past tense (synchronically, there is always devoicing of /g/ and /b/ before /j/, whose outcome is represented here as a lenis, e.g., $x = g\acute{u} > x = k\acute{u}$; $x = bi > x = pi$).¹⁹ For the third person, class I has the specific marker $x\acute{u}$.²⁰ The local values are regularly built on non-local ones (e.g., $x = B-g\acute{u} > x = g^{(B)}\acute{u} > x = g^w\acute{u} > x = k^w\acute{u}$),²¹ but the form for the third person of class I is based on the ambulative cislocative in Table 11. At this point, however, my understanding of the inner structure of the pluperfect remains poor.

5.2. *Inflectional contrasts in the irrealis mood*

The irrealis mood comprises six tenses in T-Oto. In other Otomi languages, this mood has fewer tenses. For example, in Eastern Highlands Otomi there are only two. One of these two tenses is a cognate of the present irrealis of T-Oto, which Voigtlander and Echegoyen (1979) treat as a “future”. The other one is called “potential mood”, itself a unique survival of the imperfect irrealis tense of Old Otomi, called ‘future perfect’ in Cárceres (1907 [1580]).

In T-Oto, the irrealis is justified as a mood on semantic and formal grounds. Semantically, all the tenses are used to convey nonfactive situations, typical of irrealis moods across languages, whether the speaker is talking about future or potential state of affairs that have not happened or had not happened at a specific point in the past, or to encode the narrative of nonfactive situations like dreams or instructions.²² Formally, the inflectional system enhances the syncretism of the first and the second person in the tenses of the irrealis mood. Similarly, it is in the irrealis where we find maximal marking contrasts across the three verb classes. This is shown in boldface in Table 15, involving cells of the present.²³ Finally, both the present translocative and the andative have complete paradigmatic neutralization; in the realis mood, this also happens in the past translocative.²⁴

The inflection of the present irrealis is complex. For the first person, conjugation class II verbs have the same P&TAM clitic as class I plus the first person marker $t\acute{u}$. The verbs of conjugation class III, select a different P&TAM clitic plus the formative $t\acute{i}$, characteristic of this class. For the second person, the sequence $g\acute{u} g\acute{u}$ of class II consists of the same clitic $g\acute{u}$ found for the first person and the second person marker $g\acute{u}$. In contrast, class III verbs use the same P&TAM clitic as conjugation class I plus the formative $t\acute{i}$.

Table 15. *Non-past tenses of the irrealis*

Irrealis mood		I				II				III					
Present		IMM	L.	P&TAM	L.	IMM	L.	P&TAM	L.	P.	IMM	L.	P&TAM	L.	III
Present	1st			<i>gu</i>				<i>gu</i>		<i>tú</i>			<i>gi</i>		<i>tí</i>
	2nd			<i>gi</i>				<i>gu</i>		<i>gú</i>			<i>gi</i>		<i>tí</i>
	3rd			<i>ta</i>				<i>ti</i>					<i>ti</i>		
Translocative	1st		⟨B⟩	<i>gu</i>			⟨B⟩	<i>gu</i>				⟨B⟩	<i>gu</i>		
	2nd		⟨B⟩	<i>gi</i>			⟨B⟩	<i>gi</i>				⟨B⟩	<i>gi</i>		
	3rd		⟨B⟩	<i>ta</i>			⟨B⟩	<i>ta</i>				⟨B⟩	<i>ta</i>		
Ambulative	1st			<i>ga</i>	<i>tá</i>			<i>ga</i>	<i>tá</i>	<i>tú</i>			<i>ga</i>	<i>tá</i>	
	2nd			<i>gi</i>	<i>tí</i>			<i>ga</i>		<i>gú</i>			<i>gi</i>	<i>tí</i>	
	3rd			<i>ta</i>	<i>ga</i>			<i>ta</i>	<i>ga</i>				<i>ti</i>	<i>gi</i>	
Immediative	1st	<i>xta</i>		<i>gu</i>		<i>xta</i>		<i>gu</i>		<i>tú</i>	<i>xta</i>		<i>gu</i>		
	2nd	<i>xta</i>		<i>gi</i>		<i>xta</i>		<i>gi</i>			<i>xta</i>		<i>gi</i>		
	3rd	<i>x</i>		<i>ta</i>		<i>x</i>		<i>ta</i>	<i>gi</i>		<i>x</i>		<i>ta</i>	<i>gi</i>	
Andative	1st			<i>gri</i>				<i>gri</i>					<i>gri</i>		
	2nd			<i>gri</i>				<i>gri</i>					<i>gri</i>		
	3rd			<i>ti</i>				<i>ti</i>					<i>ti</i>		

IMM. stands for “immediative”.

These cells are relevant because they represent one of the very few areas in the inflectional system where the marking of the three classes remains maximally differentiated. Because of this, it is worth exploring these markings historically.

The historical paradigm of the present irrealis found in Cárceres (1907 [1580]) is shown in Table 16, extracted from the paradigms in Palancar (2011: 14).²⁵

In Table 16, we can see that for the second person of class I verbs, there was already two encoding options, marked as A and B, which involved choosing between the P&TAM clitic *ka* or *ki*. Choosing *ka* preserves the syncretic pattern of first and second person proper of the irrealis mood, on which the marking of the other two classes is cleared based. Through time, option B won over

Table 16. *The present irrealis in Old Otomi*

	I		II		III
	P&TAM	P&TAM	P.	P&TAM	III
1st	<i>ka</i>	<i>ka</i>	<i>tá</i>	<i>ka</i>	<i>tí</i>
2nd	A	<i>ka</i>	<i>ká</i>	<i>ka</i>	<i>tí</i>
	B	<i>ki</i>			
3rd	<i>ta</i>	<i>ti</i>		<i>ti</i>	

A, giving rise to the modern markers: *ka* > *gu* (with raising of /a/ to /i/) and *ki* > *gi*. The marking of class II remained almost intact: *ka* = *tá* > *gu* = *tú* and *ka* = *ká* > *gu* = *gú*. What happened to the marking of class III is a little more interesting. The historical *ka* was not preserved, if it had been, the result would have been *ka* = *tí* > **gu* = *tú*.²⁶ What appears to have happened is that the marker of second person of class I was first extended to class III by leveling while retaining the formative *tí*, and was then extended to encode first person, preserving the characteristic syncretism between the first and the second person (see for example the translocative).²⁷ The path taken for the marking of class III suggests it is an instance of analogical leveling sensitive to inflectional class as discussed for example in Maiden (1992), who gives examples of leveling affecting the conjugation classes of Romance, especially of Italian. Such leveling involved the creation of new stem alternations to maximize the inflectional contrasts among the existing classes.

On the other hand, the immediative tense is morphologically based on the markers of the present irrealis. This tense appears to have emerged by associating the adverbial clitic *x* (the same marker of the perfect tenses in Table 14) to an irrealis marker *ta*, which later become the marker of third person present irrealis of class I. The sequence *xta* was then associated to the markers for the first and the second person, *gu* and *gi*, respectively. The markers of conjugation class I were then extended to conjugation class III. Here, like in the present irrealis, classes II and III have a distinct marker for the third person.

Finally, the internal structure of the subparadigms of the past and perfect of the irrealis is given in Table 17.

Speakers have various encoding options available for the past irrealis of the first and the second person of conjugation class II verbs. This is indicated in Table 17 by the use of the capital letters A, B and C. This phenomenon is called “overabundance” in Thornton (2010, 2011), where the different encoding

Table 17. The past tenses of the irrealis

	I			II			III		
	PF, L.	P&TAM L.	PST	PF, L.	P&TAM L.	PST	PF, L.	P&TAM L.	PST
Past	1st	A	gu	gu	gu	gu		gu	gu
		B			gu	tú			
		C			gu	tú			
	2nd	A	gi	gi	gi	gi		gi	gi
		B			gu	gú			
	3rd	ti	gi	gi	ti	gi	ti	ti	gi
Translocative	1st		gu		gu			gu	
	2nd		gi		gi			gi	
	3rd		ti ²⁸		ti			ti	
Perfect	1st	x	gu	gu	xí	gu	x	gu	gu
	2nd	x	gi	gi	xí	gu	x	gi	gi
	3rd	x	ti	gi	x	ti	x	ti	gi

options available for a given cell in a paradigm are called “cell-mates”.²⁹ Morphologically speaking, cell-mates under option A are produced out of leveling from conjugation class I, while cell-mates of option B result from a rule of referral to the present irrealis. Option C, which is only available for the first person, is built on option A plus the person formative *tú*.³⁰

The subparadigm of the perfect for conjugation classes I and III is based on the present irrealis plus the anterior clitic *x*. For class II, speakers have reanalyzed the form *xti* of the third person as an exponent of perfect irrealis, and have thus associated such a marker to the markers of the present irrealis. A similar process happened in the immediative in Table 15.

In this section, I have commented on relevant subparadigms in detail. Before moving into some of the principles ruling class membership for each inflectional class in Section 6, in the next section, I make a number of observations concerning the morphological status of class I as a morphological default.

5.3. *Morphological defaults*

When compared to class I, the amount of marking involved in the realization of certain tenses is often greater in class II and III. Unique to class II are the person formatives *tú* and *gú* for the first and the second person, respectively, while uniquely characteristic of class III is the formative *tí*. The occurrence of these formatives in a given inflectional form has a morphomic function, in the sense given by Aronoff (1994), in that they serve a purely morphological function, and indicate that the inflected lexeme in question either belongs to inflectional classes II or III (see further below for exceptions in four cells of the paradigm of a class I verb).

In the sequence of inflectional markers, the markers *tú* and *gú* and *tí* always occur at the position closest to the verb, while all other elements preceding them are for the most part the same P&TAM clitics that realize the inflected forms of class I verbs. This could be taken as a piece of evidence in favor of treating conjugation class I as a morphological default in building the inflection of the other two classes.

I take “morphological default” here in the sense it has in Baerman et al. (2005), among others, where a certain structure is seen as a morphological default if it serves as a base to derive other possible structures. A morphological default is the structure that remains when all other rules fail, so that a default typically surfaces in grammatical mistakes, but it is also the structure towards which the direction of leveling is oriented. In other words, a default is the structure whose occurrence is not predictable by any other principles, and because of this, it often has an erratic distribution across the paradigms. Such

Table 18. Degree of overlap among the conjugation classes

	I				II				III				
	same	%	diff.	%	same	%	diff.	%	same	%	diff.	%	
I	66	–	–	–	44	66.6	22	33.4	32	48.5	34	51.5	
II	67	44	65.6	23	34.4	–	–	–	–	22	32.9	45	67.1
III	64	32	50.0	32	50.0	22	34.4	42	65.6	–	–	–	–

a structure can in principle be the most frequent in some systems, but not necessarily.³¹ On the phonological side, on the other hand, a default often correlates with the least marked option.

The shading in Tables 7 and 8 above pointed to the existence of a high degree of neutralization across the paradigms of the three inflectional classes. The degree of overlap between the markers involved in the paradigms is given in Table 18. The figures include cell-mates too.³² For the paradigm of class I verbs, there are 66 different markers or sequences of markers to realize 64 cells, 67 in class II, and 64 in class III (i.e., this class lacks cell-mates). The labels “same” and “different” indicate whether the inflectional markers realizing these cells are the same or are different than the ones realizing the same cell for a different class.

Table 18 indicates that the overlaps between class I and class II, and between class I and class III are greater than those involving class II and III. For instance, 66.6% of all markers involved in the inflection of a class I verb are also used for the inflection of class II. With class III, the overlap is of 48.5%. However, only 32.9% of the markers of a class II verb are actually also used to inflect a class III verb. The figures in Table 8 further indicate that the degree of neutralization between class I and II is greater than that between class I and III.

In Section 5, I have shown that there are cases where there is complete neutralization across the three classes. In this light, Table 18 is not entirely informative about the type of overlap that happens. In other words, there are some questions that remain. For example, to what extent does the overlap between class II and III also involve class I? or to what extent are the inflectional markers involved in the paradigm of a given class distinctive of that particular class? To address such questions, please consider Table 19 first.

Table 19 shows that in the entire paradigm of a class I verb there are only eight markers that are distinctively class I markers! This means that almost 90% of the morphology realizing the paradigm of a class I verb is also used to inflect classes II and III, although to various degrees. In contrast, class II has 18 distinctive markers or sequence of markers (16 of which involve the person formatives *tú* and *gú*), while class III has up to 28, of which 23 (93%) indeed

Table 19. *Distinctive marking vs. overlaps in verbal inflection*

	I		II		III	
Distinctive	8	12.2%	19	28.4%	28	43.8%
Overlapping	58	87.8%	48	71.6%	36	56.2%
Total	66	100%	67	100%	64	100%
Ratio	1/13.8	100%	1/3.9	100%	1/1.2	100%

Table 20. *Types of overlap among classes*

	I		II		III	
= II and = III	18	31%				
= II but ≠ III	26	44.8%				
= III but ≠ II	14	24.2%				
= I and = III			18	37.5%		
= I but ≠ III			26	54.1%		
= III but ≠ I			4	8.4%		
= I and = II					18	50%
= I but ≠ II					14	38.8%
= II but ≠ I					4	11.2%
Total	58	100%	48	100%	36	100%

carry the formative *tí*. In Table 19, I have also included the ratio of distinctive markers per overlapping ones. For class I, this ratio is extremely high: for each distinctive marker we find 14 others which overlap. We find class III in the opposite end, with an almost 1:1 ratio.

Alternatively, the remaining overlapping areas in Table 19 show some distributional asymmetries, given in Table 20.

There are 18 cells showing complete neutralization across the paradigms of the three classes. This represents 31% of all overlapping cases for the inflection of class I verbs, 37.5% of class II, and up to 50% of class III! Furthermore, Table 20 also shows that those overlapping cases between class II and class III that exclude class I are rather low: 8.4% for class II and 11.2% for class III. This situation can be interpreted in two alternative ways.

One way to interpret this is to view the morphology involved in the realization of the paradigm of class I as a morphological default. Such a default would have spread over 65.6% of the paradigmatic space of class II verbs, and over 50% of class III verbs.³³ Alternatively, class II or III could be regarded as colonizers of the 87.8% of the paradigmatic space of class I.³⁴ In this article, I

adhere to the first possibility, although in reality there are arguments in favor of either direction.

Class I serves as a base to derive the inflection of classes II and III in cases where the formatives *tí* and *gí* and *tí* are added to the P&TAM clitic of class I. The role of class I as a default is further sustained in the local subparadigms. For example, the inflectional realization of the present irrealis translocative is common to all classes but it is based on the markers of conjugation class I. Similarly, the subparadigm of the immediative tense for verbs of conjugation class III is based on class I. If this were not the case, the inflected form for the first person would be something like **xta gi tí* instead of the attested form *xta gí*. The same applies for the past and the perfect irrealis.

Third person cells are especially interesting because they always show neutralization. Class I and II show neutralization in the present and the imperfect tenses of the realis, but also in the pluperfect. The same applies to class II and III in the perfect realis, in the perfect translocative realis, and in the present realis. Such neutralization makes the marker of class I stand out distinctively in three out of the only eight cases where class I receives distinct marking.³⁵ In some cases we find the same markers across the three classes, for example in many local tenses, and in the past realis. But in none of the cases we can really pinpoint the direction of the leveling in an exact way. This is only possible in a very few cases. For example, the leveling is clearly from class I to class II in the ambulative irrealis where the P&TAM marker is *ta* instead of *tí*, which is the same used in the present irrealis. But in the past irrealis (also the perfect; based on the past), the direction is exactly the opposite, i.e., from class II and III to class I, because the marker involved is *tí* instead of *ta*, although this cell may be built on the andative irrealis.

In both the ambulative cislocative realis and in its derived tense, the imperfect ambulative, there is what appears to be a leveling from class III to class I for the first and second persons. The forms in question carry the formative *tí*, which is characteristic of class III. Nevertheless, it is not entirely clear that this is an instance of leveling, because the forms of class II also have a similar morpheme *dí* for all persons, and both morphemes (the *tí* of class I and the *dí* of class II) could have in principle developed from a common directive morpheme, ultimately related to the P&TAM *tí* used for the third person of andative irrealis.

Finally, despite the arguments in favor of class I being a morphological default, the language behaves in a remarkable way concerning loanwords. One should expect that, all things being equal, the optimal way to integrate a loanword into the inflectional system is by means of the morphological default. This is what happens crosslinguistically. In T-Oto, however, loaned transitive verbs from Spanish are categorized as members of class III.³⁶ Some examples are given in (25) below.

6. The membership of the conjugation classes

For a grammarian, confronted with a linguistic system with densely populated inflectional classes, it becomes natural to seek for generalizations that may help us understand the way a given lexeme is assigned class membership in the lexicon. Against this expectation, in his defense of inflectional classes as purely morphological ontologies (i.e., morphomic), Aronoff (1994) regards class assignment to conjugation classes as a necessarily random phenomenon. As we have seen in Section 2, this is also the position taken by Corbett (2009) when defining the ideal construct of canonical inflectional classes according to criteria 6–9 in (3), to which I added in (3) the criterion suggesting a lack of semantic motivation.

Nonetheless, those in search for generalizations often find them, and there are at times phonological, morphological, as well as semantic factors that can orient class membership at least in a probabilistic way. In this section, I show that conjugation class I is very canonical in this respect, because there is no obvious specific property that most of its members share that may serve as an index for class membership. In contrast, those properties exist in class III, and especially in class II, which stands as the least canonical class of the three.

As for T-Oto, conjugation class I is a major class that serves as a lexical default. The greatest bulk of the verbs of this class are semantically heterogeneous and defy any semantic generalization. In this class, we also find all stative and inchoative verbs of the language. These verbs are intransitive verbs which are morphologically patientive (i.e., their subject is cross-referenced as an object). The verbs are inflected for TAM using the markers of the third person of conjugation class I.³⁷ An example of a stative verb is given in (13), while two inchoative verbs are illustrated in (14).³⁸

- (13) *xu=ndén-gi*
 PERF.R=swell.AS-1OBJ
 ‘I got swollen up’.
- (14) a. *b^(w)u=tú-^hpa-’k’i*
 <TRANSLOC>PST.R=INCH-heat.up.AS-2OBJ
 ‘You got hot’.
- b. *xu=tú-k’wáni*
 PERF.R=INCH-boil[3OBJ]
 ‘It’s boiled’.

In contrast to the verbs of conjugation class I, the members in classes II and III have a semantic and a morphological profile which is more readily identifiable. This profile deserves a study of its own in order to shed light onto how these classes operate in T-Oto, especially because they compete with conjugation class I in the partition of the verbal lexicon.

6.1. The verbs of conjugation class II

The 72 verbs of conjugation class II in my sample are all intransitive verbs. Most of such intransitive verbs are unergative verbs with a specific semantic profile: they express daily-life activities performed by human beings. In this sense, because many of the verbs in this class have this semantic property, being an unergative verb in the language is a property that orients class membership to class II, at least in a probabilistic way.

Similarly, the class has both basic (non-derived) and derived verbs. The list of the verbs in (15) illustrates a number of basic members of class II which typically illustrate this semantics, (*intr* stands for “intransitive”).

- (15)
- | | | |
|-------------------------|-----------|--------------------------|
| <i>’ófo</i> | (intr/II) | ‘write’ |
| <i>něj</i> | (intr/II) | ‘dance’ |
| <i>ñö</i> | (intr/II) | ‘speak’ |
| <i>pét’i</i> | (intr/II) | ‘make <i>tortillas</i> ’ |
| <i>tyúhü</i> | (intr/II) | ‘sing’ |
| <i>pé^hti</i> | (intr/II) | ‘beg’ |
| <i>’afi</i> | (intr/II) | ‘scrap agave’ |
| <i>xu</i> | (intr/II) | ‘chop wood’ |
| <i>xófo</i> | (intr/II) | ‘harvest’ |

The class also includes other intransitive verbs expressing the same type of human activities but this time morphologically derived from transitive verbs of conjugation class I. I call such verbs “antipassive”. In T-Oto, there are four rules that produce such antipassive verbs: (i) by means of the affixation of *n-*, (ii) by the affixation of *-te*, (iii) by compounding, and (iv) by changing inflectional class.³⁹

(i) A nasal prefix *n-* is found in the five verbs in (16). The three verbs in (16a) are clearly related to an existing transitive verb, but the two in (16b) are old formations which lack a transitive correlate, (*tr* stands for “transitive”).

- (16)
- | | | | | | | | |
|----|---------------------------|-----------|----------------|---|-------------------------|--------|--------------|
| a. | <i>m-fádi</i> | (intr/II) | ‘watch cattle’ | < | <i>fádi</i> | (tr/I) | ‘watch over’ |
| | <i>m-pó^hti</i> | (intr/II) | ‘patch’ | < | <i>pó^hti</i> | (tr/I) | ‘patch’ |
| | <i>n-tsöni</i> | (intr/II) | ‘crack (wood)’ | < | <i>tsöni</i> | (tr/I) | ‘hack, chew’ |
| b. | <i>n-tx’éni</i> | (intr/II) | ‘play’ | | | | |
| | <i>n-tx’önt’i</i> | (intr/II) | ‘chat’ | | | | |

(ii) A derivational affix *-te* (at times also *-púte*) is found in more verbs, given in (17). The suffix renders the action of the source verb as a customary occupational activity. Derived verbs with *-te* may be further converted into agent nouns that designate such occupations.

- (17) *xá-^hpute* (intr/II) ‘teach’/‘be a teacher’ < *xádi* (tr/I) ‘teach’
ú^htu-te (intr/II) ‘teach’/‘be a teacher’ < *’ú^hti* (tr/I) ‘show’
pá-^hpute (intr/II) ‘sell’/‘be a salesperson’ < *pá* (tr/I) ‘sell’
’óthe-te (intr/II) ‘heal’/‘be a healer’ < *’óthe* (tr/I) ‘heal’
ho-^hte (intr/II) ‘beat’/‘be a thug’ < *ho* (tr/I) ‘hit’
tsa-^hte (intr/II) ‘bite’/‘be a biter’ (for dogs) < *tsa* (tr/I) ‘bite’
ñú-^hte (intr/II) ‘have sight’/‘be a visionary’ < *ñú* (tr/I) ‘see’

(iii) Antipassive verbs are also produced by V+N compounding. The five examples in my sample are given in (18). The last example also bears the antipassive prefix *n-*.

- (18)
hón+jö (intr/II) ‘look for mushrooms’ < *hóni* ‘look for’ (tr/I) + *jö* ‘mushroom’
tsí+hme (intr/II) ‘have a meal’ < *tsí* ‘ingest’ (tr/I) + *hme* ‘tortilla’
tyü+the (intr/II) ‘be thirsty’ < *tyü* ‘die’ (tr/I) + *téhe* ‘water’
pé+’mbita (intr/II) ‘make music’ < *pé^l* ‘play instr.’ (tr/I) + *’mbita* ‘music’
n-tsón+zá (intr/II) ‘hack wood’ < *tsóni* ‘hack, chew’ (tr/I) + *zá* ‘wood’

(iv) The antipassive verbs of conjugation class II in (19) are also derived from class I by just changing the membership of the verb to class II.

- (19)
mbát’i (intr/II) ‘call’ < *mbát’i* (tr/I) ‘call’
’ráts’i (intr/II) ‘cross’ < *’ráts’i* (tr/I) ‘cross’
páxi (intr/II) ‘sweep’ < *páxi* (tr/I) ‘sweep’
pě (intr/II) ‘steal’ < *pě* (tr/I) ‘steal’
peni (intr/II) ‘do the laundry’ < *peni* (tr/I) ‘wash (clothes)’
kúni (intr/II) ‘grind corn’ < *kúni* (tr/I) ‘grind corn’
’wǒhō (intr/II) ‘plow’ < *’wǒhō* (tr/I) ‘plow’
něhi (intr/II) ‘read’ < *něhi* (tr/I) ‘read’
tyǔhū (intr/II) ‘sow’ < *tyǔhū* (tr/I) ‘sow’
xóngi (intr/II) ‘open corn-cob’ < *xóngi* (tr/I) ‘open (corn-cob)’
tsu’ti (intr/II) ‘smoke’ < *tsu’ti* (tr/I) ‘smoke’
húxi (intr/II) ‘blow instrument’ < *húxi* (tr/I) ‘blow (instrument)’
tō-beni (intr/II) ‘forget’ < *tō-beni* (tr/I) ‘forget’
’wédi (intr/II) ‘sew’ < *’wédi* (tr/I) ‘sew’
nthó-nde (intr/II) ‘kiss’ < *nthó-nde* (tr/I) ‘kiss’ (*nde* ‘mouth’)

In contrast to the cases in (16) to (18), the derivation of the intransitive verbs in (19) does not occur by means of any overt morphological marking, but instead involves a special derivational rule that implies changing the inflec-

tional class of the verbs involved. This is a possible analysis that has been already advanced for some of the conjugation classes of Hebrew in Aronoff (1994) where they are known as “binyanim”.⁴⁰ Alternatively, the verb pairs in (19) could be interpreted as being labile, that is, verbs which can both function transitively and intransitively without explicit morphological means, such as for example English *drive* in its transitive use in *I’m driving a BMW these days* contrasted with intransitive *I’m driving*. Nevertheless, I take true labile verbs in T-Oto to be those which do not change conjugation class, and such verbs abound in conjugation classes I and III, e.g., *tsoni* (tr/intr/I) ‘chew’, *joni* (tr/intr/I) ‘sweep’, *k’u^hki* (tr/intr/III) ‘burst’, *péxu* (tr/intr/III) ‘weigh’), where the intransitive verb in the pair more often than not has the same agentive meaning as the intransitive verbs of class II in (19). If the relation between the verbs in (19) were to be treated as that of labile verbs, one should expect that there would be no restriction on the source of the transitive verb in question, and as it happens, no transitive verb of conjugation class III has an antipassive intransitive verb in class II.

The class also includes verbs which, depicting actions carried out by humans, are also typically targeted by middle voice morphology according to Kemmer (1993). The verbs in (20) are the ones found in my sample which have middle semantics. These verbs are basic verbs which are not marked morphologically with the middle prefix *n-*, which as we will see in Section 6.2.1 belong to conjugation III. The verbs in (20) are organized according to the middle situation types proposed by Kemmer, which include verbs depicting actions involving the human body, as well as other cognitive and motion verbs.

(20) Involving the body:	<i>'mbũ</i>	(intr/II)	‘live’
	<i>'õhö</i>	(intr/II)	‘sleep’
	<i>ñũhũ</i>	(intr/II)	‘wake up’
	<i>káni</i>	(intr/II)	‘snore’
	<i>kõ'ts'i</i>	(intr/II)	‘breathe’
	<i>hi</i>	(intr/II)	‘have a bath’
	<i>tsũ</i>	(intr/II)	‘suckle’
	<i>théde</i>	(intr/II)	‘laugh’
	<i>mbáfi</i>	(intr/II)	‘yell’
	<i>hétxi</i>	(intr/II)	‘sneeze’
	<i>pi</i>	(intr/II)	‘defecate’
	<i>sku^{-h}pi⁴¹</i>	(intr/II)	‘urinate’
	<i>há+ndé⁴²</i>	(intr/II)	‘yawn’
	Physio-psychological:	<i>tsü</i>	(intr/II)
<i>zóni</i>		(intr/II)	‘cry’
<i>hwõ'ts'i</i>		(intr/II)	‘get startled’
<i>tõnts'i</i>		(intr/II)	‘complain’

Motion:	<i>'mba</i>	(intr/II)	'stand up'
	<i>m-p'ē'ts'i</i>	(intr/II)	'lean on one's chest'
	<i>'yo</i>	(intr/II)	'walk'
	<i>p<u>u</u></i>	(intr/II)	'walk fast'
	<i>ihi</i>	(intr/II)	'be in a hurry'

Finally, in the sample there are at least nine verbs with non-human subjects that do not readily conform to the general profile of the class: *'mbét'o* 'go first', *fôgi* 'bark',⁴³ *hüi* 'lay eggs', *ja* 'exist', *tsô* 'catch fire', *zúdi* 'for a thing to hang', *n-k'ô'ts'i* 'loom, appear', *kô* 'for a flaccid mass to be on top of something' and *te* 'grow/be alive'. The relevance of these verbs will be discussed further in Section 7.

6.2. The verbs of conjugation class III

In contrast to conjugation class II, the semantic profile of the members of class III is heterogeneous and membership is not restricted to a specific transitivity type. Nevertheless, a closer look at the verbs in this class reveals that the class functions as a depository of verbs which have a special set of grammatical properties (semantic or morphological), and consequently, many of the members of this class can still be organized in small idiosyncratic groups. This makes the principles ruling the membership of a lexeme to this class an intricate phenomenon, as I will try to show.

6.2.1. *Transitive verbs.* The class includes a few typical ditransitive (dtr) verbs. The ones found in the sample are given in (21).

- (21) *'üni* (dtr/III) 'give to a third person'
'na (dtr/III) 'give to a first or second person'
hwé^hki (dtr/III) 'give as a gift'
hmi^hti (dtr/III) 'lend'

In the sample, there are also a number of inchoative verbs of conjugation class I which originated from transitive verbs of conjugation class III by means of the inchoative prefix *tu-*.

- (22)
- | | | |
|------------------------------------|---|--|
| <i>k'wáni</i> (tr/III) 'boil' | > | <i>tu-k'wáni</i> (intr/I) 'boil' |
| <i>'mbédi</i> (tr/III) 'lose' | > | <i>tu-'mbédi</i> (intr/I) 'get lost' |
| <i>'ot'i</i> (tr/III) 'dry' | > | <i>tu-'ot'i</i> (intr/I) 'get dry' |
| <i>'ügi</i> (tr/III) 'sweeten' | > | <i>tu-'ügi</i> (intr/I) 'sweeten' |
| <i>'üxki</i> (tr/III) 'make salty' | > | <i>tu-'üxki</i> (intr/I) 'get salty' |
| <i>kút'i</i> (tr/III) 'put in' | > | <i>tu-^hkút'i</i> (intr/I) 'go in' |

<i>mbá^hki</i>	(tr/III)	‘enlarge’	>	<i>t_u-mbá^hki</i>	(intr/I)	‘get large’
<i>ntyúmi</i>	(tr/III)	‘make lukewarm’	>	<i>t_u-ntyúmi</i>	(intr/I)	‘be lukewarm’
<i>ñúts’i</i>	(tr/III)	‘fill’	>	<i>t_u-ñúts’i</i>	(intr/I)	‘fill up’
<i>pat’i</i>	(tr/III)	‘heat up’	>	<i>t_u-^hpat’i</i>	(intr/I)	‘get hot’
<i>pídi</i>	(tr/III)	‘frighten’	>	<i>t_u-^hpídi</i>	(intr/I)	‘get frightened’
<i>ihége</i>	(tr/III)	‘finish up’	>	<i>t_u-ihége</i>	(intr/I)	‘end’
<i>tō</i>	(tr/III)	‘cook’	>	<i>t_u-^htō</i>	(intr/I)	‘boil’
<i>ts’ōgi</i>	(tr/III)	‘make dirty’	>	<i>t_u-ts’ōgi</i>	(intr/I)	‘get dirty’

Besides the verbs in (22), the three verbs in (23) have an intransitive correlate in conjugation class II. The relation between the verbs in (23) is difficult to interpret, but in principle they could be taken to be old outcomes of a causative rule (no longer productive), which involved change of inflectional class.

(23)	<i>ráts’i</i>	(tr/III)	‘cross someone’	<	<i>ráts’i</i>	(intr/II)	‘cross’
	<i>tsō</i>	(tr/III)	‘set fire’	<	<i>tsō</i>	(intr/II)	‘catch fire’
	<i>ñō</i>	(tr/III)	‘talk of someone’	<	<i>ñō</i>	(intr/II)	‘talk’

Similarly, there are 14 verbs in this class which are truly labile. In such pairs, the intransitive verb most commonly depicts a human activity similar to the semantics of most verbs of class II (24a). Those that stray from this semantic generalization are given in (24b).

(24)	a.	<i>f_uts’i</i>	(tr/intr/III)	‘uproot’
		<i>gwónts’i</i>	(tr/intr/III)	‘stir up’
		<i>ts_uni</i>	(tr/intr/III)	‘prick’
		<i>k’ats’i</i>	(tr/intr/III)	‘gather cattle’
		<i>múnts’i</i>	(tr/intr/III)	‘gather, harvest’
		<i>múndo</i>	(tr/intr/III)	‘gather, heap up’
		<i>nōmi</i>	(tr/intr/III)	‘water’
		<i>gósa</i>	(tr/intr/III)	‘enjoy’
		<i>péxu</i>	(tr/intr/III)	‘weigh’
		<i>n_uxú^hki</i>	(tr/intr/III)	‘heal (with witchcraft)’
		<i>ngwá^hki</i>	(tr/intr/III)	‘rinse’
		<i>xadi⁴⁴</i>	(tr/intr/III)	‘pray’
	b.	<i>jwádi⁴⁵</i>	(tr/intr/III)	‘finish’ (tr. ‘kill’)
		<i>xüni</i>	(tr/intr/III)	‘split’
		<i>k’u^hki</i>	(tr/intr/III)	‘burst’

All transitive loanwords from Spanish also belong to this conjugation class. The verbs in (25) are the ones found in the sample. This is evidence that the rule assigning class membership to class III was still productive at least at the time when the words were borrowed.⁴⁶ The fact that such loan verbs appear in class III is surprising, for the expectation is that they should be categorized as

members of class I, which is the class that serves as both a morphological and a lexical default.

(25)

<i>ándaró</i>	(tr/III)	‘for an illness to be about to affect someone’	<	<i>andar</i> ⁴⁷
<i>gósa</i>	(tr/intr/III)	‘enjoy’	<	<i>gozar</i>
<i>hórba</i>	(tr/III)	‘annoy’	<	<i>jorobar</i>
<i>kóbra</i>	(tr/III)	‘charge’	<	<i>cobrar</i>
<i>mánsa</i>	(tr/III)	‘tame, calm somebody down’	<	<i>amansar</i>
<i>mpáréhó</i>	(tr/III)	‘make a surface even’	<	<i>emparejar</i>
<i>péxu</i>	(tr/intr/III)	‘weigh’	<	<i>pesar</i>
<i>préséntá</i>	(tr/III)	‘present, introduce’	<	<i>presentar</i>

Finally, the remaining 51 transitive verbs in the sample of conjugation class III are basic verbs with a miscellany of meanings with no apparent semantic core. The verbs include eventive verbs (e.g., *jwádi* ‘finish/kill’, *fó^hki* ‘squash’, *fó^hti* ‘nail’, *ja* ‘do’, *tho^hki* ‘bore a whole’, *tsétⁱ* ‘chop’, *tsuni* ‘prick’, *tyutⁱ* ‘crush’, etc.), verbs of carrying (e.g., *thěni* ‘load’, *wě^hki* ‘take away’, *kó^ttsⁱ* ‘return’, etc.), and many other verbs which do not involve physical impact (e.g., *tě^hti* ‘entrust’, *ja-mō-mōdi* ‘thank’, *zěngwa* ‘greet’, *’ñě* ‘have family/animals’, etc.).

6.2.2. *Intransitive verbs.* The intransitive verbs of class III are subject to more generalizations than the transitive verbs in Section 6.2.1. For example, all reciprocal verbs in the language belong to this class. Reciprocal verbs in T-Oto are intransitive verbs which are commonly derived (the only basic verb in my sample is *thō^hti* ‘get married’). Most reciprocal verbs are derived by means of the middle voice prefix *n-*, which is homophonous to the antipassive marker in (16) above.⁴⁸ The verbs in the sample are given in (26).⁴⁹ Some of such verbs are deponent verbs, that is, verbs which are marked with the middle voice prefix, but lack a transitive counterpart. These are given in (27), (*e.o.* stands for “each other”).

(26)

<i>m-pǒdi</i>	(intr/III)	‘get to know e.o.’	<	<i>pǒdi</i>	(tr/I)	‘know’
<i>n-hufi</i>	(intr/III)	‘hug e.o.’	<	<i>hufi</i>	(tr/I)	‘hug’
<i>n-thětsⁱ</i>	(intr/III)	‘hug e.o.’	<	<i>thětsⁱ</i>	(tr/I)	‘hug’
<i>n-húnta</i>	(intr/III)	‘get together’	<	<i>húnta</i>	(tr/III)	‘put together’
<i>m-préséntá</i>	(intr/III)	‘introduce to e.o.’	<	<i>préséntá</i>	(tr/III)	‘introduce’
<i>n-xüni</i>	(intr/III)	‘get divorced’	<	<i>xüni</i>	(tr/III)	‘split’
<i>n-zěngwa</i>	(intr/III)	‘greet e.o.’	<	<i>zěngwa</i>	(tr/III)	‘greet’
<i>nthó+nde</i>	(intr/III)	‘kiss e.o.’	<	<i>nthó+nde</i>	(tr/I)	‘kiss’
<i>mpu^ttsⁱ</i>	(intr/III)	‘bump into e.o.’	<	<i>mpu^ttsⁱ</i>	(tr/III)	‘bump into sth.’

- | | | | | | |
|------|---------------------------------|------------|-------------------------|--|--|
| (27) | <i>n-tyü</i> | (intr/III) | ‘fight’ | | |
| | <i>n-syëti</i> | (intr/III) | ‘be opposite to e.o.’ | | |
| | <i>n-thyë</i> | (intr/III) | ‘come across e.o.’ | | |
| | <i>n-syütu+kü</i> ⁵⁰ | (intr/III) | ‘murmur in e.o.’s ears’ | | |
| | <i>n-jo+ndé</i> ⁵¹ | (intr/III) | ‘tell e.o. secrets’ | | |

The reciprocal verbs in (28) were also produced by changing the conjugation class of the verb, just like the causative ones in (23), and the agentive verbs in (19).

- | | | | | | | | |
|------|-------------|------------|-------------------------|---|-------------|-----------|---------------|
| (28) | <i>ñö</i> | (intr/III) | ‘talk to e.o.’ | < | <i>ñö</i> | (intr/II) | ‘talk’ |
| | <i>ñü</i> | (intr/III) | ‘see each other’ (e.o.) | < | <i>ñü</i> | (tr/I) | ‘see’ |
| | <i>zófo</i> | (intr/III) | ‘talk to e.o.’ | < | <i>zófo</i> | (tr/I) | ‘talk to sb.’ |
| | <i>ñöni</i> | (intr/III) | ‘help e.o.’ | < | <i>ñöni</i> | (tr/I) | ‘help’ |

Similarly, conjugation class III holds the vast majority of middle verbs of the language. A middle verb in T-Otomi is an intransitive verb that bears the middle prefix *n-* and has the semantics associated with the middle situation types in Kemmer (1993), i.e., including actions involving the body, motion and other cognitive and psychological actions. The middle verbs found in my sample that belong to this class appear in (29). These verbs have similar middle semantics as the ones presented in (20) above which belong to class II. The only difference between these and the ones in (20) involves the presence of the middle prefix *n-*.⁵² A middle voice analysis can be also extended to account for the reciprocal verbs in (26) above. Deponent middle verbs are given in (30), and they are equivalent to the deponent reciprocal verbs in (27).⁵³

- | | | | | | | | |
|------|-----------------------------|------------|------------------------|---|-------------------------|----------|-----------|
| (29) | <i>n-tx’ë</i> | (intr/III) | ‘agree with, suit’ | < | <i>’ëhë</i> | (intr/I) | ‘come’ |
| | <i>n-k’á</i> | (intr/III) | ‘get wet’ (on purpose) | < | <i>k’á</i> | (intr/I) | ‘get wet’ |
| | <i>n-köni</i> | (intr/III) | ‘rest’ | < | <i>köni</i> | (tr/I) | ‘cool up’ |
| | <i>m-pádi</i> | (intr/III) | ‘change clothes’ | < | <i>pádi</i> | (tr/I) | ‘change’ |
| | <i>n-ʔot’i</i> | (intr/III) | ‘fasten’ | < | <i>ʔot’i</i> | (tr/I) | ‘fasten’ |
| | <i>n-ʔá^hki</i> | (intr/III) | ‘scratch (a lot)’ | < | <i>ʔá^hki</i> | (tr/I) | ‘cut’ |
| | <i>n-tx’ë^hpi</i> | (intr/III) | ‘dash’ | < | <i>’ë^hpi</i> | (tr/I) | ‘cast’ |
| | <i>n-tx’ëxki</i> | (intr/III) | ‘comb’ | < | <i>’ëxki</i> | (tr/I) | ‘comb’ |
| | <i>n-tx’óxni</i> | (intr/III) | ‘shave’ | < | <i>’óxni</i> | (tr/I) | ‘shave’ |
| | <i>m-pát’i</i> | (intr/III) | ‘get warm’ | < | <i>pát’i</i> | (tr/III) | ‘heat up’ |
| | <i>m-p’édi</i> | (intr/III) | ‘get lost’ | < | <i>mbédi</i> | (tr/III) | ‘lose’ |
| | <i>m-péxu</i> | (intr/III) | ‘weigh (oneself)’ | < | <i>péxu</i> | (tr/III) | ‘weigh’ |
| | <i>n-jwádi</i> | (intr/III) | ‘kill oneself’ | < | <i>kwádi</i> | (tr/III) | ‘kill’ |

(30)	<i>m-fɛts'i</i>	(intr/III)	'swim'			
	<i>m-pá^hki</i>	(intr/III)	'get something twisted'			
	<i>n-tx'óni</i>	(intr/III)	'hide'			
	<i>n-tx'ɛnte</i>	(intr/III)	'get involved'			
	<i>n-txüni</i>	(intr/III)	'get all muddy'			
	<i>n-hwíts'i</i>	(intr/III)	'get in bed'			
	<i>n-k'ót'i</i>	(intr/III)	'undress'			
	<i>n-kóde</i>	(intr/III)	'get dressed'			
	<i>n-syé</i>	(intr/III)	'look at oneself in the mirror'			
	<i>n-thé</i>	(intr/III)	'agree with, suit'			
	<i>n-tóxni</i>	(intr/III)	'lean on one side'			
	<i>n-tx'ó+mféni</i>	(intr/III)	'have second thoughts'			

Other intransitive verbs of class III in my sample include a number of derived antipassive verbs (31), which are similar to the verbs of conjugation class II, and the few basic atmospheric verbs (32).⁵⁴

(31)						
	<i>m-pefi</i>	(intr/III)	'work'	<	<i>pefi</i>	(tr/I) 'work, do'
	<i>zɛngwa-te</i>	(intr/III)	'greet'	<	<i>zɛngwa</i>	(tr/III) 'greet'
	<i>hmi^htu-te</i>	(intr/III)	'be lending'	<	<i>hmi^hti</i>	(tr/III) 'lend'
	<i>hwé^hku-te</i>	(intr/III)	'donate'	<	<i>hwé^hki</i>	(tr/III) 'give as a gift'
(32)	<i>hátsi</i>	(intr/III)	'dawn, wake up at the start of the day' ⁵⁵			
	<i>ne^hki</i>	(intr/III)	'dawn'			
	<i>n-jüni⁵⁶</i>	(intr/III)	'get cloudy'			
	<i>ju't'i</i>	(intr/III)	'drip, leak'			

The rest of the intransitive verbs in the sample do not readily conform to any semantic or morphological generalization. Such verbs include: *fɛt'i* 'tremble', *k'ént'i* 'get an ankle twisted', *mungi* 'get skin wrinkled', *súfri* 'suffer', *tóni* 'stroll', *xóni* 'stride', *tá^hki* 'scratch', *mituxi* 'be lodged at some place', *hmön+the* 'water by hand' (*téhe* 'water') and *ju't'i* 'drip, leak'.

7. Canonicity and the conjugation classes of Tilapa Otomi

In this section, I briefly evaluate the degree of canonicity of the conjugation classes of T-Oto against the canonical approach to inflectional classes by Corbett's (2009) presented in Section 2. I show that classes I and III are more canonical than class II, which is rather atypical. There were two principles ruling the canon: the formal principle, which attends to morphological properties of the classes, and the independence principle, which attends to the distribution

of class members. The principles were realized by a number of definitional criteria.

Canonicity is a property ascribable to actual constructions that are compared with the canon. The overall ranking of canonicity of a given conjugation class can be calculated in an impressionistic way by assigning a numerical value to the behavior of each class for each of the definitional criteria. In principle, this value can be assigned according to the following score: 1 is canonical and 0 is non-canonical.⁵⁷

However, canonicity is a gradable property. This means that the degree of canonicity of a structure is not just a question of a yes or no answer, but a question of degree. The same applies to the behavior of a certain class with respect to one of the criteria. In other words, for some criteria the behavior of the classes lie in a grey area. When this happens, it is difficult to come up with an exact numeral value, between 1 and 0, which may index the exact tinge of gray worn by the behavior of a given class. Our judgment here is based on qualitative factors.

In this article, all the gray areas, regardless of the intensity of the shading, will be given a value of 0.5. In other words, I will assign a 0.5 value to a given structure when its degree of canonicity lies somewhere in between being maximally canonical and being maximally non-canonical. The sum of the total number of criteria (10 in this case) represents maximal canonicity. This procedure works with the assumption that all criteria are equal in weight in the shaping of the canon.⁵⁸

The relevant results are given in Table 21. For convenience, I have just included the gist of each criterion. Further below, I discuss the assignment of values to each of the criteria one by one.

Criterion 1 has weak values because of the high degree of paradigmatic neutralization across the classes. The morphology involved in the paradigm of class I verbs is so erratically spread over the inflectional system that its marking is hardly unique to this class at all. In Table 19, the ratio of distinctive markers per overlapping ones was 1 to 14 for this class! The forms of class II and III are more consistent, but neither of them is very canonical. Most crucially, there are very few inflected forms in the entire system where the three classes are clearly distinguished by means of completely different markers. Despite this, I have given class III a value of 0.5 instead of a zero, because the ratio of overlapping markers in this class is the lowest of the three, with a one to one ratio.

Criterion 2 is the only one where the three conjugation classes of T-Oto rank together as canonical, as all of them realize the same morphosyntactic and morphosemantic distinctions. This does not mean that all members of such classes should have all the distinctions available. For example, class I has many stative verbs which cannot be inflected in certain tenses, for example the

Table 21. *The degree of canonicity of the three conjugation classes of T-Oto*

In a canonical situation . . .		I	II	III	System
Formal	1 Forms differ as consistently as possible across classes.	0	0	0.5	
	2 The classes realize the same distinctions.	1	1	1	
	3 Each member behaves identically.	0	0	1	
	4 Each paradigm cell is of equal status.	0	0	0.5	
Independence	5 The larger the number of members, the more canonical the class.	1	0	0	
	6 There is no phonological motivation.	1	1	1	
	7 There is no syntactic motivation.	1	0	1	
	8 There is no word-class motivation.	0	0	0	
	9 There is no pragmatic motivation.	1	1	1	
	10 There is no semantic motivation.	1	0	0.5	
Total		6	3	6.5	
Ratio of canonicity		6/10	3/10	6.5/10	15.5/30

past. Similarly a number of motion verbs of class I, such as *'ěhě* (intr/I) 'come', *kúhú* (intr/I) 'get in hither', *póhó* (intr/I) 'flow', *tsíts'i* (tr/I) 'bring a person', *tsú* (tr/I) 'bring', *tyũ* (tr/I) 'carry', etc. are only inflected in the local tenses. Such inflectional restrictions are not a characteristic or a handicap of the class, they respond to the lexical properties of specific members. Similarly, the classes rank with very low canonicity for criteria 3 and 4.

Criterion 3 has internal scope and implies that there should be no stem alternants or other subclasses within one given inflectional class. Here many verbs in classes I and II have stem alternants required by their inflection. This may be seen in the paradigms in Table 7 and 8, where the model verbs 'fall' and 'grind corn' occur in three different stems, respectively (e.g., *tági*, *tági*, *ntági* and *kúni*, *kúni*, *nkúni*). This alternation is not available for class III verbs. In this respect, class III ranks as canonical.

Criterion 4 dictates that for the canonical situation each cell is of equal status. If there are implications between cells in the form of paradigm-structure

conditions (Wurzel 1989 [1984]; Carstairs-McCarthy 1998; etc.) or of any other type, these can reduce the differences between classes, so that the paradigms are not kept maximally independent. In this respect, the erratic dispersion of the morphology realizing class I across classes II and III, regardless of how we account for the direction of leveling, makes this class rank non-canonically. Class II is a little more canonical than class I because if a given form has the first person formative *tú*, this will predict that the form for the second person will have *gú*, and vice versa. The correlation works for most cases except for the immediative irrealis. But all in all, there are only 16 instances in the paradigm where this could apply. Class III is the most canonical of the three. Here the occurrence of the formative *tí* in any form predicts that a *tí* will also be used for all persons.⁵⁹ For the inflection of the past tenses of class III (also the perfect) speakers could be further aided by a rule of referral that directs them to the inflection of class I verbs, at least for the first and second persons. Because of this, I have given class III a value 0.5.

Criterion 5 is about distribution. In the canonical situation, the larger the number of members of an inflectional class the more canonical that class. The equal share of three classes is 33.3% (100 divided by three). In my sample, 62.5% of the verbs belong to class I. This represents +29.2 points above the equal share, making class I canonical. In contrast, the other two classes are non-canonical in this respect. 11.6% of the verbs in the sample belong to class II, ranking -21.7 points below the equal share! On the other hand, while class III has more than a quarter of the verbs in the sample (25.8%), this share still ranks -7.5 points below the equal share.

The rest of the criteria suggest that in the canonical situation speakers should remember the assignment of a particular lexeme to a particular class without further clues. Weakening of the criteria may apply precisely when a given property associated with the lexeme serves as a token signaling its class membership. In this respect, while the three classes rank canonically according to criteria 6 and 9, they rank low in canonicity for criterion 8 because the classes are limited to verbs.⁶⁰ Class I and III are also canonical in respect of criterion 7. In contrast, class II is non-canonical because all its members are intransitive verbs.

Finally, criterion 10 dictates that a canonical class should not be semantically motivated. Class I is canonical in this respect. The same applies to the vast majority of the members of class III, but as middle verbs (marked with the middle prefix *n-*) are found in this class, the class is not entirely canonical. This is why I have given class III a value of 0.5 instead of 1. Most problematic in this respect is the status of class II.

In Section 6.1, we saw that the vast majority of the verbs in conjugation class II, whether basic or derived, depict actions performed by human beings: 63 verbs out of a total of 72, constituting the 87%. According to criterion 10,

the semantic homogeneity of class II would rank so low that it could even jeopardize its treatment as a genuine conjugation class. The alternative could be to treat such verbs as a subclass of class I verbs. There are at least three pieces of evidence that point to class II being a conjugation class.

Firstly, there is not a single marker in the paradigm of a class II verb that can be reasonably conceived of as functioning as an exponent of any agentive, intransitive or high animacy feature value. Secondly, in (19) we saw a number of instances where verbs of class II could be treated as derived from class I in the very same fashion as the Semitic *binyanim* advocated by Aronoff (1994). This phenomenon also happens between class I and III for most reciprocal verbs in (28), and between class II and III for the old causative pairs in (23). Similarly, verbs with agentive semantics are also found in other conjugation classes, so that agentivity or high animacy is not a feature value that forces membership to class II. For example *tsagi* (intr) ‘jump’, *tets’i* (intr) ‘go up’, *tsoni* (tr/intr) ‘chew’, *’ode* (intr) ‘hear, obey’, etc. belong to conjugation class I, while *k’ats’i* (tr/intr) ‘gather (cattle)’, *múnts’i* (tr/intr) ‘gather (corn), harvest’, *nõmi* (tr/intr) ‘water (plants)’, and others like them belong to class III. Finally, I have pointed out that class II also includes a number of verbs with non-human subjects. While these verbs are still few, they constitute 12% of the sample; a portion of the lexicon which may become significant as more members were added.

Table 11 also indicates the overall score of canonicity of the classes when all the values are put together. I take any score above 5 as being canonical, while below 5 is non-canonical.⁶¹ Then there is the question of degree, which is approached by the numerical value. The most canonical of the conjugation classes of T-Oto is class III: it misses maximal canonicity by only 3.5 points. Class III is closely followed by class I. In contrast, class II is typically non-canonical, as it does not only rank below 5, but misses maximal non-canonicity by just 3 points, almost the opposite of class III.

In principle, we could also take the ranking of the three conjugation classes together and evaluate them as a morphological system. Because of the low canonicity of class II, the overall ranking of the system is relatively low: it is roughly canonical because it barely scores over half of the spectrum (15.5 represents the 51% of 30).⁶²

8. Final remarks

In this article, I have studied in detail the three conjugation classes of T-Oto focusing both on their inflectional realizations and on the factors determining class membership. I have shown that conjugation class I works as a default major class, while classes II and III are less canonical inflectional classes

because their formal distinctiveness is based on elaborations based on the realizations of class I, and the vast majority of their members are not assigned arbitrarily, but present a number of properties that make probabilistic predictions possible. This is especially the case for class II. A number of such predictions are given in (33) in the form of implicational hierarchies similar to those proposed in Blevins (2006):

- (33)
- a. If a verb depicts a property concept or an inchoative state, it belongs to class I.
 - b. If a verb has reciprocal semantics, it is intransitive and belongs to class III.
 - c. If a verb has middle semantics, it is intransitive.
 - d. If a verb with middle semantics bears a nasal prefix, it belongs to class III; if it does not have a prefix, it will most likely belong to class II.
 - e. When an intransitive verb depicts a human agentive activity, it belongs to class II (only exceptionally could it belong to class III),
 - (i) if it has a *-te* ending, or
 - (ii) if it has the antipassive prefix *n-*, or
 - (iii) if it is a compound.
 - f. When an intransitive verb depicts a human agentive activity, and
 - (i) if its transitive pair belongs to class III, it then belongs to class III, or
 - (ii) if its transitive pair belongs to class I, it most likely belongs to class II.
 - g. If a verb is transitive, it belongs to either class I or class III.
 - h. If a transitive verb is a loanword from Spanish, it belongs to class III.

Morphologically, the classes of T-Oto constitute an interesting object of study for a theory of inflectional classes because they are made possible by loose inflectional formatives, which are neither bound affixes in the traditional sense nor stem extensions closer to the root. Similar contrasts appear in distantly related Chichimeco Jonaz as described in Angulo (1932), but the extent to which there are systemic similarities remains an open question for further research. Oto-Manguenan languages display rich morphological complexity, and conjugation classes appear to be one of the axes where such complexity is resolved, but much of the morphological treasures of these languages remain veiled and awaiting discovery.

Received 9 February 2011
Revised version received 7 July 2011

University of Surrey

Notes

1. This article was written under the auspices of the ESRC/AHRC project RES-062-23-3126 “Endangered Complexity: Inflectional classes in Oto-Manguean languages” as well as a collaboration to the project “Meso-American Morpho-Phonology” (MAMP) by the IUF (Institut Universitaire de France). First, I want to express my deepest gratitude to Néstor H. Green and Selene Hernández for helping me to compile the sample of verbs used in this article, without their help, the present work would have appeared in a later publication. To Benito Mendoza† and Petra Cruz for sharing their knowledge of their ancient language with us. I am also immensely grateful to Greville G. Corbett, Matthew Baerman, Dunstan Brown, Anna M. Thornton and Doris Bartholomew for their invaluable comments and suggestions on earlier drafts and to Jean-Léo Léonard for allowing me to collaborate with the MAMP project, thus allowing the retrieval of more data to validate the analysis. I especially want to thank the two anonymous referees, whose constructive criticism inspired me to make a number of important changes in both the presentation and the analysis which contributed substantially to the improvement of the article. I also want to thank Penelope Everson for proof-reading the English in the text quickly and efficiently. All errors and deficiencies remain my responsibility. Correspondence address: Surrey Morphology Group, Faculty of Arts and Human Sciences, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom. E-mail: e.palancar@surrey.ac.uk
2. A linguistic variety is defined as “(1)a variante lingüística . . . se define como una forma de habla que: a) presenta diferencias estructurales y léxicas en comparación con otras variantes de la misma agrupación lingüística; y b) implica para sus usuarios una determinada identidad sociolingüística, que se diferencia de la identidad sociolingüística de los usuarios de otras variantes.” (CLIN 2008: 37), [“a variety of speech (i) which has structural and lexical differences in comparison with other varieties within the same linguistic group, and (ii) which has a distinct sociolinguistic mark of identity for their users, different from the sociolinguistic identity born by speakers of other varieties.” (my translation)].
3. Santiago Tilapa is located at 2,768 m above sea level at coordinates 19° 11' 12.86" N and 99° 25' 18.30" W.
4. Consonants: *C* /Cʔ/; *C4* /FORTIS/ (word initially); ^h*C* /FORTIS/ (word internally); *f* /p^h/; *j* /k^h/; ^ʔ /ʔ/; *ñ* /ɲ/; *tx* /tʃ/; *r* /r/; *x* /ʃ/; *y* /j/. Vowels: *a* /a/ [ɔ]; *e* /e/; *o* /o/; *u* /i/; [~] nasal vowel. Abbreviations: < > represents an infix; = clitic; 1/2/3 grammatical persons; AS adjusted stem; AMBU ambulative; B bound shape of a verb; CONT continuous; DEF definitive; EMPH emphatic; EXCL exclusive; IMPF imperfect; INCL inclusive; *intr* intransitive verb; INCH inchoative; IRR irrealis; NEG negative; OBJ object; PL plural; PRES present; PST past; R realis; SG singular; SS secondary stem; ST stative; TRANSLOC translocative; *tr* transitive verb.
5. Verb stems like *túgi* ‘fall’ and *kúni* ‘grind corn’ are words with an onset fortis consonant /t/ and /k/. Fortis consonants are commonly realized by pre-aspirated voiceless allophones (i.e., [t^h] and [k^h]) when preceded by an open syllable, like the verbal stems in (9a) and (9b). They are not pre-aspirated when preceded by a close syllable, like in the subparadigm of the ambulative andative realis in Table 7 below. In turn, the stems *túgi* and *kúni* serve both as primary stems and as citation forms. See Example (12) and Note 7 for alternative stems.
6. The same softening occurs in the homophonous P&TAM clitic *tú* in Example (11).
7. The verbs *túgi* ‘fall’ and *kúni* ‘grind corn’ appear here in their ‘secondary stem’, which is built by lenisizing the onset fortis consonant, e.g., *túgi* and *kúni*. The secondary stem is used when the verb inflects for the third person (at times also the second person) in a number of specific tenses, the present irrealis being one of them. There is a third stem I treat as the “perfect stem” which is used for the form of third person in the perfect realis and the ambulative irrealis.

8. The sample is mainly based on a database of about 7,000 inflected forms from different lexemes. While for some lexemes, only a number of contrastive forms were collected to register their class, for many, the database has the entire paradigm. For the majority of the lexemes in the database there are a number of informative forms. The paradigms in Section 4 were constructed with the information provided in this database.
9. Transitive verbs are used here as a convenient cover category to include both monotransitive and ditransitive verbs.
10. The tenses I treat here as “imperfect distal realis” and “pluperfect realis”, and to a certain extent the “ambulative irrealis” are also commonly used with fronted adverbials. In this respect, their occurrence may respond to other types of morphosyntactic requirements which at the time of writing this article were not yet properly understood. Nevertheless, this does not affect the pertinence of the present analysis.
11. As pointed out in Section 3, the imperfect tenses are morphologically built on present tenses, and they both behave as a block for distributional purposes.
12. The formative has the phonologically conditioned alternant *tú* [tí], which results from the harmonization of /i/ in /tí/ to the central vowel /i/ of the syllabic nucleus of the preceding syllable.
13. Like in the other two present tenses in Table 9, the first and the second person carry a high tone here as well. The TAM marker *a* receives a high tone in the third person when preceded by the cislocative affix *b-*.
14. Historically, the labial affix was the marker *ba*, which in the third person was associated with the bare stem. In reality, a form such as *g^wá* for the second person resulted from a left-driven metathesis in an original sequence */g-á-ba/ > */g-á-b/ > */gá/ > [g^wá].
15. While the cislocative value is conveyed for the second and third person by the labial affix, for the first person, there is ambiguity between an ambulative and a cislocative reading except with option B of class I, which have *tí*.
16. The translocative of the third person is realized as follows: /B-bi/ > /bi/ > /bwi/ > [b^wi] (*b^wú*). The /i/ is centralized through assimilation to the velar component of the labialization.
17. In the sequence *tú* = *tú*, the lenis of the second formative, i.e., the person marker, is always realized as [d], e.g., *tú* = *dú*.
18. The old adverb survives in T-Oto as the adverbial clitic ‘already’. This can combine freely with other tenses to express anteriority, as for example in (i), where it is used with the imperfect continuous realis.

(i) *'ně a nána x=má='ǒh=á*
 and DEF.SG woman already=IMPF.CONT.R=[3.PRES]sleep.AS=PHRASE.FINAL
 ‘And the woman was already sleeping.’ (Txt)

19. Alternatively, /g/ and /b/ could be seen as the voiced phones [g] and [b] of a lenis /k/ and /t/ for initial position, whose voiceless (non-preaspirated) phones [k] and [t] are used before /j/. Such an account is more coherent with what appears to have happened historically.
20. This marker corresponds to the marker *xǎ*, already found in Cárceres. It may be an early weak pronunciation of the adverb *xǎ* ‘already’ (i.e., centralized) occurring with an inflected form with a zero P&TAM clitic. Alternatively, it may well result from associating *xǎ* to the inflected form of third person of the present habitual realis marked with a clitic *i*, which still survives as a present marker in Eastern Highlands Otomi (Voigtlander and Echegoyen 1979), which then underwent centralization.
21. As mentioned in Note 15, in the local tenses an /i/ gets centralized after the velar component of the labialization, e.g., /f = B-gí/ > /f = gi/ > /f = gwí/ > [f = k^wi] (*xk^wú*).
22. The present irrealis is also one of the most common encodings of the imperative.

23. The maintenance of maximal differentiation in the present irrealis may also respond to the high frequency of this tense.
24. Furthermore, in contrast to counterpart tenses of the realis, the present translocative irrealis is also used for the first person. Defectiveness in person in the local paradigms may respond to semantic pressures (e.g., it is more natural for the speaker to be able to talk about potential motion events that involve the speaker as subject rather than actual ones), but such a restriction does certainly not explain the defectiveness of the pattern in the past tense.
25. In Cárceres' description, Old Otomi is said to have two conjugation classes, which are called the *Tana* and the *Tati* conjugations. Palancar (2011) argues for a reinterpretation of the historical data, and provides evidence that the *Tana*-conjugation class in Cárceres in reality consists of two other classes, which are the historical ancestors of class I and II of T-Oto.
26. See Note 11 for the alternant *tú* [ti] of the formative *ti*.
27. In the translocative, the markers of the first and the second person surface as homophonous [g^wi] because of the centralization of /i/ before the labialization, (see Notes 15 and 19 for further examples). For the third person the marker surfaces as either *t^wú* [t^wi] or as *tú* [ti], with optional labialization. I take this marker to represent *ta* and not *ti*. Although both sources are equally possible, the leveling appears to be from class I to classes II and III, and not the other way around. This could be seen in that the markers used for the first and second persons are those of class I, and in that a *ta* marker is used for the third person of the present ambulative. Despite all this, the direction of the leveling in the past tense is unequivocally from classes II and III to class I.
28. This sequence always surfaces phonetically as [ti=g^wi] instead of [*t^wi=g^wi].
29. Lastra (1997) reports much variation for the encoding of past irrealis for the distant language of Ixtenco Otomi. This could also be a case of overabundance.
30. Thornton (2010, 2011) takes a canonical approach to overabundance. The case of T-Oto ranks highly canonical in Thornton's proposal because of various reasons: (a) the cell-mates are restricted to specific cells in a subparadigm (i.e., here only in the first and second persons); (b) they have a similar frequency in usage (i.e., speakers have an equal preference for one or the other, that is, they may produce one form at a given point, and then produce some other at a different time under equal circumstances); (c) they are not conditioned by any stylistic or pragmatic factors; and (d) they are morphomic, in that they are confined to the verbs of a specific inflectional class.
31. I thank Matthew Baerman for a fruitful discussion on this matter.
32. In Table 11, I showed that the ambulative cislocative realis has two cell-mates for the first and the second persons of class I verbs. In class II, the past irrealis has three cell-mates for the first person, and two for the second.
33. For class II, overlap with class I happens in 44 inflectional forms out of a total of 67, for class III, in 32 out of a total of 64.
34. For class I, we have overlap with either class II or class III, or with both, in 58 forms out of a total of 66.
35. A neutralization between class I and II in third person also occurs, e.g., the imperfect ambulative realis and the pluperfect.
36. I thank one of the anonymous referees for suggesting placing some emphasis on this remarkable behavior.
37. In past tenses of the realis and in the irrealis, the marking of third person in P&TAM clitics cannot be disassociated from the marking of TAM. In the present and imperfect tenses of the realis, it is zero.
38. Inchoative verbs select local tenses for their inflection, motivated by a metaphorical construal that views state change as change of location.

39. The class also contains two verbs derived from intransitive verbs, the compound and middle verb *m-p'g+ñö* (intr/II) 'crouch' (from '*mbg* 'stand' (intr/I) + *ñö* 'head') and the verb *n-tx'o* (intr/II) 'accompany', derived from '*yo* (intr/II) 'walk'.
40. The binyanim are traditionally regarded as a set of phonological templates. Aronoff views such templates as organized paradigms of different inflectional classes.
41. The element *skü-* is a diminutive prefix.
42. This verb is a compound with an opaque verb stem. The stem *ndé* is a nominal meaning 'mouth'.
43. Despite the action not involving a human being, this verb could be included in the list in (20) as involving the body. This was suggested by one of the anonymous referees.
44. The verb derives from *xqdi* (tr/I) 'teach'.
45. The transitive verb meaning 'kill' gave rise to the derived middle verb *n-jwádi* (intr/III) 'kill oneself'.
46. I lack the appropriate data to support whether the rule is still productive. The list could also include the verb *sélé* 'turn around', which is a stem of obscure origin that does not have native phonology (i.e., /s/ and /l/ are very rarely found in Otomi words).
47. All Otomi verbs have a similar meaning to the Spanish verb, except *andar* 'walk'.
48. For the diachronic relation between the middle voice prefix and the antipassive marker, see Palancar (2006).
49. The last two verbs in (26) derive from a transitive verb with a complex onset involving a nasal. The association of the nasal prefix is sanctioned by the phonotactics.
50. The element *kü* is a nominal meaning 'ear'.
51. See Note 39.
52. I have so far been only able to identify one middle verb of class II (*m-p'g+ñö* 'crouch') and two of class I (*m-peni* 'remember' and *m-pidi* 'get ill of a fright').
53. The nasal in these examples is interpreted as an old prefix and not as a phonological nasal in a complex onset because of the existence of cognate stems in other Otomi languages that do not bear the prefix.
54. Atmospheric verbs are also found in class I, e.g., '*wö* 'rain', '*mbg^hki* 'dawn', '*to^hni* 'thunder', '*zo^l-nte* 'go dark at dusk', '*zo^l-xü* 'get dark', etc.
55. The semantics of the verb *hátsi* is equivalent to that of the verb *amanecer* 'dawn' in Spanish, which can be used to refer to the act of waking up at the start of one's day, commonly in a given concomitant state.
56. The verb derives from the noun *jüni* 'cloud'.
57. An alternative approach to measure degree of canonicity is found in Brown et al. (2009).
58. I want to thank Dunstan Brown for this observation. While in principle we could make some of the criteria play a more significant role than others in the shaping of what is canonical, I wanted to make all things equal in order to avoid the risk of imposing any bias on the canon.
59. This happens in all cases except in the present irrealis, where the third person is not **titi* but just *ti*.
60. This is, nonetheless, crosslinguistically expected. Corbett (2009: 7) himself acknowledges that this criterion is likely to be met rarely. Although conjugation classes are commonly restricted to verbs only, there are also certain inflectional classes that rank more canonical in this respect, like the ones involving nouns and adjectives in some Indo-European languages.
61. Had the number of criteria been nine, the flexing point for canonicity would have then been 4.5.
62. However, this would perhaps not be adequate because a morphological system of inflectional classes constitutes a different ontology than the inflectional classes it is made of and should perhaps be ranked according to its own different canon defined by its own principles, including for example an evaluation of the formal contrasts among the existing classes.

References

- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Andrews, Henrietta. 1993. *The function of verb prefixes in Southwestern Otomi* (Summer Institute of Linguistics and University of Texas at Arlington Publications in Linguistics 115). Dallas, TX: Summer Institute of Linguistics and University of Texas at Arlington.
- de Angulo, Jaime. 1932. The Chichimeco language (Central Mexico). *International Journal of American Linguistics* 7. 152–193.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes* (Linguistic Inquiry Monograph 22). Cambridge, MA: MIT Press.
- Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.
- Bartholomew, Doris. 2004. Notas sobre la gramática. In Luis Hernández Cruz, Moisés Victoria Torquemada & Donald Sinclair Crawford (eds.), *Diccionario del hñähñu (otomi) del Valle del Mezquital, Estado de Hidalgo* (Vocabularios Indígenas 45), 473–495. Mexico City: Instituto Lingüístico de Verano.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. III: Grammatical categories and the lexicon*, 169–240. 2nd edn. Cambridge: Cambridge University Press.
- Blevins, James. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573.
- Blevins, Juliette. 2005. Yurok verb classes. *International Journal of American Linguistics* 71(3). 327–349.
- Brown, Dunstan, Carole Tiberius, Marina Chumakina, Greville G. Corbett & Alexander Krašovitsky. 2009. Databases designed for investigating specific phenomena. In Martin Everaert, Simon Musgrave & Alexis Dimitriadis (eds.), *The use of databases in cross-linguistic studies* (Empirical Approaches to Language Typology 41). Berlin & New York: Mouton de Gruyter.
- de Cárceres, Fray Pedro. 1907 [1580]. *Arte de la lengua Othomi*. Edited by Nicolás León. *Boletín del Instituto Bibliográfico Mexicano* 6. 38–155.
- Carstairs-McCarthy, Andrew. 1998. Paradigmatic structure: inflectional paradigms and morphological classes. In Andrew Spencer & Arnold Zwicky (eds.), *The handbook of morphology*, 322–334. Oxford: Blackwell.
- CLIN (Catálogo de las Lenguas Indígenas Nacionales): Variantes lingüísticas de México con sus autodenominaciones y referencias geostadísticas. 2008. *Diario de la Nación* (14–01–2008). Mexico DF. http://www.inali.gob.mx/pdf/CLIN_completo.pdf.
- Comrie, Bernard. 2003. When agreement gets trigger-happy. *Transactions of the Philological Society* 101(2). 313–337.
- Corbett, Greville G. 2005. The canonical approach in typology. In Zygmunt Frajzyngier, David Rood & Adam Hodges (eds.), *Linguistic diversity and language theories*, 25–49. Amsterdam & Philadelphia: John Benjamins.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2007. Canonical typology, suppletion and possible words. *Language* 83. 8–42.
- Corbett, Greville G. 2009. Canonical inflectional classes. In Fabio Montermini, Gilles Boyé & Jesse Tseng (eds.), *Selected proceedings of the 6th décembrettes*, 1–11. Somerville, MA: Cascadia Proceedings Project. <http://www.lingref.com/cpp/decemb/6/index.html>.
- Corbett, Greville & Matthew Baerman. 2006. Prolegomena to a typology of morphological features. *Morphology* 16. 231–246.
- Donohue, Mark. 2001. Animacy, class and gender in Burmeso. *Pacific Linguistics* 514. 95–115.

- Echegoyen, Artemisa & Katherine Voigtlander. 2007. *Diccionario yuhú: Otomí de la Sierra Madre Oriental (Estados de Hidalgo, Puebla y Veracruz, México)* (Vocabularios Indígenas: Preliminary version). Mexico City: Instituto Lingüístico de Verano.
- Evans, Nicholas. 2003. Typologies of agreement: Some problems from Kayardild. *Transactions of the Philological Society* 101(2), 203–234.
- Finkel, Raphael & Gregory Stump. 2009. Principal parts and degrees of paradigmatic transparency. In James P. Blevins & Juliette Blevins (eds.), *Analogy in grammar: Form and acquisition*, 13–53. Oxford: Oxford University Press.
- Hekking, Ewald. 1995. *El otomí de Santiago Mexquititlán: Desplazamiento lingüístico, préstamos y cambios gramaticales* (Studies in Language and Language Use 17). Amsterdam: Institute for Functional Research into Language and Language Use.
- Hernández Cruz, Luis, Moisés Victoria Torquemada & Donaldo Sinclair Crawford (eds.). 2004. *Diccionario del hñähñu (otomí) del Valle de Mezquital, Estado de Hidalgo* (Vocabularios Indígenas 45). Mexico City: Instituto Lingüístico de Verano.
- Hess, H. Harwood. 1968. *The syntactic structure of Mezquital Otomi* (Janua Linguarum Series Practica 43). The Hague: Mouton.
- Kaufman, Terrence. 1990. The reconstruction of TAM in Proto Oto-Manguan. Unpublished ms.
- Kemmer, Suzanne. 1993. *The middle voice* (Typological Studies in Language 23). Amsterdam & Philadelphia: John Benjamins.
- Lastra, Yolanda. 1997. *El otomí de Ixtenco*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Lastra, Yolanda. 2001. *Unidad y diversidad de la lengua: Relatos otomíes*. Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- Maiden, Martin. 1992. Irregularity as a determinant of morphological change. *Journal of Linguistics* 28(2), 285–312.
- Maiden, Martin. 2005. Morphological autonomy and diachrony. *Yearbook of Morphology* 2004, 137–175.
- Pace, Wanda Jane. 1990. Comaltepec Chinantec verb inflection. In William R. Merrifield & Calvin R. Rensch (eds.), *Syllables, tones and verb paradigms* (Studies in Chinantec Languages 4), 21–62. Arlington, Texas: The Summer Linguistic Institute and the University of Texas at Arlington.
- Palancar, Enrique L. 2006. Middle voice and intransitivity in Otomi. *Linguistics* 43(1), 613–643.
- Palancar, Enrique L. 2009a. *Gramática y textos del hñöhñö: Otomí de San Ildefonso Tultepec, vol. I: Gramática*. Mexico City: Plaza y Valdés.
- Palancar, Enrique L. 2009b. The TAM morphology in Tilapa Otomi verb conjugations. Paper presented at Morphology of the World's Languages I, University of Leipzig, 11–13 June.
- Palancar, Enrique L. 2011. The conjugations of Colonial Otomi. *Transactions of the Philological Society* 109, 1–19.
- Polinsky, Maria. 2003. Non-canonical agreement is canonical. *Transactions of the Philological Society* 101(2), 279–312.
- Smith-Stark, Thomas C. 2002. Las clases verbales del zapoteco de Chichicapan. In Rosa María Ortiz Ciscomani & Zarina Estrada Fernández (eds.), *VI Encuentro Internacional de Lingüística en el Noroeste: Memorias*, vol. 2, 165–212. Hermosillo, Sonora: UniSON.
- Spencer, Andrew. 2005. Extending deponency. In Matthew Baerman, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.), *Deponency and morphological mismatches* (Proceedings of the British Academy 145), 45–70. Oxford: British Academy & Oxford University Press.
- Stump, Gregory T. 2005. A non-canonical pattern of deponency and its implications. In Matthew Baerman, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.), *Deponency and morphological mismatches* (Proceedings of the British Academy 145), 71–95. Oxford: British Academy & Oxford University Press.

- Stump, Gregory and Raphael Finkel. 2007. Principal parts and morphological typology. *Morphology* 17. 39–75.
- Thornton, Anna M. 2010. Towards a typology of overabundance. Paper presented at Décembrettes 7: International Conference on Morphology, University of Toulouse, 2–3 December.
- Thornton, Anna M. 2011. “Overabundance (multiple forms realizing the same cell): a non-canonical phenomenon in Italian verb morphology”. In Maiden, Martin, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds.), *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*. Oxford, Oxford University Press, 2011, 358–381.
- Voigtlander, Katherine & Artemisa Echegoyen. 1979. *Luces contemporáneas del otomí*. Mexico City: Instituto Lingüístico de Verano.
- Wurzel, Wolfgang U. 1989 [1984]. *Inflectional morphology and naturalness* (Studies in Natural Language and Linguistic Theory). Translated by Manfred Schentke. Dordrecht: Kluwer.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington, IN: Indiana University Linguistics Club.

THE CONJUGATIONS OF COLONIAL OTOMI

By ENRIQUE L. PALANCAR
Surrey Morphology Group, University of Surrey

ABSTRACT

Pedro de Cárceres, a Spanish friar from New Spain (Mexico), wrote the first grammar of Otomi (Oto-Pamean, Oto-Manguéan) in 1580. The stage of the Otomi language reflected in Cárceres is called ‘Colonial Otomi’ (sixteenth and seventeenth centuries). Later, dialectal diversification gave rise to a family of different languages. Cárceres describes Colonial Otomi as having two verb conjugations. However, modern languages displaying great morphological innovation (e.g. Mezquital Otomi) show no trace of such conjugations. This led some authors to discredit Cárceres’ work by portraying it as confusing and inaccurate. In this article, I show that the analysis by Cárceres is fundamentally right, and I propose that many of the inconsistencies in his work can be solved by positing three conjugations instead of two, especially in the light of the evidence from modern languages like Tilapa Otomi, which has preserved three such conjugations.

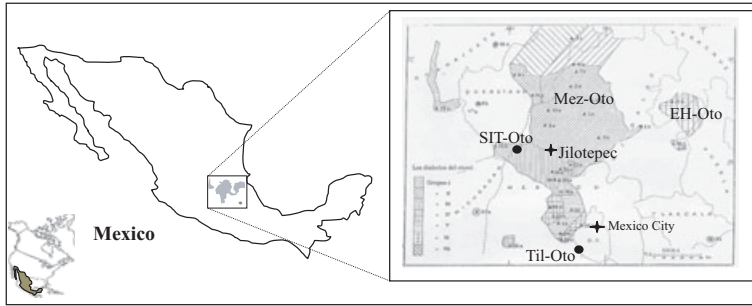
1. INTRODUCTION¹

In this paper, I revisit the analysis by Pedro de Cárceres (1580/1907) regarding the existence of two verbal conjugations in Colonial Otomi. I propose the existence of a third one. Positing three conjugations, instead of two, resolves the inconsistencies already detected in previous works, and the historical data are made more coherent with the descriptive facts of a number of modern Otomi languages.

Otomi (Oto-Pamean; Oto-Manguéan) is a family of Mesoamerican Indian languages spoken in Central Mexico by about a quarter of a million people. Map 1 indicates the location of the four modern languages I mention in this article: San Ildefonso Tultepec Otomi (SIT-Oto), Mezquital Otomi (Mez-Oto), Eastern Highlands Otomi (EH-Oto), and Tilapa Otomi (Til-Oto).

While the internal linguistic diversity of Otomi is uncontroversial (summarized in Lastra 2001; 2006), there is no established consensus as to which of the different linguistic subsystems identified so far as Otomi should be treated as independent languages and which as dialects of those languages. This is why Otomi has traditionally been described as forming a dialectal continuum rather than a family, and why authors in Mexico refrain from using terms such as ‘language’ and ‘dialect’ and prefer ‘variety’ as a convenient cover term. In contrast to this tradition, the current, official stance of the National Institute of Indigenous Languages (INALI) is to treat Otomi innovatively as forming a ‘linguistic group’ with nine different ‘linguistic varieties’ (CLIN 2008). INALI’s linguistic varieties are treated as genuine languages

¹A preliminary version of this article was read at the Language, Text and History Workshop at the University of Cambridge in March 2010. I want to thank the organizers, Wendy Ayres-Bennett and Sylvia Adamson, for encouraging me to submit this work for publication in *TPhS*. I also want to thank the three anonymous referees for their invaluable comments and suggestions and for taking the time to read through the Old Otomi data in painstaking detail. All errors and deficiencies remain my responsibility.



Map 1. The Otomi languages of Mexico (from Soustelle 1937: 597)

for official – i.e. educational, administrative, judicial, social and informative – purposes. This is also the stance I take here.²

Otomi emerged as a linguistic family relatively recently; we can date some of the dialectal isoglosses from about the second half of the seventeenth century. The dialectal spread that took place was favoured by the isolation of the linguistic communities under the new geopolitical order imposed by the Spaniards on the conquest territory as the colonial territory of New Spain.

Pedro de Cárceres was a Franciscan friar who lived in New Spain in the sixteenth century. He was the Guardian of the Monastery of the Holy Cross in Querétaro, where he worked as a preacher of the new faith. Most relevant for our purposes, he was also the author of the grammar titled *Arte de la Lengua Othomi* ('Art of the Otomi Language'), completed in 1580. Missionary grammars were called *Artes* because grammars were seen as the *art* of speaking a language properly. This grammar was never published because the printers could not afford to cast the new stamps of letters for a variety of novel characters that Cárceres uses to represent the different sounds of Otomi, which he judged could not be rendered accurately by the Latin alphabet.³ The grammar survived up to the beginning of the twentieth century in the form of two manuscripts. Nicolás León, an educated bibliophile physician, having been granted access to both of them, edited and published in 1907 the one manuscript he thought had been written by Cárceres himself. The two manuscripts were later lost, possibly in the years of the Mexican Revolution (which began in 1910).

The *Art* by Cárceres is the first grammar of Otomi as a common language, and has a considerable historical value for this reason. Other works such as Urbano (1605/1990) were based on this work. The grammar describes aspects of the historical stage of Otomi commonly known as Colonial or Old Otomi, which was spoken for about two centuries between 1500 and 1700 before dialectal diversification. The dialect of Colonial Otomi described in Cárceres' grammar corresponds to a variety spoken in the large northern area under the influence of the political centre of Jilotepec (see Map 1); an important Otomi state in pre-Columbian times. This may be seen in the phonological shape of many words which represent innovative

²A linguistic variety is defined as 'a variety of speech (i) which has structural and lexical differences in comparison with other varieties within the same linguistic group, and (ii) which has a distinct sociolinguistic mark of identity for their users, different from the sociolinguistic identity born by speakers of other varieties' (translation mine). Original text in *CLIN* (2008: 37), 'La variante lingüística [...] se define como una forma de habla que: a) presenta diferencias estructurales y léxicas en comparación con otras variantes de la misma agrupación lingüística; y b) implica para sus usuarios una determinada identidad sociolingüística, que se diferencia de la identidad sociolingüística de los usuarios de otras variantes.'

³This was a general problem for the printers of New Spain (Garone Gravier 2010).

isoglosses of the area and in the use of toponyms of nearby places. Other dialects of Colonial Otomi with no historical attestation are the Southern and the Eastern varieties.⁴

2. VERBAL TENSES IN COLONIAL OTOMI

As in other Oto-Pamean languages, categories of tense/aspect/mood (TAM) in Colonial Otomi are grammatically encoded by means of a complex array of proclitics and prefixes.

Table 1. The principal grammatical tenses of Colonial Otomi

	Tense/aspect	Subject	Verb form	Translatable as
realis	Present (habitual) (<i>Present indicative</i>)	1st	<i>tana-xohnabate</i>	'I teach'
		2nd	<i>cana-xohnabate</i>	'you teach'
		3rd	<i>na-xohnabate</i>	'he teaches'
	Present continuous (<i>Another present indicative</i>)	1st	<i>ti-xohnabate</i>	'I am teaching'
		2nd	<i>qui-xohnabate</i>	'you're teaching'
		3rd	<i>(i-)xohnabate</i>	'he's teaching'
	Imperfect (habitual) (<i>Imperfect preterite ind.</i>)	1st	<i>tomana-xohnabate</i>	'I used to teach'
		2nd	<i>comana-xohnabate</i>	'you used to teach'
		3rd	<i>mana-xohnabate</i>	'he used to teach'
	Imperfect continuous (<i>Another imp. pret. ind.</i>)	1st	<i>tomi-xohnabate</i>	'I was teaching'
		2nd	<i>comi-xohnabate</i>	'you were teaching'
		3rd	<i>mi-xohnabate</i>	'he was teaching'
	Past (or Preterite) (<i>Perfect preterite indicative</i>)	1st	<i>tota-xohnabate</i>	'I taught'
		2nd	<i>coca-xohnabate</i>	'you taught'
		3rd	<i>pin-xohnabate</i>	'he taught'
	Perfect cislocative (<i>Another perf. pret. ind.</i>)	1st	<i>ta-xohnabate</i>	'I came to teach'
		2nd	<i>ca-xohnabate</i>	'you came to teach'
		3rd	<i>a-xohnabate</i>	'he came to teach'
Pluperfect (<i>Plusquamperfect pret. ind.</i>)	1st	<i>(xo)tata-xohnabate</i>	'I had taught'	
	2nd	<i>cocaca-xohnabate</i>	'you had taught'	
	3rd	<i>(xo)ca-xohnabate</i>	'he had taught'	
irrealis	Present (<i>imperfect future indicative</i>)	1st	<i>cata-xohnabate</i>	'I'll teach'
		2nd	<i>caca-xohnabate</i>	'you'll teach'
		3rd	<i>tin-xohnabate</i>	'he'll teach'
	Present continuous (<i>Another imperf. fut. ind.</i>)	1st	<i>cati-xohnabate</i>	'I'll be teaching'
		2nd	<i>cati-xohnabate</i>	'you'll be teaching'
		3rd	<i>tin-xohnabate</i>	'he'll be teaching'
	Present cislocative (<i>Another imperf. fut. ind.</i>)	1st	<i>cata-xohnabate</i>	'I'll come to teach'
		2nd	<i>cata-xohnabate</i>	'you'll come to teach'
		3rd	<i>ta-xohnabate</i>	'he'll come to teach'
	Imperfect A (<i>imperfect preterite optative</i>)	1st	<i>comana-xohnabate</i>	'(if) I were to teach'
		2nd	<i>comana-xohnabate</i>	'(if) you were to teach'
		3rd	<i>tomana-xohnabate</i>	'(if) he were to teach'
	Imperfect B	1st	<i>comata-xohnabate</i>	
		2nd	<i>comaca-xohnabate</i>	
		3rd	<i>toma-xohnabate</i>	
	Imperfect continuous (<i>Another imperfect pret. opt.</i>)	1st	<i>comin-xohnabate</i>	'(if) I were teaching'
		2nd	<i>comin-xohnabate</i>	'(if) you were teaching'
		3rd	<i>tomin-xohnabate</i>	'(if) he were teaching'
	Past (or preterite) (<i>Perfect preterite optative</i>)	1st	<i>cota-xohnabate</i>	'I'd have taught'
		2nd	<i>coca-xohnabate</i>	'you'd have taught'
		3rd	<i>toco-xohnabate</i>	'he'd have taught'
	Pluperfect cislocative (<i>Another perfect pret. opt.</i>)	1st	<i>co(ca)ta-xohnabate</i>	'I'd come to teach'
		2nd	<i>co(ca)ta-xohnabate</i>	'you'd come to teach'
		3rd	<i>tocota-xohnabate</i>	'he'd come to teach'

Note: The imperfect preterite optative has two sub-paradigms in the Colonial grammar, which I have called imperfect A and B. Nothing more is said about their distribution or function.

⁴As the phenomenon I study here is a diasystemic feature of the old language, I will not elaborate further on what may have characterised these dialects.

As the traditional rendition of such inflectional formatives is as if they all were prefixes, I will not attempt to make any further formal distinctions here.⁵ Often, the TAM prefixes also encode reference to the person of the subject. In Table 1, I introduce the most representative TAM paradigms described by Cárceres in his grammar for the intransitive verb *xohnabate* /ʃ^hṇapaṭe/ ‘teach, preach’.⁶ The tenses are set out according to a modern reinterpretation of the data, based on existing formal contrasts and on the function the historical reflex of such tenses have in the modern languages (see e.g. Palancar 2009).⁷

In Table 1, we can see that the tenses are organised according to a mood distinction based on a realis/irrealis contrast, with an equal number of tenses in each mood. In the realis, there are two presents (and two imperfects). These correspond roughly to a distinction between a tense conveying a habitual action (e.g. ‘I write’) and another conveying actions that are taking place at the very moment of the speech act situation, which present some degree of continuative aspect (e.g. ‘I am writing’). The modern Otomi languages of the south have preserved this morphological distinction. However, most of the modern languages have reduced the tense inventory in the irrealis mood. In general, the more morphological tenses a language has, the more conservative it is, morphologically speaking.

For the sake of brevity, in Table 1 I have not analysed the inner morphological structure of the prefixal material in boldface. A number of analyses are viable. In the present (habitual) realis for example, the affixes *tana-*, *cana-* and *na-* all have the element *na-* in common, a fact that suggests that the prefixes *ta-*, *ca-* and *Ø-* could be interpreted as exponents of grammatical person, leaving *na-* to encode other functions, perhaps just tense or aspect. However, the formatives *ti-*, *qui-* and *i-* or *Ø-* of the present continuous suggest that an analysis of *t-* /t̥/ and *c-/qu-* /k̥/ as exponents of 1st and 2nd person is perhaps more viable, leaving the prefixes *a-* and *i-* to encode other functions. In the tenses of the irrealis, there is neutralization between 1st and 2nd person in favour of *c-* /k̥/ marker.

2.1. Verbal conjugations in Colonial Otomi

In Cárceres’ analysis, Colonial Otomi verbs are organized in two large verb classes or conjugations, which are called the ‘*TANA*’ conjugation and the ‘*TATI*’ conjugation. The names *TANA* and *TATI* respond mnemonically to the grammatical exponence used for the 1st person of the present realis. This is illustrated in (1) with instances of the two verbs consistently used in the grammar to illustrate the paradigms, the intransitive verbs *xohnabate* ‘teach’ and *māte* ‘love’. While these verbs are intransitive, both conjugations also include transitive verbs.

⁵I use the term ‘formative’ in the same sense as in Bickel & Nichols (2007) as a cover term for any marker of inflectional information.

⁶The phonological transcription of Colonial Otomi is an informed reconstruction. I use the extensions to the IPA symbols for fortis Ć and lenis Ć̣. The acute accent represents high tone, while the circumflex indicates raising tone. The orthographic usage in Cárceres is as follows: <*c*> before <*a, o*> and <*qu*> before <*i*> and <*e*> represent lenis /k̥/ or fortis /ḳ/; <*t*> lenis /t̥/ or fortis /ṭ/; <*k*> glotalized /ḳʔ/; <*x*> /ʃ/. Nasal vowels are marked with a diacritic, e.g. <*ã*> /ã̃/. The abbreviations to the morpheme-by-morpheme glosses to the Otomi examples follow the convention of the Leipzig Glossing Rules: 1/2/3: grammatical persons, A: morphophonologically adjusted form, AMBU: ambulative, ANIM: Animate, IMPER: Impersonal, INCHO: Inchoative, IRR: irrealis, OBJ: Object, PASS: Passive, PRES: Present, R: realis.

⁷In the paradigm of Table 1, I have not included forms with local prefixes conveying spatial information about the location of the subject (e.g. perlocative *ni-*, distal *ba-*, ‘down’ *ba-na-*), whose reflexes are still found in Eastern Highlands Otomi in what Voigtlander & Echegoyen (1979) call the ‘local aspect’.

1. 1st person realis

- a. TANA conjugation **tā-nā**-xohnābate 'I teach'
 b. TATI conjugation **tā-tī**-māti 'I love, I am loving'⁸

If we remove the paradigms of the pluperfect, which are incomplete, together with the set of tenses involving location and motion about which my understanding is still poor,⁹ we can establish a number of formal contrasts between the exponence of these two inflectional classes. The morphological contrasts that we observe by comparing the means of exponence of the two inflectional classes are the following:

(i) We have four tenses whose paradigms show a similar amount of marking in the two classes, e.g. in (2). Here both classes have paradigms with a set of prefixes in common, to which other more specific inflectional prefixes are added, i.e. *na-* for the *TANA* conjugation and *ti-* for the *TATI* conjugation. I use the term 'base' unconventionally to refer to any set of prefixes shared between different classes as a way of emphasising the inflectional morphology they have in common. As I show in §4, these common prefixes are the inflectional exponents of a third class of verbs.

2. Base	<i>TANA</i>		<i>TATI</i>
	Present (habitual) realis		
ta-	1st ta- na -xohnabate	'I teach'	ta- ti -māti 'I love'
ca-	2nd ca- na -xohnabate	'you teach'	ca- ti -māti 'you love'
Ø-	3rd Ø- na -xohnabate	'he teaches'	Ø- ti -māti 'he loves'
	Imperfect (habitual) realis		
to-ma-	1st to-ma- na -xohnabate	'I used to teach'	to-ma- ti -māti 'I used to love'
co-ma-	2nd co-ma- na -xohnabate	'you used to teach'	co-ma- ti -māti 'you used to love'
Ø-ma-	3rd Ø-ma- na -xohnabate	'he used to teach'	Ø-ma- ti -māti 'he used to love'
	Present irrealis		
ca-	1st ca- ta -xohnabate	'I'll teach'	ca- ti -māti 'I'll love'
ca-	2nd ca- ca -xohnabate	'you'll teach'	ca- ti -māti 'you'll love'
ti-	3rd ti- n -xohnabate	'he'll teach'	ti-māti 'he'll love'
	Imperfect B irrealis		
co-ma-	1st co-ma- ta -xohnabate	'(if) I were to teach'	co-ma- ti -māti '(if) I were to love'
co-ma-	2nd co-ma- ca -xohnabate	'(if) you were to teach'	co-ma- ti -māti '(if) you were to love'
to-ma-	3rd to-ma-xohnabate	'(if) he were to teach'	to-ma- ti -māti '(if) he were to love'

(ii) There is only one tense where the two conjugations have exactly the same base paradigm.

3. Base	<i>TANA</i>		<i>TATI</i>
	Imperfect continuous irrealis		
co-mi-n-	1st co-mi-n-xohnabate	'(if) I were teaching'	co-mi-n-māti '(if) I were loving'
co-mi-n-	2nd co-mi-n-xohnabate	'(if) you were teaching'	co-mi-n-māti '(if) you were loving'
to-mi-n-	3rd to-mi-n-xohnabate	'(if) he were teaching'	to-mi-n-māti '(if) he were loving'

(iii) Despite (2) and (3), in at least five tenses the paradigm of the *TATI* class is morphologically constructed on the *TANA* conjugation by adding the prefix *ti-*. In these cases, the means of exponence of the *TANA* conjugation serve as the default case, as shown in (4).

⁸The form in (1b) depicts the act of being a loving person, rather than loving someone in particular.

⁹See fn. 7 above.

4. Base	<i>TANA</i>			<i>TATI</i>	
			Present continuous realis		
ti-	1st	ti-xohnabate	‘I am teaching’	ti- ti -māti	‘I am loving’
qui-	2nd	qui-xohnabate	‘you’re teaching’	qui- ti -māti	‘you’re loving’
(i-)	3rd	(i-)xohnabate	‘he’s teaching’	(i-) ti -māti	‘he’s loving’
			Perfect cislocative realis		
ta-	1st	ta-xohnabate	‘I came to teach’	ta- ti -māti	‘I came to love’
ca-	2nd	ca-xohnabate	‘you came to teach’	ca- ti -māti	‘you came to love’
a-	3rd	a-xohnabate	‘he came to teach’	a- ti -māti	‘he came to love’
			Imperfect continuous realis		
to-mi-	1st	to-mi-xohnabate	‘I was teaching’	to-mi- ti -māti	‘I was loving’
co-mi-	2nd	co-mi-xohnabate	‘you were teaching’	co-mi- ti -māti	‘you were loving’
Ø-mi-	3rd	Ø-mi-xohnabate	‘he was teaching’	Ø-mi- ti -māti	‘he was loving’
			Present continuous irrealis		
ca-ti-	1st	ca-ti-xohnabate	‘I’ll be teaching’	ca-ti- ti -māti	‘I’ll be loving’
ca-ti-	2nd	ca-ti-xohnabate	‘you’ll be teaching’	ca-ti- ti -māti	‘you’ll be loving’
ti-	3rd	ti- n -xohnabate	‘he’ll be teaching’	ti- ti -māti	‘he’ll be loving’
			Present cislocative irrealis		
ca-ta-	1st	ca-ta-xohnabate	‘I’ll come to teach’	ca-ta- ti -māti	‘I’ll come to love’
ca-ta-	2nd	ca-ta-xohnabate	‘you’ll come to teach’	ca-ta- ti -māti	‘you’ll come to love’
ta-	3rd	ta-xohnabate	‘he’ll come to teach’	ta- ti -māti	‘he’ll come to love’

(iv) The paradigms in the two tenses in (5) show exactly the opposite tendency to the one given in (4). Here the means of exponence of the *TATI* conjugation serve as the default case on which the paradigms of the *TANA* conjugation are based by adding a variety of other prefixes, including *ta-* and *ca-*, which appear to be reduplicative formatives, a nasal *n-* and the prefix *na-*, shown in other tenses of this conjugation, as in (2) above.

5. Base	<i>TANA</i>			<i>TATI</i>	
			Past realis		
to-	1st	to- ta -xohnabate	‘I taught’	to-māti	‘I loved’
co-	2nd	co- ca -xohnabate	‘you taught’	co-māti	‘you loved’
pi-	3rd	pi- n -xohnabate	‘he taught’	pi-māti	‘he loved’
			Imperfect A irrealis		
co-ma-	1st	co-ma- na -xohnabate	‘(if) I were to teach’	co-ma-māti	‘(if) I were to love’
co-ma-	2nd	co-ma- na -xohnabate	‘(if) you were to teach’	co-ma-māti	‘(if) you were to love’
to-ma-	3rd	to-ma- na -xohnabate	‘(if) he were to teach’	to-ma-māti	‘(if) he were to love’

(v) Finally, there is one tense where the marking pattern is uneven in the paradigms of both conjugations, as shown in (6). In the tense in (6), the *TANA* conjugation has the reduplicative prefixes operating in the past realis, as illustrated in (5), while the *TATI* conjugation departs from the base by having the prefix *que-* /*ḱe*/ in the 2nd person.

6. Base	<i>TANA</i>			<i>TATI</i>	
			Past irrealis		
co-	1st	co- ta -xohnabate	‘I’d have taught’	co-māti	‘I’d have loved’
co-	2nd	co- ca -xohnabate	‘you’d have taught’	co- que -māti	‘you’d have loved’
to-co-	3rd	to-co-xohnabate	‘he’d have taught’	to-co-māti	‘he’d have loved’

The contrasts illustrated in (2)–(5) indicate that both inflectional classes are two independent paradigmatic sets that keep the morphological formatives involved in their realization maximally distinct by showing very few overlaps as a whole. On some occasions, one of the conjugations appears to be morphologically more basic than the other; on some other occasions, the reverse is true. The presence of the prefix *ti-* characterizes the paradigm of the *TATI* conjugation, while the presence of *na-* in a given verb form may help to identify its lexeme as belonging to the *TANA* conjugation.

2.2. Inconsistencies within the conjugations

Despite the apparent consistencies of the inflection given in §3.1 below, many verbs in Cárceres' grammar do not readily fit such paradigms. For example, many verbs said to belong to the *TANA* conjugation do not take the expected inflection in some tenses. This happens for example with all transitive verbs (but also with a number of intransitive ones) of *TANA*, which in the 1st and 2nd persons of the present realis do not take the prefix *na-*, as shown in (7), while the 3rd person behaves as expected, e.g. (8).¹⁰

- | | | |
|-------|--|-----------------------------|
| | Conjugation <i>TANA</i> | Expected for present realis |
| 7. a. | tā-xohnikī
/tā-ʃ ^h ni-k'i/
<i>1.PRES.R-teach-2OBJ</i>
'I teach you.' (p. 73) | * tā-na-xohnikī |
| | b. | cā-ēntǰigui |
| | /ká-ʔēntǰi-ki/
<i>2.PRES.R-lift-1OBJ</i>
'You lift me up.' (p. 89) | * cā-na-ēntǰigui |
| 8. | nā-huixīguī
/ná-hwǰ ⁱ -ki/
<i>3.PRES.R-push.A-1OBJ</i>
'They push me around.' (p. 90) | nā-huixīguī |

This discrepancy is somewhat inconvenient, because the inflection of the present realis is the paradigmatic key for determining membership of the *TANA* conjugation. Nevertheless, it also appears that transitive verbs in the 1st and the 2nd person can equally take the prefix *na-* in a present tense.¹¹ However, when they do so, the verb now conveys motion information about the subject: it depicts the subject doing the action as he is moving along. I call this tense 'ambulative', and for convenience it has not been included in the paradigms above. An example is given in (9).¹²

9. **tānaxohnibi**
/tā_{na}-ʃ^hni-pi/
1.PRES.R.AMBU-teach-3OBJ.ANIM
'I teach him as we go.' (p. 90)

¹⁰The page number in the examples in the paper indicates the page where the examples can be found in Cárceres (1580/1907).

¹¹At this point, the grammar has only the transitive verb *xohni* 'teach, preach'. However, this is also the case for most paradigms, and I take this tense to be an inflectional property extendable to all transitive verbs, providing, of course, that the actions depicted bear a certain degree of semantic compatibility with motion.

¹²The prefix *na-* has not been analysed here as marking 'ambulative' because its basic locative function is of an inessive marker. In the other examples throughout the paper where this prefix is found, it already works as an aspectual marker.

Another important inconsistency is found in the past of realis, first noticed by Soustelle (1937: 239) at least for the verb in (10). Here a number of verbs (namely, all transitive verbs and a number of intransitive ones), which are said to pertain to the *TANA* conjugation, inflect like a *TATI* verb. The key example is given in (10) with the transitive verb *ne* ‘want’, which should in principle have the expected inflection given in (5) above.

10.	<i>TANA</i>		Expected form
		Past realis	
	1st	to-ne	‘I wanted’ * to-ta-ne
	2nd	co-ne	‘you wanted’ * co-ca-ne
	3rd	pi-ne	‘he wanted’ pi-n-ne

Inconsistencies like the ones noted, which Cárceres himself did not himself recognise or characterise as inflectional oddities, plus others scattered throughout the data of the Colonial grammar, made Soustelle (1937: 249–50) criticise the work by Cárceres very harshly (emphasis and translation mine):¹³

Nothing is more confusing and more difficult to understand than the description of the verb by Cárceres. [...] There is an extraordinary confusion in the Otomi verb as it is described in Cárceres. [...] Besides, in each conjugation, the examples given by Cárceres differ considerably from each other. [...] To state it clearly, *it seems that he artificially grouped all forms that came to his notice in two separate ‘conjugations’, but that the grouping itself did not correspond to any reality*, because in each conjugation there is no rule that can bind the forms together. *Cárceres is the only author who talks about conjugations in Otomi, and his evidence can be considered suspect*, for the reasons given above.

Words like these illustrate, on the other hand, the general feeling towards the Colonial grammar amongst linguists working with Oto-Pamean languages who come to know this historical work. For example, Lastra (2000) also comments negatively on the confusing style of the grammar in presenting the data. I interpret Soustelle’s words as being a reflection of his own frustration at not understanding the data or the grammatical system as a whole. While a little philological patience with the text would have perhaps made things a little clearer, Soustelle was definitely not helped by the fact that he happened not to know modern Otomi languages with more than one conjugation, having carried out linguistic fieldwork only on Otomi languages which lack such inflectional contrasts. In other words, if Soustelle had come across an Otomi language with conjugation classes, he might have understood the data by Cárceres and would perhaps not have discredited his work as promptly as he did.

One such Otomi language lacking different conjugations is San Ildefonso Tultepec Otomi (see Map 1 above), where verbs are inflected in one large default class. The case is illustrated with the present tense of realis and irrealis, in (11) and (12), respectively.¹⁴

¹³Rien n’est plus confus et plus difficilement intelligible que la description du verbe par Cárceres. [...] [U]ne extraordinaire confusion n’en subsiste pas moins dans le verbe otomi tel que le décrit Cárceres. [...] Mais, à l’intérieur de chacune de ces conjugaisons, les exemples que donne Cárceres diffèrent considérablement les uns des autres. [...] A vrai dire, il semble que l’on ait groupé artificiellement les formes qui se présentaient à l’observation en deux «conjugaisons», sans que ce groupement réponde à la réalité, puis qu’à l’intérieur de chaque conjugaison aucune règle ne paraît lier ces formes les unes avec les autres. Cárceres est le seul auteur qui parle de conjugaisons en otomi, et son témoignage peut être considéré comme suspect, pour les raisons exposées plus haut.

¹⁴The exponents of TAM in Otomi languages are commonly proclitics instead of prefixes. I will treat them here as prefixes for the sake of coherence with the way the historical material is presented elsewhere in the paper.

11. 1st person present realis
- dí-tági** /dí-tógi/ 'I fall'
 - dí-gáni** /dí-góni/ 'I snore'
 - dí-hwé^hki** /dí-h^wéki/ 'I pity (him)'
12. 1st person present irrealis
- ga-tági** /ga-tógi/ 'I will fall'
 - ga-gáni** /ga-góni/ 'I will snore'
 - ga-hwé^hki** /ga-h^wéki/ 'I will pity (him)'

Other languages with no conjugations are Mezquital Otomi, the Otomi language with the greater number of speakers, and all the other varieties around the historical centre of Jilotepec (see Map 1). Interestingly, these languages are precisely the historical descendants of the variety of Colonial Otomi which was described by Cárceres in the sixteenth century. Because of their lack of different conjugations, such languages should be regarded now as morphologically innovative, having lost a large amount of the inflectional exponence present in the old language. All later grammars of Otomi were based on innovative languages, and thus on systems which lacked different conjugations, e.g. Neve y Molina (1767), *Luces del Otomí* (1893/1767?) and López Yepes (1826). Because of this, Cárceres is indeed 'the only author who talks about conjugations in Otomi', but Soustelle's claim that Cárceres' evidence is not to be trusted is utterly unjustified in the light of modern evidence from conservative Otomi languages, which have preserved a number of the contrasts described for Colonial Otomi.

3. VERB CONJUGATIONS IN MODERN OTOMI

Some modern Otomi languages have preserved at least the two inflectional classes present in Cárceres' analysis of Colonial Otomi. One such language is Eastern Highlands Otomi (EH-Oto) (see Map 1 above). The data I present in this section come from Voigtlander & Echegoyen (1979). The two conjugations of this Otomi language can be readily observed in the formal contrast characterizing the means of exponence of the 1st person present realis, as illustrated in (13). Here conjugation II bears the prefix *dí-*, which is the historical reflex of the *ti-* prefix from the *TATI* conjugation we saw in §2.1.

13. 1st person present realis
- | | | | |
|---------------|--------------------------------|----------------------------|--------------------|
| I | dí-tági | /dí-tógi/ | 'I fall' |
| II (Option A) | dí-dí-hwé^hki | /dí-dí-h ^w éki/ | 'I care (for him)' |
| (Option B) | da-dí-hwé^hki | /da-dí-h ^w éki/ | 'I care (for him)' |

As indicated in (13), conjugation II has two encoding options available for the present realis. The options are diatopic: A is found at the village of San Antonio el Grande and B at San Gregorio. Leaving aside a number of details of the conjugations,¹⁵ as well as other tenses such as the perfect or the pluperfect, which are based on one of the tenses illustrated, the basic paradigms of conjugations I and II in Eastern Highlands Otomi are given in (14).

¹⁵The paradigms in (14) correspond to subclass (a) of such conjugations; Eastern Highlands Otomi has other subclasses characterised by the presence of a nasal morpheme in a complex distribution.

14.		I	II	
			Option A	Option B
Present realis	1st	di-	dí-dí-	da-dí-
	2nd	gí-	gí-dí-	ga-dí-
	3rd	(i-)	(i-)dí-	
Imperfect realis	1st	d-mí-	d-mí-dí-	ma-dí-
	2nd	g-mí-	g-mí-dí-	ma-dí-
	3rd	Ø-mí-	Ø-mí-dí-	ma-dí-
Past realis	1st			dá-
	2nd			gá-
	3rd			bi-
Present irrealis	1st			ga-
	2nd			gi-
	3rd	da-		di-

In (14), we may observe that the major paradigmatic contrast between the classes is limited to the present and the imperfect realis where verbs of conjugation II have the prefix *dí-*. Of the two inflectional classes, conjugation I is the less marked. In the Option A dialect, conjugation II is based on conjugation I, which serves as a default. The other minor contrast is the marker of the 3rd person present irrealis, where conjugation I has *da-*, while conjugation II has the marker *di-*.

We can trace most formatives in (14) back to Colonial Otomi. In this sense, the means of exponence of the present realis are the reflexes of the historical present continuous realis, as shown in (15).

15.	Present continuous realis > present realis				
	<i>TANA</i>	I	<i>TATI</i>	II (Option A)	
	1st	ti-	> dí-	ti-ti-	> dí-dí-
	2nd	qui-	> gí-	qui-ti-	> gí-dí-
	3rd	(i-)	> (i-)	(i-)ti-	> (i-)dí-

The formatives of Option B of present realis of conjugation II stem from the present (habitual) realis of *TATI*. This is illustrated in (16).

16.	Present (habitual) realis > present realis (Option B)				
	<i>TATI</i>		II (Option B)		
	1st	ta-ti-	> da-dí-		
	2nd	ca-ti-	> ga-dí-		

The morphology involved in the expression of the imperfect is the one found in Colonial Otomi for the imperfect continuous realis, e.g. (17a), but Option B for the imperfect of conjugation II stems from the levelling of the 3rd person found in the historical imperfect (habitual), as shown in (17b).

17. a. Imperfect continuous realis > imperfect
- | | <i>TANA</i> | | I | <i>TATI</i> | | II (Option A) |
|-----|-------------|---|-------|-------------|---|---------------|
| 1st | to-mi- | > | d-mí- | to-mi-ti- | > | d-mí-dí- |
| 2nd | co-mi- | > | g-mí- | co-mi-ti- | > | g-mí-dí- |
| 3rd | Ø-mi- | > | Ø-mí- | Ø-mi-ti- | > | Ø-mí-dí- |
- b. Imperfect habitual realis > imperfect (Option B)
- | | | | | | | |
|-----------------|--|---|----------|---|--------|------------|
| 3rd > 1st & 2nd | | > | Ø-ma-ti- | > | ma-dí- | (Option B) |
|-----------------|--|---|----------|---|--------|------------|

In the past tense, both conjugations receive the same marking. The current morphology may come from the inflection of the past realis of *TATI* seen in (5) above, as illustrated in (18a), but it may well be that the markers for the 1st and 2nd person have stemmed from the perfect cislocative of *TANA* seen in (4) above, as illustrated in (18b).

18. a. Past realis
- | | <i>TATI</i> | | I & II |
|-----|-------------|---|--------|
| 1st | to- | > | dá- |
| 2nd | co- | > | gá- |
| 3rd | pi- | > | bi- |
- b. Perfect cislocative realis
- | | <i>TANA</i> | | I & II |
|-----|-------------|---|--------|
| 1st | ta- | > | dá- |
| 2nd | ca- | > | gá- |

Finally, the prefixes conveying present irrealis in Eastern Highlands Otomi are the only ones which are not traceable to any of the markers present in the inflection of the two main classes of Colonial Otomi, except the prefix *di-* of the 3rd person of conjugation II, which is, as expected, a reflex of the *TATI* conjugation. The case is shown in (19).

19. Present irrealis > present irrealis
- | | <i>TANA</i> | | I | <i>TATI</i> | | II |
|-----|-------------|---|----------------|-------------|---|------------|
| 1st | ca-ta- | > | ga- | ca-ti- | > | ga- |
| 2nd | ca-ca- | > | gi- | ca-ti- | > | gi- |
| 3rd | ti-n- | > | da- ti- | | > | <i>di-</i> |

The origin of the modern markers in (19) is puzzling. Before attempting to provide an answer to this puzzle, in the next section I first introduce more important data of another modern Otomi language which is more conservative than the variety from the Eastern Highlands.

3.1. *Verb conjugations in Tilapa Otomi*

Another language with conjugations is Tilapa Otomi (Til-Oto) (see Map 1 above). In this language, as in the Eastern Highlands, we can establish the existence of two major inflectional classes based again on the marking of the 1st person of the present continuous realis. This is illustrated in (20). The data in this section come from my own fieldwork on this language. These data serve as evidence that Cárceres' analysis of Colonial Otomi was fundamentally right. For the sake of convenience, I will treat the conjugations in (20) preliminarily as 'A' and 'B', instead of using roman numbers.

20. 1st person present continuous realis

A	txá- ^h tági	/tʃá-tʃógi/	‘I’m falling’
	txá-káni	/tʃá-kógi/	‘I’m snoring’
B	txá-tí-hwé ^h ki	/tʃá-tí-h ^w éki/	‘I’m giving it as a present’

In this respect, Tilapa Otomi appears to be no different from the previous language, but when the paradigm of the present irrealis is taken into account, we can see that this Otomi language displays three conjugations, not two, as illustrated by the contrasts in (21). The classes are now given with roman numbers.

21. 1st person present irrealis

I	gu- ^h tági	/gi-tʃógi/	‘I will fall’
II	gu-tu-káni	/gi-tí-kógi/	‘I will snore’
III	gi-tí-hwé ^h ki	/gi-tí-h ^w éki/	‘I will give it as a present’

The full paradigm of this tense is given in (22).

22. Present irrealis

	I	II	III
1st	gu-	gu-tu-	gi-tí-
2nd	gi-	gu-gu-	gi-tí-
3rd	ta-		tí-

We can treat the inflectional formatives of conjugations II and III as the historical reflexes of the markers in the present irrealis of both *TANA* and *TATI*, respectively, as shown in (23).¹⁶

23. Present irrealis of *TANA* & *TATI* > present irrealis of conjugations II and III

	<i>TANA</i>		II		<i>TATI</i>		III			
1st	ca-ta-	/ka-ta-/	>	gu-tu-	/gi-tí-/	ca-ti-	/ka-tí-/	>	gi-tí-	/gi-tí-/
2nd	ca-ca-	/ka-ka-/	>	gu-gu-	/gi-gí-/	ca-ti-	/ka-tí-/	>	gi-tí-	/gi-tí-/
3rd	ti-n-	/ti-n-/	>	tí-	/tí-/	ti-	/tí-/	>	tí-	/tí-/

But, as in the Eastern Highlands in (19) above, it is difficult to reconstruct the origin for the exponence of the present irrealis of conjugation I for Tilapa Otomi. This problem is given in (24) and I provide a solution for it in the next section where I posit that we should increase the number of conjugations in Colonial Otomi.

24. *? > Present irrealis of conjugation I

			Til-Oto	EH-Oto
1st	*?	>	gu-	ga-
2nd	*?	>	gi-	gi-
3rd	*?	>	ta-	da-

¹⁶Colonial lenis /k/ voiced to /g/ in inflectional morphemes in Tilapa Otomi, while lenis /t/ remained lenis. Unstressed /a/ in the prefix /ta-/ centralized to /i/. In an inflectional sequence, the vowel of the prefix further away from the stem harmonised to the vowel of the prefix closer to the stem, e.g. /gi-tí-/ and /gi-tí-/.

4. REVISITING THE CONJUGATIONS OF COLONIAL OTOMI

The inconsistencies raised by the analysis of the conjugations in Cárceres (1580/1907) are solved if one proposes the existence of a third distinct conjugation from the ones already established. This new conjugation can accommodate the particulars of the inflection of many verbs in the Colonial grammar whose inflection does not fit the canonical paradigmatic profiles of *TANA* and *TATI*. With this proposal, one can not only solve the historical problems posited by the two conjugations but also understand the paradigmatic distribution of the markers across the conjugations as given in §2.1. In my proposal, the conjugations are characterised as in (25). The forms instantiating the classes are again in the 1st person of the present realis.

25. I *ta-xohnibi* ‘I teach him’
 II *ta-na-xohnabate* ‘I teach’
 III *ta-ti-māti* ‘I love’

In this article, I am proposing that the verbs registered in both Cárceres (1580/1907) and Urbano (1605/1990) as belonging to the *TANA* conjugation should be split into two conjugations: conjugation I and II. I will treat the old *TATI* conjugation as conjugation III. Through detailed examination of the Colonial grammar, I have been able to identify most of the markers involved in the inflection of conjugation I. To do this, I recorded 1,435 inflected verbal forms and classified them both morphologically and lexically.¹⁷ In my sample, there are 502 inflectional tokens of conjugation I (253 transitive and 249 intransitive), 587 of conjugation II (all intransitive) and 346 of conjugation III (320 intransitive and only 26 transitive uses). These tokens include tokens of the active, the indefinite/impersonal, and the passive voice. The figures should not be taken to indicate frequency of usage but just the number of grammatical forms used in the grammar to illustrate certain inflectional properties of Otomi lexemes. In this respect, Cárceres uses four verbs to illustrate most paradigms: transitive *xohmi* /ʃɔ́h̄ni/ ‘teach’ of conjugation I in 174 tokens; its derived intransitive *xohnabate* /ʃɔ́h̄n̄ap̄ate/ ‘teach’ of conjugation II in 344 tokens, and of conjugation III, transitive *māti* /m̄ati/ ‘love’ in 37 tokens and its derived intransitive *māti* /m̄ate/ ‘love’ in 271 tokens. Besides these four lexemes, which already constitute 57 per cent of all inflected forms in my sample, we find other lexemes: conjugation I has 36 (discernible) transitive verbs and 15 intransitive ones; conjugation II has 42 intransitive verbs, of which only 4 are basic (the rest are derived from conjugation I); and conjugation III has 10 transitive verbs and 6 intransitive ones (only 2 of the latter are basic).

Morphologically speaking, conjugation I is the least marked of the three classes, and serves in most cases as the default on which the inflectional exponents of the other two conjugations are based. We can observe this in a number of descriptive facts:

(i) Conjugation I serves as the default for conjugations II and III in the present realis and irrealis, as shown in (26).

¹⁷The forms in question were recorded as they occurred in the section on verbal morphology. There are, of course, many other verbal forms to be found in sections devoted to Otomi phraseology, which I have not included here. Similarly, I did not include nominalisations.

26.		I		II		III	
		xohnibi	‘teach him’	xohnabate	‘teach’	māti	‘love’
	Default	Present (habitual) realis					
	ta-	1st	ta-xohnibi	ta- na -xohnabate	ta- ti -māti	‘I do’	
	ca-	2nd	ca-xohnibi	ca- na -xohnabate	ca- ti -māti	‘you do’	
	Ø-	3rd	Ø-(na-)xohnibi	Ø- na -xohnabate	Ø- ti -māti	‘he do’	
		Present irrealis					
	ca-	1st	ca-xohnibi	ca- ta -xohnabate	ca- ti -māti	‘I’ll do’	
	ca-	2nd	ca-xohnibi	ca- ca -xohnabate	ca- ti -māti	‘you’ll do’	
	qui-		qui-xohnibi				
	ta-	3rd	ta-xohnibi				
	ti-			ti- n -xohnabate	ti-māti	‘he’ll do’	

The use of the prefix *ta-* in (26) for the 3rd person of the present irrealis of conjugation I is yet another of the inconsistencies found in the Colonial grammar. But its occurrence can help us reconstruct the source for the inflectional formatives for the present irrealis in the modern languages, as shown in (27).

27.	I		I	I	(default)
	Present irrealis		Til-Oto	EH-Oto	SIT-Oto
	1st	ca-	> <i>gu-</i>	<i>ga-</i>	<i>ga-</i>
	2nd	qui-	> <i>gi-</i>	<i>gi-</i>	<i>gi-</i>
	3rd	ta-	> <i>ʔa-</i>	<i>da-</i>	<i>da-</i>

The morphology in (27) shows that in modern Otomi languages, the inflectional formatives involved in the largest conjugation (i.e. conjugation I for those languages with different conjugations) are commonly the historical reflexes found in Colonial Otomi as exponents of conjugation I.

(ii) In the imperfect continuous of irrealis, the paradigms of conjugations II and III are levelled in favour of conjugation I.

28.		I		II		III	
		xohnibi	‘teach him’	xohnabate	‘teach’	māti	‘love’
	Default	Imperfect continuous irrealis					
	co-mi-n-	1st	co-mi-n-xohnibi	co-mi-n-xohnabate	co-mi-n-māti	‘(if) I were doing’	
	co-mi-n-	2nd	co-mi-n-xohnibi	co-mi-n-xohnabate	co-mi-n-māti	‘(if) you were doing’	
	to-mi-n-	3rd	to-mi-n-xohnibi	to-mi-n-xohnabate	to-mi-n-māti	‘(if) he were doing’	

(iii) In some tenses, the paradigms of conjugation II are levelled to I, the default, while III remains distinct and more marked, e.g. (29).

29.		I	II	III	
		xohnibi ‘teach him’	xohnabate ‘teach’	māti ‘love’	
	Default	Present continuous realis			
	ti-	1st ti-xohnibi	ti-xohnabate	ti- ti -māti	‘I am doing’
	qui-	2nd qui-xohnibi	qui-xohnabate	qui- ti -māti	‘you’re doing’
	i-	3rd i-xohnibi	i-xohnabate	i- ti -māti	‘he’s doing’
		Imperfect continuous realis			
	to-mi-	1st to-mi-xohnibi	to-mi-xohnabate	to-mi- ti -māti	‘I was doing’
	co-mi-	2nd co-mi-xohnibi	co-mi-xohnabate	co-mi- ti -māti	‘you were doing’
	Ø-mi-	3rd mi-xohnibi	Ø-mi-xohnabate	Ø-mi- ti -māti	‘he was doing’
		Perfect cislocative realis			
	ta-	1st ta-xohnibi	ta-xohnabate	ta- ti -māti	‘I came to do’
	ca-	2nd ca-xohnibi	ca-xohnabate	ca- ti -māti	‘you came to do’
	a-	3rd a-xohnibi	a-xohnabate	a- ti -māti	‘he came to do’

(iv) In other tenses, the paradigms of conjugation III are levelled up to conjugation I, while the inflection of verbs of conjugation II remains distinct and more marked.

30.		I	II	III	
		xohnibi ‘teach him’	xohnabate ‘teach’	māti ‘love’	
	Default	Past realis			
	to-	1st to-xohnibi	to- ta -xohnabate	to-māti	‘I done’
	co-	2nd co-xohnibi	co- ca -xohnabate	co-māti	‘you done’
	pi-	3rd pi-xohnibi	pi- n -xohnabate	pi-māti	‘he done’
		Past irrealis			
	co-	1st co-xohnibi	co- ta -xohnabate	co-māti	‘I’d have done’
	co-c(a)-e	2nd co-qu-e-xohnibi	co- ca -xohnabate	co-qu-e-māti	‘you’d have done’
	to-co-	3rd to-co-xohnibi	to-co-xohnabate	to-co-māti	‘he’d have done’
		Imperfect A irrealis			
	co-ma-	1st co-ma-xohnibi	co-ma- na -xohnabate	co-ma-māti	‘(if) I were to doe’
	co-ma-	2nd co-ma-xohnibi	co-ma- na -xohnabate	co-ma-māti	‘(if) you were to doe’
	to-ma-	3rd to-ma-xohnibi	to-ma- na -xohnabate	to-ma-māti	‘(if) he were to doe’

(v) In the imperfect realis, and in the 3rd person of the present realis, the paradigm of conjugation I has already been levelled to conjugation II.

31.		I	II	III	
		xohnibi ‘teach him’	xohnabate ‘teach’	māti ‘love’	
		Present (habitual) realis			
	3rd	Ø- na -xohnibi	Ø- na -xohnabate	Ø-ti-māti	‘he does’
		Imperfect (habitual) realis			
	1st	to-ma- na -xohnibi	to-ma- na -xohnabate	to-ma-ti-māti	‘I used to do’
	2nd	co-ma- na -xohnibi	co-ma- na -xohnabate	Ø-ma-ti-māti	‘you used to do’
	3rd	Ø-ma- na -xohnibi	Ø-ma- na -xohnabate	ma-ti-māti	‘he used to do’

(vi) Finally, in at least the three tenses in (32), we lack information about what markers could have been involved in the inflection of conjugation I. This is because this class does not constitute a separate category in the Colonial grammar.

32.	I	II	III	
	xohnibi ‘teach him’	xohnabate ‘teach’	māti ‘love’	
	Present continuous irrealis			
1st	?	ca-ti-xohnabate	ca-ti- ti -māti	‘I’ll be doing’
2nd	?	ca-ti-xohnabate	ca-ti- ti -māti	‘you’ll be doing’
3rd	?	ti- n -xohnabate	ti- ti -māti	‘he’ll be doing’
	Present cislocative irrealis			
1st	?	ca-ta-xohnabate	ca-ta- ti -māti	‘I’ll come to do’
2nd	ca-ta-xohnibi	ca-ta-xohnabate	ca-ta- ti -māti	‘you’ll come to do’
3rd	?	ta-xohnabate	ta- ti -māti	‘he’ll come to do’
	Imperfect B irrealis			
1st	?	co-ma- ta -xohnabate	co-ma- ti -māti	‘(if) I were to do’
2nd	?	co-ma- ca -xohnabate	co-ma- ti -māti	‘(if) you were to do’
3rd	?	to-ma-xohnabate	to-ma- ti -māti	‘(if) he were to do’

4.1. More about conjugation I

I can point to a number of specific parts of the paradigm which consistently reveal conjugation-class membership. Whereas inflected forms with the prefixes *to-ta-* or *co-ca-* for the 1st and 2nd person, respectively, of the past realis point invariably to a verb of conjugation II, forms with *to-* or *co-* point to a verb that belongs to either conjugation I or III. If that verb has a *ti-* formative in the present realis,¹⁸ it is most likely a verb of conjugation III. If for the present realis, the verb has *ta-* or *ca-* for the 1st and 2nd person, respectively, then it is invariably a verb of conjugation I, just as much as when a given inflected form has the prefix *ta-* for the 3rd person of present irrealis. But conjugation I verbs may also take the prefix *na-* for the present realis in all persons, making the inflected forms of such a verb look as though they belonged to a conjugation II verb, whereas in reality such forms would realise the motion components illustrated in (9) above.

The valency of a verb is also a useful indicator of class membership if the verb is transitive, as it can only belong to conjugation I or III. Derived intransitive verbs bearing the element *-te* can either belong to conjugation II or III, depending on the class membership of their transitive source.

The analysis of the Colonial data I have advanced has the advantage of solving the inconsistencies pointed out by Soustelle (1937). Additionally, a number of intransitive verbs (possibly a rather substantial number after further research is carried out) treated in Cárceres (1580/1907) and Urbano (1605/1990) as belonging to the *TANA* conjugation should be reallocated to conjugation I. The 15 verbs I have been able to identify so far in my sample are given in Table 2.

Other intransitive verbs that belong to conjugation I are the ‘patientive’ verbs, which are one of the active/patient split intransitivity types proposed in Mithun (1991), and consequently encode their only argument as an object. Two examples are given in (33) with the verbs *beti* /ʔbétʃi/ ‘get lost’ and *ṭso* /tsɔ/ ‘be able’.

¹⁸In the grammar, one also finds occasional forms with two other homophonous prefixes, one serving as a directional, the other as an inchoative marker, illustrated in example (34b).

Table 2. Some intransitive verbs of conjugation I

<i>beni</i>	/b ^h eni/	‘remember’
<i>cāy</i>	/k ^h ai/	‘go down’
<i>cha</i>	/k ^h a/	‘happen, exist’
<i>ēhē</i>	/ʔēhē/	‘come’
<i>hūānni</i>	/h ^w āni/	‘pacify’
<i>hūy</i>	/hūi/	‘sit’
<i>nangui</i>	/nāngi/	‘get up’
<i>pa</i>	/pā/	‘go’
<i>phuēti</i>	/p ^h ūti/	‘start’
<i>quētti</i>	/k ^h iti/	‘go in’
<i>tētge</i>	/t ^h éte/	‘go up’
<i>thege</i>	/t ^h ége/	‘finish’
<i>togui</i>	/t ^h ogi/	‘fall’
<i>tū</i>	/tū/	‘die’
<i>vāy</i>	/ʔwāi/	‘rain’

33. a. *tabequi*
 /ʔa-ʔb^hé-ki/
 (3.) PRES.IRR-get.lost.A-IOBJ
 ‘I will get lost.’ (p. 80)
- b. *nāʔgoqui*
 /nā-ʔso-ki/
 (3.) PRES.R-be.able-IOBJ
 ‘I can’t (do it).’ (p. 80)

Passive forms, which are also patientive like the examples in (32), are inflected following conjugation I regardless of whether the transitive verb already belongs to conjugation I, as in (34a), or originally belongs to conjugation III, as in (34b) (there are no transitive verbs in conjugation II).

34. a. *tāʔhohnigui*
 /ʔa-ʔs^hō^hni-ki/
 (3.) PRES.IRR-IMPER/PASS.teach-IOBJ
 ‘I’ll be taught.’ (p. 78)
- b. *notīmātiki*
 /n^h-o-ʔi-māt^h-i-k^hi/
 (3.) PRES.R-IMPER-INCHO-love-2OBJ
 ‘You are loved.’ (p. 83)

The example in (34a) comes from the transitive verb *xohni* /ʔs^hni/ ‘teach’ of conjugation I, (34b) from *māti* /māti/ ‘love’ of conjugation III. For their passive form, the verbs of conjugation I undergo consonant mutation, while the verbs of conjugation III take the inchoative prefix *ti-* (which is homophonous with the prefix *ti-* of conjugation III) plus the impersonal prefix *o-*, which also functions as an exponent of passive voice for this conjugation.

All remaining intransitive verbs from the *TANA* conjugation form conjugation II. In contrast to conjugation I, conjugation II stands as a marked inflectional class containing inergative intransitive verbs which most commonly depict human actions. In this conjugation – apart from basic verbs, of which I have been able to identify only four in my sample of the grammar (*buæy* /ʔb^whi/ ‘live’, *ʔohni* /ʔso^hni/ ‘arrive there’, *ʔiya* /ʔsiya/ ‘be greedy’, *ʔa* /ʔsá/ ‘have a meal’) – we have many derived antipassive verbs from conjugation I formed by adding the suffix *-(ba)te* /(-pa)-te/. A few of such verbs are given in Table 3.

The derived verbs in Table 3 depict the action of the transitive verb as an agentive activity. This is why I treat them as antipassive verbs. Additionally, the verbs can be used via conversion as agent nouns depicting a role or a craft. This derivation is no longer productive in the modern innovative languages. The pattern is still observable in conservative ones, but

Table 3. A few derived activity verbs of conjugation II

I (tr)		>	II (in)	
<i>āxi</i>	‘shear’	>	<i>āx-te</i>	‘be a shearer’
<i>chūtti</i>	‘pay’	>	<i>chūtta-te</i>	‘be a payer’
<i>ḡihi</i>	‘bring’	>	<i>ḡi-te</i>	‘be bringing’
<i>ḡū</i>	‘fear’	>	<i>ḡū-te</i>	‘be afraid’
<i>ēntti</i>	‘push’	>	<i>ēntta-te</i>	‘be a pusher’
<i>ēntḡtti</i>	‘lift’	>	<i>ēntḡa-te</i>	‘be a lifter’
<i>huānni</i>	‘choose’	>	<i>huānna-bate</i>	‘be a chooser’
<i>nu</i>	‘see’	>	<i>nu-te</i>	‘be able to see’
<i>pātḡi</i>	‘smell’	>	<i>pātḡa-te</i>	‘be a smeller’
<i>pehni</i>	‘send’	>	<i>pehna-te</i>	‘be an envoy’
<i>photḡi</i>	‘help’	>	<i>phoxa-te</i>	‘be a helper’
<i>puāmi</i>	‘sacrifice’	>	<i>puānā-bate</i>	‘be a sacrificer’
<i>theni</i>	‘follow’	>	<i>tenna-te</i>	‘be a follower’
<i>uti</i>	‘show’	>	<i>uta-te</i>	‘show away’
<i>xóhni</i>	‘teach’	>	<i>xóhna-bate</i>	‘be a teacher’
<i>yātti</i>	‘promise’	>	<i>yātta-te</i>	‘promise away’

not to the extent it appears to have been productive in Cárceres’ time, where we can only conjecture that it must have been rather common, making the author of the Colonial grammar believe that the verbs involved belonged to one unified inflectional class, i.e. the *TANA* conjugation.

Traces of the three conjugations of Colonial Otomi have been preserved with different outcomes in the conservative languages of the south of Mexico State and Tlaxcala. In these languages, conjugation I completed its levelling to conjugation II in both the present and imperfect of the realis. The varieties spoken in the Eastern Highlands reduced the contrast to only two classes: conjugation III was preserved in the present and imperfect, and in the 3rd person of the present irrealis. In this language, conjugation II was assimilated to conjugation I, but still survives in the form of an inflectional subclass which Voigtlander & Echegoyen (1979) treat as conjugation Ib. The paradigm of this conjugation Ib has a nasal prefix conditioned by tense, which is probably the reflex of the *na-* prefix we see in the Colonial data.

On the other hand, the languages of the north and the northern part of Mexico State, which are the linguistic descendants of the Colonial dialect described in Cárceres, were also the systems that underwent more drastic reductions. In these varieties, the three classes were reduced to one in favour of conjugation I. Conjugation III disappeared altogether, but, as in Eastern Highlands Otomi, remnants of conjugation II may be seen in a rather small inflectional class comprising 20–30 intransitive verbs depicting human activities (for more details, see the activity verbs in Palancar 2009).

5. FINAL REMARKS

I have proposed that Colonial Otomi had three major inflectional classes of verb, or conjugations, instead of the two we find in the analysis by Pedro de Cárceres as author of a sixteenth-century grammar of the language. With this proposal, the inconsistencies of the inflection of verbs in his grammar are solved, and we can understand the origin and development of the markers in modern languages. The sixteenth- and seventeenth-century missionary grammars of Mesoamerican languages are difficult to interpret when the modern linguist lacks the proper linguistic context of how to interpret their data; for example, Soustelle (1937) had not come across a modern Otomi language with conjugations, and was too quick (in my opinion) to discredit the invaluable work by Cárceres, who had conducted a

pioneering description of the common language 300 years before him. But missionary grammars are also treasures for historical reconstruction, which are worth digging up when philological and linguistic efforts are properly combined.

Surrey Morphology Group
Faculty of Arts and Human Sciences
AC Building, University of Surrey
Guildford, Surrey GU2 7XH
Email: e.palancar@surrey.ac.uk

REFERENCES

- BICKEL, BALTHASAR & NICHOLS, JOHANNA, 2007. 'Inflectional morphology', in Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. III: *Grammatical Categories and the Lexicon*, 2nd edn, Cambridge: Cambridge University Press, 169–240.
- CÁRCERES, FRAY PEDRO DE, 1580/1907. *Arte de la Lengua Othomi*, ed. Nicolás León, *Boletín del Instituto Bibliográfico Mexicano* 6, 38–155.
- CLIN: CATÁLOGO DE LAS LENGUAS INDÍGENAS NACIONALES, 2008. 'Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas', *Diario de la Nación* (14 Jan.), Mexico City. Accessible at: http://www.inali.gob.mx/pdf/CLIN_completo.pdf
- GARONE GRAVIER, MARINA, 2010. 'Las lenguas otomangués y tipografía colonial poblana: comentarios en torno a ediciones en zapoteco, mixteco, y mixe', invited paper at IV Coloquio Lenguas Otomangués y sus Vecinas: Homenaje a Thomas C. Smith Stark, Oaxaca, Mexico, 16–18 April.
- LASTRA, YOLANDA, 2000. 'El arte de la lengua otomí de Fray Pedro de Cárceres', in Otto Zwartjes (ed.), *Las gramáticas misioneras de tradición hispánica (siglos XVI–XVII)*, Amsterdam: Rodopi, 97–106.
- LASTRA, YOLANDA, 2001. *Unidad y diversidad de la lengua: relatos otomíes*, Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- LASTRA, YOLANDA, 2006. *Los otomíes: su lengua y su historia*, Mexico City: Instituto de Investigaciones Antropológicas, Universidad Autónoma Nacional de México.
- LÓPEZ YEPES, JOAQUÍN, 1826. *Catecismo y declaración de la doctrina cristiana en lengua otomí, con un vocabulario del mismo idioma*, Mexico City: Alejandro Valdés.
- Luces del Otomí: O Gramática del idioma que hablan los indios otomíes en la República Mexicana*. 1893/1767? (Anonymous), ed. Eustaquio Buelna, Mexico City: Imprenta del Gobierno Federal.
- MITHUN, MARIANNE, 1991. 'Active/agentive case marking and its motivations', *Language* 67.3, 510–46.
- NEVE Y MOLINA, LUIS, 1767. *Reglas de ortografía, diccionario y arte del idioma othomí: breve instrucción para los principiantes*, Mexico City: Imprenta de la Bibliotheca Mexicana.
- PALANCAR, ENRIQUE L., 2009. *Gramática y textos del Hñöhñö: Otomí de San Ildefonso Tultepec*, vol. I, Mexico City: Plaza y Valdés.
- SOUSTELLE, JACQUES, 1937. *La famille otomí-pame du Mexique central*, Paris: Institut d'Ethnologie.
- ÚRBANO, ALONSO DE, 1605/1990. *Arte breve de la lengua otomí y vocabulario trilingüe*, ed. René Acuña, Mexico City: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.
- VOIGTLANDER, KATHERINE & ECHEGOYEN, ARTEMISA, 1979. *Luces contemporáneas del Otomí: gramática del Otomí de la Sierra*, Mexico City: Instituto Lingüístico de Verano.

