CLEFTS IN OTOMI: EXTENDED USES OF THE COPULAR CONSTRUCTION

Enrique L Palancar

CLEFTS IN OTOMI:
EXTENDED USES OF THE COPULAR CONSTRUCTION

ENRIQUE L. PALANCAR
Structure et Dynamique des Langues, CNRS
International Journal of American Linguistics 84(1): 93–145

Abstract
In this article, I study clefts in Otomi, a small family of languages of Mexico from the Oto-Manguean stock. Clefts represent an area of Otomi syntax that has not been previously reported on. I compare the clefts we find in two very different languages: Tilapa Otomi and Northern Otomi. Clefts in Otomi are based on the copular construction and have three typologically expected elements (i.e., a copula, a focus phrase and a relative-like clause encoding the common ground), but they also have two elements that are unique to Otomi: a pronominal enclitic that cross-references the focus phrase and a focus pronoun. The former is a property of the copular construction, but the latter only occurs in clefts. Clefts in Otomi also display a great amount of surface variation in natural discourse. In the article, I also explore the diachrony of the system by studying the historical sources from the Colonial period.

1. Introduction[1]

In this article, I study cleft constructions in Otomi, a small family of Amerindian languages of Mexico from the Oto-Manguean stock. The goal of the article is to attain a first and comprehensive description of cleft structures in Otomi, which constitute an area of Otomi syntax that has not yet been previously reported on. Clefts in Otomi are based on the copular construction, which as I will show in §3.2 is practically only used with a specificational or identificational function. The copular construction has an interesting syntax of its own and most of its constituting elements can be omitted under the right circumstances. This optionality is inherited in clefts, making their analysis challenging. Clefts in Otomi also present a pronoun (*ko* or *go,* depending on the language) which only occurs in clefts and which I take to be a focus pronoun with an unusual typological behavior (see §5 for details). As a way to introduce clefts as an object of study, in this section I give a brief overview of the relevant literature on the subject before I concentrate on Otomi.

1.1. A brief introduction to clefts. The modern study of clefts starts with Jespersen (1949), itself a counter-proposal to the position in Jespersen (1937). The object of Jespersen's study is the English construction known as the 'it-cleft', illustrated in (1a), which Lambrecht (2001) prefers to treat as 'IT-cleft' (with capitals) in order to encapsulate a type of cleft known to exist in other languages, such as for example in French, as shown in (1b), (example 1a is from Lambrecht, 1994: 230).

(1) a. it's my car that broke down (IT-cleft)
b. c'est ma voiture qui est en panne

According to Lambrecht (2001:467) "A cleft construction is a complex sentence structure consisting of a matrix clause headed by a copula and a relative or relative-like clause whose relativized argument is co-indexed with the predicative argument of the copula." In this way, a canonical cleft is constituted of at least the following three elements: (i) a copula or copular element; (ii) a relative or relative-like clause, which I will call the 'clefted clause' (henceforth CC) and which I indicate with curly brackets in (2); and (iii) a predicative argument of the copula that is co-indexed with the relativized argument of the CC and which appears in square brackets.

(2) it's$_{COP}$ [my car] {that broke down}$_{CC}$

But Lambrecht's definition crucially continues as: "Taken together, the matrix and the relative express a logically simple proposition, which can also be expressed in the form of a single clause without a change in truth conditions." In other words, the cleft in (2) is semantically equivalent to the simple clause in (3), by which it can be paraphrased.

(3) my car broke down

The key difference between clefts and simple clauses is that clefts are specialized focus constructions. With this in mind, Jespersen's (1949:147) definition of an 'it-cleft' in English is particularly revealing: "A cleaving of a sentence by means of *it is* (often followed by a relative pronoun or connective) serves to single out one particular element of the sentence and very often, by directing attention to it and bringing it, as it were, into focus, to mark a contrast."[2]

As contrastive focus constructions, the CC is the part of the cleft that encodes the background or common ground, which establishes a given state of affairs as known and relevant to the speech act participants at the time of the speech act. In example (2), the common ground is that something has broken down, and this information has some relevance in the actual context where an utterance like (2) could be said. However, the common ground is presented with a VARIABLE, a piece of information that is open for interpretation in context: in (2) we know that something has broken down but what that might be is under discussion, so the variable is the entity whose identification is under contention. The entity that serves to reveal the identity of the variable is the VALUE: the phrase *my car* in (2) is identified as being the same thing that has broken down, and this is done by means of the copular structure. In doing so, the new information provided by *my car* is introduced with assertive force and conveys focus. In this context, the syntactic phrase *my car* in (2) can be treated as the focus phrase (henceforth FP) of the cleft. In clefts, the value is further evaluated against a range of potential alternatives which get cancelled; this is why they commonly encode nuances of contrastive focus.

Apart from IT-clefts, the literature on clefts has also recognized the existence of other types of clefts. Particularly interesting for our purposes are the so-called 'pseudo-clefts', as shown in (4), which are commonly divided in two subtypes depending on the position of the FP in the sentence. Commonly, the FP receives a special prosodic treatment, by way of stress. This is represented by the small caps.

(4) a. {what broke down}$_{CC}$ is [my CAR]$_{FP}$ (pseudo-cleft)
b. [my CAR]$_{FP}$ is {what broke down}$_{CC}$ (reversed pseudo-cleft)

According to Quirk *et al*. (1985:1387), pseudo-clefts are "essentially an SVC [Subject-Verb-Complement] sentence with a nominal relative clause as subject or complement". In other words, pseudo-clefts are built on the syntax of a copular construction that has a specificational or identificational function of the type instantiated in (5), where following Mikkelsen (2005:1) (5a) is specificational, while (5b) is identificational or equative.[3]

(5) a. The lead actress in that movie is Ingrid Bergman
b. She is Ingrid Bergman

However, the pseudo-clefts in (4) are not mere instances of the copular construction, but also instances of clefts because they can be paraphrased in one simple clause (i.e., their propositional content is similar).[4] The pseudo-cleft in (4a) is a more natural construct in English than (4b), because the FP occurs in the same position where focus is placed by default in an English sentence, that is, towards the end of the clause. For example in (6), the syntactic subject phrase *my car* is topical, but the predicate is in focus (i.e., broad focus if the speaker just asserts that something happened to the car, or contrastive, if it broke down as opposed to being stolen.)

(6) my car broke DOWN

In this connection, the syntax of the pseudo-cleft in (4a) is uncontroversial: the CC is taken to be subject and topical, just like *my car* in (6), whereas the FP is analyzed as the complement of the copula and is in focus, just like *broke down* in (6). However, the same is not true about the so-called reverse pseudo-cleft in (4b), which is taken by most as a marked structure, in the same fashion as (7) would be, where the subject is in focus. In other words, in contrast to (6), the predicate in (7) now encodes the common ground where it is presupposed that something broke down. With respect to such common ground, the speaker asserts that the entity that broke down is his/her car and not something else.

(7) MY CAR broke down

But whereas it is unquestionable that the status of the phrase *my car* in (7) is subject, when it comes to its counterpart phrase in the pseudo-clefts in (4b), authors are less in agreement, mainly because for most the pragmatic notion of focus and the syntactic notion of subject are not compatible notions. This is so to the extent that instances of clefts in (4b) have been analyzed as involving a syntactic movement of predicate inversion, of the same type postulated for copular sentences (den Dikken 2006, Moro 1997, etc.). The position that the FP is not subject in a pseudo-cleft can be seen for example in Lambrecht (2001:467), apropos his use of the term 'predicative argument of the copula' to refer to the FP in his definition of a cleft (see above), when he explicitly states (emphasis mine): "The term 'predicative argument' (instead of the more straightforward 'object' or 'non-subject' argument)[5] is a hedge used in order to account for sentences like *CHAMPAGNE is what I like best*, in which the co-indexed argument occurs in preverbal position." Note that the construction at issue is the reversed pseudo-cleft in (4b). In this paper, I have no theoretical qualms about a subject being in focus in

a cleft, and as for Otomi, I show in §3.3 onwards that there is ample evidence that the FP in Otomi clefts is treated as the subject of the copula.

In summary, clefts share structure with copular constructions, but also have a peculiar syntax for two reasons: (i) they express a simple proposition via biclausal syntax (Lambrecht 2001: 466), and (ii) this syntax is linked to the realization of focus (which is commonly contrastive, but not exclusively).[6] Due to their idiosyncrasies, clefts have not ceased to attract the attention of linguists from different approaches; consequently the literature on clefts in better known-languages is vast. But despite considerable advances in our understanding of how they work, the syntax and pragmatics of clefts continue to defy us. In this respect, in Drubig and Schaffer (2001:1091) –a state of the art on focus constructions– the authors acknowledge that "the grammatical structure of this type of construction is an unsolved problem." The challenge is greater for the lesser-known languages.

1.2. Goals and structure of this article. This paper is a contribution to fill in gaps in our knowledge of the syntax and morphosyntax of clefts in Mesoamerican languages. Mayan is by far the Mesoamerican language family with a renowned literature on the syntax of focus, to mention a few, Aissen (1992), which is seminal and referential, Trechsel (1993), Tonhauser (2003; 2007), and more recently Bohnemeyer (2014), Verhoeven and Skopeteas (2015), Gutiérrez Bravo (2015) and Aissen (to appear). In Oto-Manguean, there have been studies on the syntax of focus mainly in Zapotec languages (Ward 1987; Lee 1997; Broadwell 1999; Black 2000; etc.) and Mixtec (Hollenbach 1995), but as for other branches of Oto-Manguean, such as Oto-Pamean, which is very a distinctive, geographically isolated branch in Central Mexico, I am unaware of any study dedicated to the topic.

To compensate for lacunae in our knowledge of the syntax of these languages, in this article I study cleft structures in Otomi, which is a small language family within the Oto-Pamean branch of Oto-Manguean consisting of at least six languages (see Figure 1 below). My focus is on Tilapa Otomi (henceforth T-Oto), which is the Otomi language of the village of Santiago Tilapa, State of Mexico. This language is currently spoken by about five elderly people, and it is thus one of the most critically endangered Oto-Manguean languages. To give a comprehensive view of clefts in Otomi, I compare the clefts in T-Oto with the structures we observe in Northern Otomi. The choice of these two languages is convenient because they represent the two ends of a continuum of inflectional complexity in the Otomi family and can serve as optimal view-points for the structural diversity in other areas of the grammar, such as the syntax. In this continuum, T-Oto is most conservative, having an inflectional system that has changed little from the situation we observe in the historical sources from the 16th century, which are based on an Otomi that we can take to be the common ancestor of the modern languages (see Palancar 2011, 2012). In contrast, Northern Otomi is an innovative continuum of dialects, some of which have a large number of speakers. I illustrate cleft structure here in two representative dialects: San Ildefonso Tultepec Otomi and Mezquital Otomi. To complete the picture, I also present historical data from Colonial sources and evaluate them in comparison with the synchronic data to understand the modern situation.

INSERT FIGURE 1

A typical instantiation of a cleft in T-Oto is given in example (8), where we have the three typologically expected elements present: a copula; a FP (which in example 9 is a pronominal enclitic hosted on the copula); and a CC.[7]

(8) kẹh=a[=k'u$_i$]$_{FP}$ {ko$_i$ tú=ñü='mbe}$_{CC}$=k'u$_i$
COP.AS=CL=3PL$_{PRO}$ FOC 1.CPL.R=see=PL.EXCL=3PL$_{PRO}$
'It's them we saw.' (Txt)

There are two other elements in (8) that are specific to Otomi: One is the element *ko*, which in §5 I analyze as a focus pronoun that stands for the FP in the CC. *Ko* is an interesting type of pronoun typologically in T-Oto, because against the impression that (8) may give, it is neither a canonical relative pronoun nor does it occur in initial position in the CC. The other element is the pronominal enclitic *=k'u*, which is part of a set of pronominals at the left edge of the CC that cross-reference the FP as subject. In §3.3, I show that such pronominals are a feature of the copular construction.

Clefts in Otomi are an interesting object of study because all their constituting elements (i.e., the copula, the FP, the CC, the focus pronoun, and the pronominal enclitic) can be omitted under the right circumstances. Optionality produces a great amount of surface variation in natural discourse; making the analysis of clefts particularly challenging. For instance, example (9), which is a prayer, is a minimal instantiation of a cleft in T-Oto consisting of just a CC with the focus pronoun *ko*.

(9) *Dios_mio* este... tŕá='mb^{w}u 'a di 'ye
my.God HES 1.INCPL.R=exist P 2POSS hand
{ko xu grá=^{h}pödi}$_{CC}$
FOC indeed 2.INCPL.R=know
'My god...I am in your hands, you're indeed the one who knows.' (Txt)

The data on which I base my analysis come from a large corpus of oral texts (marked as 'Txt') compiled in both T-Oto and San Ildefonso Tultepec Otomi as outcomes of various documentation projects in which I have been involved. I use elicited examples only occasionally. For Mezquital Otomi, I have used the written materials in the ethnographical book by Salinas Pedraza (1984), which was originally written in Otomi and contains c. 130,000 words.

The paper has a straightforward structure. In the next section, I introduce some basics of T-Oto grammar which can serve to illustrate Otomi grammar in general, and which are useful to understand the syntax of both the copular construction and of clefts. In §3 I introduce and analyze the copular construction in T-Oto, which is relatively the same in other Otomi languages. Sections 4-6 are about clefts in T-Oto. Section 4 introduces clefts as based on the copular construction; Section 5 is about the focus pronoun and Section 6 provides an analysis of the different types of surface variants we find in clefts. Section 7 is a brief comparison with Northern Otomi. In section 8, I introduce historical data from Old Otomi to put forward a diachronic proposal about the way clefts evolved in this family. Section 9 concludes.

2. Basics about Tilapa Otomi grammar

In this section, I introduce basics of the grammar of T-Oto that are relevant to understand what will follow about copular structure and cleft structure in Otomi. I focus on T-Oto but in doing so, I hope to illustrate some of the main principles behind the grammar of an Otomi language, although I also give a few relevant data from San Ildefonso Tultepec Otomi (henceforth SIT-Oto), which is a variety of Northern Otomi. The section consists of three parts. First, I give general notions of Otomi verbal inflection, because the language is head-marking and the verb encodes a great deal of grammatical information about the clause. In the two subsequent sections, I give basic notions of word order and relative clause structure. Both are important to understand the structure of clefts.

Verbs in Otomi inflect for tense-aspect-mood (henceforth TAM) values by means of markers that always precede the verbal stem. I call such markers 'inflectional formatives'. Inflectional formatives also index the person of the subject in a cumulative way (although 3rd person subject is indexed only indirectly). Table 1 illustrates a representative set of subparadigms for the (cognate) verb *në* 'dance' in both T-Oto and in SIT-Oto.[8] Brackets represent optional material.[9]

INSERT TABLE 1

Monosyllabic inflectional formatives are phonologically dependent. I treat them as clitics rather than affixes, because they can also be hosted on other elements apart from the verb. This may be seen in the contrast between (10) and (11): in (10), the formative is a proclitic hosted on the verbal stem, while in (11) the same formative is an enclitic hosted on the negative marker. As an enclitic, the onset of the formative undergoes a voicing adjustment before a nasal element (for the use of the approximation sign see further below before Table 3). In contrast, formatives consisting of more than one syllable are independent words, although they may still cluster with preceding elements.[10]

(10) **tŕá**='mb^{w}u̲h=ka 'a Txapultepe
1.INCPL.R=exist.AS=1[SG]$_{PRO}$ P CH.
'I'm at Chapultepec.' (Txt)

(11) hi'na, hin≈**dŕá** ndẹ='mbe i serbesa
no NEG≈1.INCPL.R want=PL.EXCL PL beer
'No, we don't want beer.' (Txt)

Person of object is encoded by suffixes, as illustrated in (12). A bare stem is used to index an object of 3rd person, as shown in (12c) (i.e., there is zero anaphora for an object of 3rd person).

(12) a. b^{w}u̲=^{h}tsix**-ki**=a
[3]CPL.CISLOC=bring.AS-1OBJ=CL
'He brought me.' (Txt)

b. ra=hon**-k'i**
[3]INCPL.R=look.for.AS-2OBJ
'They are looking for you.' (Txt)

c. 'mba ti=kha
then [3]CPL.IRR=do[3OBJ]
'Then he's going to do it.' (Txt)

Number of subject (as well as of possessor and of object) is encoded by means of enclitics, as shown in (13). Some languages have a number distinction involving singular, dual and plural, although in most the dual is marginal and has collocation restrictions (see Palancar 2013 for details). There is also a clusivity distinction involving the first person.

(13) a. tŕámátí zo**='mbe**
1.IMPF talk=PL.EXCL
'We were talking.' (Txt)

b. giti tsoh**=wi** a kitheni
2.CPL.R light.up.AS=PL DEF.SG censer
'You (PL) are going to light up the censer.' (Txt)

For person, Otomi languages have bound pronominal forms that I treat as enclitics. The paradigm of personal pronouns in T-Oto is given in Table 2. The number of the first and the second person is encoded by means of the same number enclitics we find in verbs. In the absence of such markers, singular is the default interpretation, like in (10) above or (14) below. There are dedicated forms for the 3rd person singular and plural. Demonstratives are often used anaphorically as pronouns.

INSERT TABLE 2

Pronominal enclitics cross-reference the person of the subject (or less commonly so, the object).[11] They may occur associated with the verb, like in example (14), where they often encode nuances of contrastive focus.[12]

(14) pues gu=xǫx**=ka** porke rá='wahki=thǫ=gwa
so 1CPL.IRR=lift.AS[3OBJ]=1[SG]$_{PRO}$ because INCPL.ST=be.in.the.open[3S_O]=DEL=here
'Well, I'm going to put it up myself because it (the courtyard) is all in the open here.' (Txt)

Pronominal enclitics may also occur in other positions in the clause. For example, in clefts we find them to the right edge of the matrix clause, as shown in (15) (also in example (8) above). But they may also occur fronted to the left edge like in (16), where the enclitic is hosted on the conjunction, (in example (16) the pronominal is in focus, but see example (22) below for an example where it is topic).

(15) kẹh=a {bi=ñö ra ñühü}$_{CC}$**=’ä**
COP.AS=CL[3] [3]CPL.R=speak NMLZ otomi=3SG$_{PRO}$
‘He was the one who spoke Otomi.’ (Txt)

(16) pe=**ga=’mbe** hints’e gata ñü=’mbe
but=1$_{PRO}$=PL.EXCL no.longer 1.AND.IRR see[3OBJ]=PL.EXCL
‘But WE are no longer going to see it.’ (Txt)

Example (16) further illustrates the occurrence of the negative polarity item *hints’e* ‘no longer’. This is one of a set of elements occurring preverbally in fixed positions. The elements in question have a range of different meanings; apart from polarity markers, they may also include pronominals and adverbials. For analytical convenience, I conceive of the various preverbal positions as constituting a syntactic area in the clause I call ‘the preverbal zone’ (henceforth ‘pre-V’). Some of these items may occur in isolation as words; especially when they are disyllabic, but others arrange themselves in clitic clusters involving the inflectional formatives we saw in Table 1. Such clusters function in turn as phonological words. I use the approximation sign ‘≈’ to indicate morphological junctures in such clusters. Conservative languages like T-Oto preserve a greater number of such items in the pre-V than do innovative ones like SIT-Oto. The two sets of elements occurring in these two languages are given in Table 3 (the capital N indicates a homorganic nasal that assimilates to the onset of the inflectional formative with which they form a cluster).

INSERT TABLE 3

Some examples of pre-V elements from T-Oto in Table 3 are given in (17).

(17) a. hi’na, **hin**≈gata ’ñöndy=a
no NEG≈1.ADLAT.IRR SS/hide.AS=CL
‘No, I’m not going to move.’ (Txt)

b. **han**≈gi ’yotu=wi k’ani!
again≈2.CPL.IRR SS/make.AS=PL wild.vegetable
‘Prepare another dish of greens!’ (Txt)

c. ’neh=ka **xun**≈dŕá hẹ=tho mi pahni
and.AS=1[SG]$_{PRO}$ also≈1.INCPL.R dress=DEL 1♀POSS shirt
‘I also wear my shirt.’ (Txt)

Finally, example (17a) further illustrates the discourse enclitic *=a* (glossed as CL). This enclitic is only found in the Southern Otomi languages (for example in Acazulco Otomi, Hernández-Green, 2014), where it is very common.[13]

2.1. Basics about word order in Tilapa Otomi. Conservative Otomi languages are V-initial, which is a widely accepted trait of Mesoamerican languages (Campbell *et al.* 1986). When two NP arguments are expressed in a transitive clause, which admittedly happens extremely rarely in natural discourse, the preference in T-Oto is to have a VOS

word order under broad focus. This is shown in (18) where the object is a nominalized complement phrase.[14]

(18) madikha hin≈a ʰpö≈[ra syö]$_{\textbf{OBJ}}$ [i kha'ni]$_{\textbf{SUBJ}}$
long.ago NEG≈[3]INCPL.R know.AS≈NMLZ speech PL person
'A long time ago, people didn't know how to speak.' (Txt)
(Lit. 'People didn't know speech.')

A VSO word order is also attested, shown in (19), but as Hernández-Green (2015) points out for Acazulco Otomi, it is only possible when there is no doubt as to which participant plays the role of agent. In other situations, when the referents of the two NPs are equally able to play the role of agent or patient, the order VOS is a fixed one. This is shown in (20); an example that can only have the interpretation in (a).

(19) kasi bi=zọ-ġi [a sku dihuntọ ni mi ʰta]$_{\textbf{SUBJ}}$
almost [3]CPL.R=TS/leave.AS-1OBJ DEF.SG DIM late DEM.SG 1♀POSS father
[nt'a=r nkü]$_{\textbf{OBJ}}$
one=SG house
'My late father almost left me a house.' (Txt)
(Lit. 'The deceased of my father almost left me a house.')

(20) bʷu=ʰtsi [a nana]$_{\textbf{OBJ}}$ [mu ʰta]$_{\textbf{SUBJ}}$
[3]CPL.CISLOC=take.animate DEF.SG woman 1♂POSS father
a. 'My dad went to fetch the woman.' (Txt)
b. *'The woman went to fetch my dad.'

Examples (18) and (19) additionally show that a clause may have constituents placed to the left of the predicate for informative prominence. Most adverbs and all quantifiers occur in that position, as seen in (18) with the temporal topical adverb *madikha* 'long ago' and in (19) with the quantifier *kasi* 'almost'.

Subject phrases (by way of DPs or pronominals) can also be found to the left of the predicate, like in (21) (fronted objects occur more rarely). When a subject is fronted, it may function as a contrastive topic. The extract in (22) is an illustration of this. Here the speaker establishes a contrast as to who will be the person (herself or the addressee) to do what.

(21) 'ne [a kha'ni]$_{\textbf{SUBJ}}$ bi='yehpi-gi
and DEF.SG person [3]CPL.R=SS/throw-1OBJ
'And the man threw me away.' (Txt)

(22) a. 'ne gu=ʰkundy=a 'ne gu=ʰpet'i=a,
and 2.CPL.IRR=grind.maize.AS=CL and 2.CPL.IRR=make.tortillas=CL
'And grind the maize, and make tortillas,'

b. mientras[=**ga**]$_{\textbf{SUBJ}}$ gu='ot'u≈ni, 'i
while=1[SG]$_{\text{PRO}}$ 1.CPL.IRR=make.AS≈DEM.SG food
'While I will make the food,'

c. kha [ñü=**k'e**]$_{\text{SUBJ}}$ nt'öntho ^{h}kundy=a i ^{h}tsuni
and(.in.contrast) PRTCL=2[SG]$_{\text{PRO}}$ quickly grind.AS=CL PL lime.soaked.maize
'And you grind the lime-soaked maize quickly.' (Txt)

In (22b), the enclitic *=ga* is hosted on the first element of the clause (i.e., conjunctions and complementizers; more examples of this behavior are examples (16) and (17c) above). In contrast, in (22c) the enclitic *=k'e* is hosted on the particle *ñü*, which serves as a morphological base to produce an independent pronoun (i.e., a pronominal that functions as a syntactic word), *e.g.* from *=k'e* → *ñü=k'e* 'you (SG)', from *=gwa* 'here' → *ñü=gwa* 'here', etc.[15]

In T-Oto, independent pronouns (i.e. based on *ñü*) always occur fronted in the clause, whether subject (23a) or object (24a). This means that they cannot occur after the predicate; as illustrated by the ungrammaticality of the (b) examples.

(23) a. 'nẹ=a syakaso [ñü=ga='mbe]$_{\text{SUBJ}}$ tŕá=má 'mb^{w}u='mbe ndöxi
and=CL in.case PRTCL=1$_{\text{PRO}}$=PL.EXCL 1.INCPL.R=PST exist=PL.EXCL Monday
'And WE were (here) on Monday just in case.' (Txt)

b. *'nẹ=a syakaso tŕá=má 'mb^{w}u='mbe [ñü=ga='mbe]$_{\text{SUBJ}}$ ndöxi
and=CL in.case 1.INCPL.R=PST exist=PL.EXCL PRTCL=1$_{\text{PRO}}$=PL.EXCL Monday
Intended reading: 'And we were (here) on Monday just in case.'

(24) a. pues [ñü=ga='mbe]$_{\text{OBJ}}$ ti=^{h}to^{h}karu-gu='mbe[16]
well/so PRTCL=1$_{\text{PRO}}$=PL.EXCL [3]CPL.IRR=befall.turn.on-1OBJ.AS=PL.EXCL
'Well, it's OUR turn.' (Txt)
(Lit. 'Well, (on) US it (the turn) will befall on us.')

b. *pues ti=^{h}to^{h}karu-gu='mbe [ñü=ga='mbe]
well/so [3]CPL.IRR=befall.turn.on-1OBJ.AS=PL.EXCL PRTCL=1$_{\text{PRO}}$=PL.EXCL
Intended reading: 'Well, it's our turn.'

A definite DP may also be introduced by *ñü*, and it often is, as shown in (25). Like independent pronouns, a DP marked with *ñü* can only occur in a fronted position (although also when extraposed to the right). To show this, compare example (19) above with its ungrammatical counterpart in (26).

(25) [ñü k'i sku tyü]$_{\text{SUBJ}}$ nkhonts'e ra=pe'ts'i
PRTCL DEM.PL DIM dead nobody.no.more [3]INCPL.R=have
to=ta... to=ta to'mi
who=[3]CPL.IRR who=[3]CPL.IRR wait.for
'Those dead people have nobody who would wait for them.' (Txt)

(26) *kasi bi=zọ-ġi [**ñü** a sku dihuntọ ni mi ^{h}ta]$_{SUBJ}$
almost [3]CPL.R=TS/leave.AS-1OBJ PRTCL DEF.SG DIM late DEM.SG 1♀POSS father
[nt'a=r nkü]$_{OBJ}$
one=SG house
Intended reading: 'My late father almost left me a house.'
(Lit. 'The deceased of my father almost left me a house.')

The function of the particle *ñü* is challenging and it deserves an independent study of its own. For the purposes of this paper, I propose to treat the particle *ñü* as a presentational marker that serves to (re)introduce a definite DP as topic. A clear example is (27) which is an instance of an external topic. In (27), there is a pause after the topic phrase separating it from the rest of the utterance (introduced by the discourse particle *pues* 'well/so' (Spanish *pues*) that always occurs in initial position). Further note in (27) that there is a pronominal enclitic on the verb cross-referencing the extracted object. Such enclitics are optional, but they are nonetheless common in this context. The use of *ñü* with a topicalized DP is facultative when the DP functions as subject, but it is obligatory when it functions as object (for exceptions see note [17]).

(27) [ñü i suntabo$_i$]$_{OBJ}$, pues hin≈gata thah=ka=yu$_i$
PRTLC PL cent well NEG≈1.ADLAT.IRR meet/find.AS=1[SG]$_{PRO}$=DEM.PL$_{PRO}$
'The cents, well so I'm not going to find them.' (Txt)

However, when the particle *ñü* is associated with a pronoun, I treat it as a morphological base devoid of pragmatic force. This is based on the observation that fronted pronouns can also be in focus and thus abound in focus constructions like clefts, as for example in (28). For such cases, I will consider that an independent pronominal phrase gains its pragmatic force, whatever it might be, from the pragmatic context and from its privileged syntactic position to the left of the predicate, but not from *ñü*.

(28) pero [ñü=a=ya]$_{FP}$ ke=a {giti k'oty=a}$_{CC}$=ya
but PRTCL=CL=DEM.PL$_{PRO}$ COP=CL[3] 2.CPL.IRR.ADV clean.AS=CL=DEM.PL$_{PRO}$
'But it's with that you're going to clean it.' (Txt)
(Lit. 'But THOSE (the ashes) are what you're going to wash it with.'

2.2. Basics about relative clause structure in Tilapa Otomi. The example of the cleft in (28), which includes the relative clause *giti k'otya* 'what you're going to wash it with' as CC brings us to relative clauses in Otomi. Relative clauses in Otomi are post nominal. The most common relative clause type is one that has a gap and is not introduced by a relative pronoun or a complementizer, that is, it is a contact or asyndetic relative clause. Examples in (29) illustrate a gap strategy of a relativized subject and object, respectively.

(29) a. tó='ötu='mbe [ni nkü$_i$ {ra=kha=ni __SUBJ$_i$}]
1.CPL.R=paint.AS=1PL.EXCL DEM.SG house [3]INCPL.R=exist=there
'We painted the house that is there.' (Txt)

b. pe rá=^{h}k<u>u</u> [a ’i$_{i}$ {t<u>ú</u>ti kha=’mbe ___OBJ$_{i}$}]
but INCPL.R.ST=be.flavory[3S$_{O}$] DEF.SG chili 1.HAB make=PL.EXCL
‘But the chili we make has full flavor.’ (Txt)

Crucially for the understanding of clefts in Otomi, headless relatives can also be asyndetic, and they often are. Examples are given in (30): in (30a) the clause functions as subject, and in (30b) it is an oblique encoding the stimulus.

(30) a. hin=á z<u>a</u> {rati kha}
NEG=INCPL.R.ST be.good[3S$_{O}$] [3]INCPL.R make[3OBJ]
‘What he does is not good.’ (Txt)

b. porke t<u>ú</u>d<u>ú</u> ^{h}tsü=a {bi=mbe^{h}ti-gi}
because 1.CPL.R get.scared=CL [3]CPL.R=SS/tell.DTR-1DAT
‘Because I got scared of what he told me.’ (Txt)

Up to this point, I have presented basics about Otomi grammar that are relevant to understand the grammar of a cleft sentence. But as clefts in Otomi feed on the syntax of the copular construction, the study of clefts in Otomi makes it imperative to study first the syntax of the copular construction, which as I show in the next section is far from being simple, because it involves a great amount of material that can be omitted.

3. Intransitive predication and the copular construction in Tilapa Otomi.

Clefts, especially in the form of IT-clefts (e.g. *it’s her that committed the crime*) or pseudo-clefts (e.g. *champagne is what I like best*) are focus constructions with an identificational function that are based on copular structure. Because of this, to understand clefts in Otomi, like in any other language, it is important to understand the constructions that are used in the language to convey identification, and to do so, it is convenient to place such constructions in the typological context of intransitive predication. For this, I follow Stassen’s (1997) typological proposal.

Stassen (1997) departs from a semantic domain that involves the traditional Western taxonomy of intransitive predication in four basic types of predicates: (a) event predicates (commonly encoded by verbs); (b) property or quality predicates (commonly encoded by adjectives); (c) locational predicates (commonly encoded by adverbs or adpositional phrases); and (d) class predicates (commonly encoded by nouns). Stassen’s typology is based on the different linguistic types that obtain from the distributional range of a given construction in a given language regarding the encoding of such predicates. When a language has an extended copular construction like English *be* in examples (31), it is found to encode the type of predicates in (b-d). I show below that the Otomi copular construction is mainly restricted to specificational instances, which are related to the class predication in (d), although they are not exactly the same.

(31) a. She runs
b. She is tall
c. She is here
d. She is a doctor

Otomi languages lack adjectives in predication (see Palancar 2006 for more details),[18] so Otomi uses verbs to predicate on a property.[19] In T-Oto, property verbs are patientive verbs (i.e., their S argument is indexed by the same markers that index the O of transitive verbs, hence the glossing 'S$_O$'). In such predications, the inflectional formative does not encode person.

(32) a. rá=hwë-k'i
INCPL.R.ST=be.sad-2S$_O$
'You're sad.' (Txt)

b. x≈á 'öxki rú nkü
ANT≈INCPL.R.ST be.nice[3S$_O$] SG.3POSS house
'Her house is quite nice.' (Txt)

To predicate location, Otomi languages use the locative existential verb *'mb^{w}u* which has a polysemy of meanings including 'exist or be present in a given place', 'live in a certain place' and 'be located at a certain place' (as in examples (9), (10) and (23)). The grammatical treatment of class predication is more relevant for clefts and I treat it in the following section.

3.1. The nominal predication construction in Tilapa Otomi. Class predication in Stassen's typology is an exemplary case of ascriptive predication. Loosely speaking, in ascriptive predication an entity is ascribed to a class so that as a member of that class it gains the attributes of the class. For class predication, the default encoding option in T-Oto is the nominal predication construction. This construction consists of a nominal converted into a non-verbal predicate that is inflected by means of some special inflection, which splits singular from plural. For the singular, the construction uses the same inflectional formatives that verbs use for the habitual aspect (see Table 1). This is shown in (33a). In the plural, as shown in (33b), a person marker (*t-* for the 1st person, *g-* for the 2nd person, and zero for the 3rd) is prefixed to the plural determiner in the incompletive.[20] A pronominal enclitic pointing to the person of the subject is also often used, especially with a 3rd person subject, like in (33c), but it is facultative. I indicate the predicate with parentheses and the subject with brackets.

(33) a. 'ne (tŕú=dentista)$_{PRED}$=[ga]$_{SUBJ}$
and 1.HAB=dentist=1[SG]$_{PRO}$
'And I am a dentist.' (Txt)

b. (t-i sku ñühü='mbe)$_{PRED}$
[INCPL.R]1-PL DIM otomi=PL.EXCL
'We're Otomi.' (Txt)

c. porke [ñü ni nk'inkhwa$_i$]$_{SUBJ}$ este=ru... (ru='nitxi)$_{PRED}$=ni$_i$
because PRTCL DEM.SG armadillo HES=[3]HAB [3]HAB=medicine=DEM.SG$_{PRO}$
'Because the armadillo is medicine.' (Txt)

While the default function of this construction is ascriptive, there are attested instances in my corpus with identificational function. A clear example is (34), which has an equative reading. Here the speaker states the identity between a man known by the nickname of 'the Mole' and himself as being the same person.[21] In the dialogue in (34a), the first speaker was aware that 'the Mole' was around somewhere, but she didn't know him so she was not able to recognize him. The dialogue happens when they meet. In (34b), the identity of the Mole is revealed.

(34) a. — ...hín=dŕá=má ʰpöh=a mas gwu=má 'ë=k'u,
NEG=1.INCPL.R=PST know.AS=CL if 2.IRR.CISLOC=PST come.AS=there
mas (grú=ʰTs'ubi)$_{PRED}$=[k'e]$_{SUBJ}$=a
if 2.HAB=Mole=2[SG]$_{PRO}$=CL
'...I didn't know if you'd be coming from over there, if you would be the Mole.'

b. — bweno, pues (tŕú=ʰTusa)$_{PRED}$=[ga]$_{SUBJ}$
well so 1.HAB=Mole=1[SG]$_{PRO}$
'Well, so I am the Mole.' (Txt)

But such identificational uses of the nominal predication construction are exceptional. By far the most natural way to establish a specificational or equative relation in Otomi is by way of the copular construction.

3.2. The copular construction in Tilapa Otomi. A canonical example of the copular construction in T-Oto is (35). The structure is schematized in (36) as 'Variant 1'.

(35) kẹh=a [ni mi t'yu$_i$]$_{SUBJ}$ (a mbahkö)$_{COMPL}$=ni$_i$
COP.AS=CL[3] DEM.SG 1♀POSS son DEF.SG priest=DEM.SG$_{PRO}$
'My son is the priest.'

(36) VARIANT 1

COP	[DP$_i$] SUBJ	[DP]$_{COMPL}$	=pro$_i$

Example (35) instantiates Variant 1 with its four main elements: (i) an overt copula *ke*, which is a non-verbal predicate; (ii) a DP that functions as subject; (iii) a DP that functions as the complement of the copula; and (iv) a pronominal enclitic that cross-references the subject.

I employ the term 'canonical' referring to example (35) in the sense this term receives in Canonical Typology, as developed in Corbett (2000). In this context, 'canonical' refers to an idealized structure that is meant to serve as a reference point to measure actual structures of the language, which might or might not be close to this ideal or canon. Canons are conceived of as tools to understand structural diversity. I treat variant (36) as a canon because it has all the constituting elements overt. However, the variant is not naturally found in Otomi discourse (note that example (35) is elicited). The fundamental reason behind the rarity of (36) is the fact that it involves two overt nominal references. Other variants of the construction are more common

In this connection, variants of the construction often involve a topical subject, whose referent can be expressed overtly by an anaphoric pronoun like in (37) or covertly through the indexation of person on the copula, like in (38) which is a common phrase

used for closure in storytelling. The schema for both these possibilities appears in (39) and (40), where I have only represented material that is phonologically overt. This is done in order not to impose an analysis as to where the covert elements are to be located in the structure (see next section for a possible analysis).

(37) kẹ=a=[k'u$_i$]$_{SUBJ}$ (i loka)$_{COMPL}$=k'u$_i$
COP=CL=3PL$_{PRO}$ PL mad.woman=3PL$_{PRO}$
'They are the mad women.'[22] (Txt)

(38) pues kẹh=a [___ $_i$]$_{SUBJ}$ (a rraso)$_{COMPL}$='ä $_i$
well/so COP.AS=CL[3$_i$] DEF.SG reason=3SG$_{PRO}$
'That's the story.' (Txt)

(39) VARIANT 2 | COP [=pro$_i$] $_{SUBJ}$ [DP]$_{COMPL}$ =pro$_i$ |

(40) VARIANT 3 | COP [DP]$_{COMPL}$ =pro$_i$ |

Similarly, as pointed out above in §2.2, we may have the subject in fronted topical position, as shown in (41a) with an independent pronoun and in (41b) with a definite DP marked with *ñü*. The schema of this variant is given in (42).

(41) a. pe [ñü=a=ni$_i$]$_{SUBJ}$ hingu≈kẹh=a [a doktora]$_{COMPL}$=ni$_i$
but PRTCL=CL=DEM.SG$_{PRO}$ NEG≈COP.AS=CL[3] DEF.SG she.doctor=DEM.SG$_{PRO}$
'But she is not the doctor (of the village clinic).'[23] (Txt)

b. [ñü na kha'ni$_i$]$_{SUBJ}$ kẹh=a [a mbahkö]$_{COMPL}$=na$_i$
PRTCL DEM.SG man COP.AS=CL[3] DEF.SG priest=DEM.SG$_{PRO}$
'This man is the priest (a specific person in the village).'

(42) VARIANT 4 | [pro/DP$_i$] $_{SUBJ}$ COP [DP]$_{COMPL}$ =pro$_i$ |

In other examples, we have no overt copula.[24] When this happens along with an overt subject phrase, I interpret that the subject phrase is always fronted. For example, in (43) we have a fronted pronoun: an enclitic hosted on a conjunction (44a) or an independent pronoun (43b). In (44) we have a fronted DP. In (44a), the occurrence of the presentational particle *ñü* is expected, but so is its absence in (44b), because NPs headed by possessors do not take *ñü*. I take such cases as instances of Variant 5 in (45), which is a version of Variant 4 in (42) but with an elided copula; a relatively common cross-linguistic pattern (Pustet 2003).

(43) a. 'ne[=a=ni $_i$]$_{SUBJ}$ [a mba^{h}kö]$_{COMPL}$=ni$_i$
and=CL= DEM.SG$_{PRO}$ DEF.SG priest=DEM.SG$_{PRO}$
'And he's the priest.' (Txt)

b. [ñü='a$_i$]$_{SUBJ}$ [yi yú ^{h}pa i sku tyü]$_{COMPL}$='a$_i$
PRTCL=3SG$_{PRO}$ DEM.PL PL.3POSS day PL DIM dead=3SG$_{PRO}$
'It (that celebration being talked about) is the days of the dead.' (Txt)

(44) a. pe [ñü a karro$_i$]$_{SUBJ}$ [rú karro mi t'yu]$_{COMPL}$
but PRTCL DEF.SG car SG.3POSS car 1♀POSS son=3SG$_{PRO}$
'But the car is my son's car.' (Txt)

b. [mi t'yu$_i$]$_{SUBJ}$ [a mbahkö]$_{COMPL}$=ni$_i$
1♀POSS son DEF.SG priest=DEM.SG$_{PRO}$
'My son is the priest.'

(45) VARIANT 5 [pro/DP$_i$]$_{SUBJ}$ [DP]$_{COMPL}$ =pro$_i$

Finally, there are also instances of the copular construction where neither the subject phrase nor the copula is present, as in (46), representing Variant 6 in (47).

(46) (a gu=pë)$_{COMPL}$=ni$_i$
DEF.SG CLF.HUM=SS/steal=DEM.SG$_{PRO}$
'That (pointing to a man) is the thief.' (Txt)

(47) VARIANT 6 [DP]$_{COMPL}$ =pro$_i$

3.3. The syntax of the copular construction in Tilapa Otomi. In the previous section, I have shown that the copular construction consists of various elements, but because some of them can be omitted under certain circumstances, we obtain the different surface variants in Table 4.

INSERT TABLE 4

It is desirable to attain a unified account of these six variants. In this section, I propose one such possible analysis, there could be others. I will start with the pronominal enclitic that cross-references the subject, which I take to play a fundamental role in revealing the inner structure of the construction.

The enclitics are not obligatory when the copula is overt (Variants 1-4). However, I should emphasize that using them is by far the speakers' default option, which means that in natural discourse they are more often present than absent. But interestingly, when the copula is covert as in Variant 5 the enclitic is obligatory (unless the predicate is a heavy constituent like in (44a) above, in which case it can be omitted). In Variant 6 it must, nonetheless, occur. In my analysis, the obligatory character of the enclitic in Variants 5 and 6 speaks of its structural significance.

To choose what enclitic to use, the default mechanism is to have a copy of the subject pronoun, as in (48a), repeated from (43a), or a copy of the head determiner of the DP subject, like in (48b). Failing that, the default elsewhere option is the pronoun *='a* for singular and *=k'u* for plural, like in examples (43b) and (37), respectively.[25]

(48) a. 'ne [=a=**ni** $_i$]$_{SUBJ}$ (a mba^{h}kö)$_{COMPL}$=**ni**$_i$
and=CL= DEM.SG$_{PRO}$ DEF.SG priest=DEM.SG$_{PRO}$
'And he's the priest.' (Txt)

b. [ñü **na** kha̲'ni$_i$]$_{SUBJ}$ kẹh=a (a mba̲hkö)$_{COMPL}$=**na**$_i$
PRTCL DEM.SG man COP.AS=CL[3] DEF.SG priest=DEM.SG$_{PRO}$
'This man is the priest.'

Examples like (48) raise the important question as to what DP is subject in the copular construction. In this regard, I propose that the default word order of the Otomi copular construction is one in which the complement of the copula occurs in final position. Apart from the enclitics, which for the most part are informative about what DP is subject, evidence that the complement occurs in final position also comes from the rare and exceptional uses of the copular construction as an extended frame to encode class predication. In such cases, like in (49), the predicate always occurs in the same position as the complements in the identificational examples above.

(49) a. pero komo[=yu̲$_i$]$_{SUBJ}$ (ga=nde̲zna)$_{COMPL}$=yu̲$_i$
but as=DEM.PL$_{PRO}$ CLF.HUM=be.Nahuatl=DEM.PL$_{PRO}$
'But as those are Nahuatl indians.' (Txt)

b. kha [ñü=a=yu̲$_i$]$_{SUBJ}$ (simarro-gá-te̲thö)$_{COMPL}$=yu̲$_i$
and PRTCL=CL=DEM.PL$_{PRO}$ violet.maize-LIG-corncob=DEM.PL$_{PRO}$
'And those (the maize cobs) are violet maize corncobs.' (Txt)

To account for the syntax of the copular construction, I propose that the six different variants in Table 4 can be interpreted as surface outcomes of the constructs in Table 5. In my analysis, all such constructs are seen as departing from an underlying structure that is not realized on surface level, which I treat as 'Construct I'. This departing structure, as it were, is based on the natural VOS word order of the language.

INSERT TABLE 5

From the underlying VOS structure in Construct I, transposed here as COP-COMP-SUBJ,[26] we obtain the structure in Construct II as a first arrangement of the underlying syntax. The change from Construct I to II is a specific requirement of the syntax of the copular construction by which the subject phrase has to move to a fronted position. Such a movement is perhaps motivated by the nature of the specificational semantics of the construction that favors a topical treatment of the entity to be identified and a focus treatment of the identifying complement.

In this movement of the subject phrase from Construct I to II, a pronominal enclitic is left *in situ* at its place serving as a trace. This would explain two things: (i) the occurrence of the pronominal enclitic in the construction, which is an interesting feature of Otomi syntax; and (ii) the lexical copying that the pronominal commonly manifests from determiner/pronominal material present in the fronted phrase.

The omission of the enclitic is only allowed in cases where there are enough structural cues in the utterance (by way of other overt elements) that would make the utterance an unambiguous instance of the copular construction. For instance, in example (44a) above –an instance of Variant 5– while its elision is prompted by the heavy constituent, the juxtaposition of the two DPs is sufficient to trigger an interpretation as a copular structure. A further illustration of how relevant the enclitic is shown in examples

(50) and (51). Example (50) is a case of a simple predication, and it cannot be anything else but that. However (51), with a pronominal enclitic, is a cleft and cannot be anything else but that.

(50) porke [ñü ni rú ngopho]$_{SUBJ}$ (bi=kokhi)$_{PRED}$
because PRTCL DEM.SG SG.3POSS brain [3]CPL.R=bleed
'Because her brains bled.' (Txt)
*'Because it was her brains that bled.'

(51) porke [ñü ni rú ngopho$_i$]$_{FP}$ {bi=kokhi}$_{CC}$=**ni**$_i$
because PRTCL DEM.SG SG.3POSS brain [3]CPL.R=bleed=DEM.SG$_{PRO}$
'Because it was her brains that bled.'
*'Because her brains bled.'

This means that if speakers have the cleft structure in mind, the enclitic must occur if they choose to omit the copula like in (51).[27] The structure in (51) –an instance of Construct (d)– indicates that all the functional load of the copular construction lies on the pronominal enclitic alone. The same happens in Construct (b). This is a related to a situation observed cross-linguistically in Stassen (1997:76ff), where determiners are analyzed as having copular functions.[28] In the Otomi case, because the language has an overt copula, the copular analysis of the clitic *=ni* in cases like (51) (or its counterparts) is not desirable, even though it serves as a fundamental token to hint at the fact that the copular structure is at use.

While the movement of the subject phrase from Construct I gives us Construct II, the latter also operates at a more abstract level, because a fronted DP in T-Oto is based on *ñü*, like in (52), repeated from (41a) above, which I analyze as an instance of Construct (c) and not Construct II. The next surface realization of Construct II is Construct (a), where the copula has moved to its natural initial position and it is consequently followed by a subject phrase, which occurs internally and cannot be introduced by *ñü*. This is shown in (53), repeated from (35) and (37) above.

(52) pe [ñü=a=ni$_i$]$_{SUBJ}$ hingu≈kẹh$_j$=a [___t_i] [___t_j] [a doktora]$_{COMPL}$=ni$_i$
but PRTCL=CL=DEM.SG$_{PRO}$ NEG≈COP.AS=CL[3] DEF.SG she.doctor=DEM.SG$_{PRO}$
'But she (that one) is not the doctor.' (Txt)

(53) a. kẹh$_j$=a [ni mi t'yu$_i$]$_{SUBJ}$ [___t_j] [a mbahkö]$_{COMPL}$=ni$_i$
COP.AS=CL[3] DEM.SG 1♀POSS son DEF.SG priest=DEM.SG$_{PRO}$
'My son is the priest.'

b. kẹ$_j$=a=[k'u$_i$]$_{SUBJ}$ [___t_j] [i loka]$_{COMPL}$=k'u$_i$
COP=CL=3PL$_{PRO}$ PL mad.woman=3PL$_{PRO}$
'They are the mad women.' (Txt)

Having accounted for the syntax of the copular construction, the syntax of clefts in T-Oto is relatively straightforward, as we will see in the next sections.

4. Clefts in Tilapa Otomi.

In the previous section, I have shown that the copular construction in Otomi is primarily specificational and that it has multiple surface variants. Clefts in T-Oto are based on the copular construction. Consider the cleft in (54), which is based on Construct (a) as instantiated in example (55), repeated from (35) above.

(54) kẹh=a [Papa Boniphasio$_i$]$_{FP}$ {bi=kha}$_{CC}$='a$_i$
COP.AS=CL[3] Pope Boniface [3]CPL.R=do[3OBJ]=3SG$_{PRO}$
'It was Pope Boniface the one who did it.' (Txt)
(Lit. 'POPE BONIFACE was the one who did it.')

(55) kẹh=a [ni mi t'yu̲$_i$]$_{SUBJ}$ (a mba̲hkö)$_{COMPL}$=ni$_i$
COP.AS=CL[3] DEM.SG 1♀POSS son DEF.SG priest=DEM.SG$_{PRO}$
'My son is the priest.'

I analyze example (54) as a cleft, and not just as a simple instantiation of the copular construction, because of its semantic and pragmatic nature. Following Lambrecht (1994; 2001), I take clefts to be biclausal structures specialized for the marking of focus, which can be paraphrased in simple clauses with which they share the same propositional content. In this regard, (54) is semantically similar to (56).[29]

(56) [ñü a Papa Boniphasio]$_{SUBJ}$ bi=kha
PRTCL DEF.SG Pope Boniface [3]CPL.R=do[3OBJ]
'Pope Boniface did it.'

Based on the structural identity between (54) and (55), I analyze the components of clefts as having exactly the same syntactic function as the ones we find in the copular construction. In the cleft, the FP has the role of subject and the CC introducing the common ground is treated as the complement of the copula.

Similarly, there is surface variation in the realization of clefts. For instance, examples in (57) are instances of FPs encoded with pronominals. In (57a-b), the pronoun refers back to a participant whose referent is recoverable from the discourse context. In (57c), *gwa* 'here' introduces a location. In the examples in (58) the FP is fronted: (58a) is based on Construct (c) and (58b), without an overt copula, on Construct (d).

(57) a. kẹ^{y}n=ts'e̲=a=['a$_i$]$_{FP}$ {grá=nde=^{g}wi}$_{CC}$='a$_i$?
COP.AS=just=CL=3SG$_{PRO}$ 2.INCPL.R=want=PL=3SG$_{PRO}$
'Is it just that what you (PL) want?' (Txt)
(Lit. 'Is IT just what you (PL) want?')

b. kẹh=[na$_i$]$_{FP}$ {ra='a̲-k'i}$_{CC}$=na$_i$
COP.AS=DEM.SG$_{PRO}$ [3]INCPL.R=ask.for.AS-2OBJ=DEM.SG$_{PRO}$
'This is what she's asking you.' (Txt)

c. ken[=gwa$_i$]$_{FP}$ {'ab^{w}u tú=m-pe='mbe}$_{CC}$=gwa$_i$
COP.AS=here where 1.CPL.R=ANTI-work.AS=PL.EXCL=here
'It's here where we worked.' (Txt)
(Lit. 'HERE is where we worked.')

(58) a. 'ne=a [ñü=a$_i$]$_{FP}$ kẹh=a {tŕá='a-k'u=wi}$_{CC}$='a$_i$
and=CL PRTCL=3SG$_{PRO}$ COP.AS=CL[3] 1.INCPL.R=ask.for.AS-2OBJ.AS=PL=3SG$_{PRO}$
'And it's that what I am asking you.' (Txt)
(Lit. 'And THAT is what I am asking you (PL).'

b. [ñü=gwa$_i$]$_{FP}$ {ra=^{h}po≈ra tehẹ}$_{CC}$=gwa$_i$
PRTCL=here [3]INCPL.R=spring.AS≈SG water=here
'It's here where the water springs.' (Txt)
(Lit. 'HERE is where the water springs.')

Because Otomi clefts mimic the copular construction, in a typology of clefts they are to be categorized as pseudo-clefts. However, there is at least one element occurring in clefts which is unique to them. The element in question is the pronoun *ko* in T-Oto (its cognate is *go* in other Otomi languages). Examples (8) and (9) above have *ko*, another example is (59), an instance of Construct (a).

(59) kẹh=a [na mi mimukha$_i$]$_{FP}$ {**ko**$_i$ bi=^{h}paty=a}$_{CC}$=na$_i$
COP.AS=CL[3] DEM.SG 1♀POSS sister.in.law FOC [3]CPL.R=heat.up[3OBJ]AS=CL=DEM.SG$_{PRO}$
'It's my sister-in-law here that burnt it (the hut).' (Txt)

Because *ko* is unique to clefts, its occurrence suggests that the syntax of clefts is somehow independent from the copular construction. In the next section, I analyze *ko* in T-Oto as a 'focus extraction pronoun'[30] (glossed as FOC) that encodes the variable in the CC.

5. The element *ko* in Tilapa Otomi.

Clefts in Otomi, bear the element *ko* in the CC, like in (60), an instance of Construct (d).

(60) 'neh=a [mi ts'ihwe$_i$]$_{FP}$ {**ko**$_i$ ru=^{h}pety=a}$_{CC}$='a$_i$
and.AS=CL 1♀POSS daughter.in.law FOC [3]HAB=make.tortillas.AS=CL=3SG$_{PRO}$
'It's my daughter in law who makes the tortillas.' (Txt)
(Lit. 'MY DAUGHTER IN LAW is who makes the tortillas.')

In §5.3, I propose that *ko* is a non-canonical relative pronoun that stands in the CC for the noun that has been extracted to focus in the construction. But to show what *ko* is, it is convenient to show first what it is not, because in examples like (59) or (60) *ko* can raise the natural expectation that it is a canonical relative pronoun occurring in the canonical syntactic position for relative pronouns, which in T-Oto would be to the left edge of the relative clause. In what follows in §5.1 and §5.2, I show that neither assumption is true. In other words, *ko* is neither a canonical relative pronoun nor does it occur in initial position. Let us start with relative pronouns.

5.1. The element *ko* is not a canonical relative pronoun. Drubig and Schaffer (2001) judge that the grammatical structure behind clefts remains unsolved, but they acknowledge that the only existing agreement about them is that the subordinate structure in them must be analyzed as a case of a relative clause. For Otomi, the relative clause status of the CC is straightforward. For example, as pointed out in §2.3 both headed and headless relative clauses in Otomi can be asyndetic relatives (i.e., not introduced by a complementizer or a relative pronoun). This provides good leverage for an analysis of the CC in (61) as a relative clause.[31]

(61) abeses ke [thehwö$_i$]$_{FP}$ {**tríti** **’ün$_i$**}$_{CC}$
sometimes COP fish 1.HAB give[3OBJ]
‘Sometimes it's fish what I give him.’ (Txt)
(Lit. ‘Sometimes FISH is what I give him.’)

Similarly, a special inflection I treat as ‘adverbial inflection’ (analyzed as having a registration function in Hernández-Green 2016) is used in a relative clause in T-Oto to index the relativization of instrumentals.[32] This happens in headed relatives (62), in headless relatives (63), and in the CC (64).[33]

(62) xpi=’nah-ku≈[ru **xabǫ$_i$** {**giti$_i$** ^{h}pendy=a ____ INSTR$_i$}]
[3]PRF=give.to.1/2.AS-1OBJ.AS≈SG soap 2CPL.IRR.ADV wash.clothes.AS[3OBJ]=CL
‘She gave me the soap with which you'll wash it.’ (Txt)

(63) hin=grá ^{h}pödi [{**te$_i$** **giti$_i$** xuti=a ____ INSTR$_i$}]
NEG=2.INCPL.R know what 2CPL.IRR.ADV wash.dishes.AS[3OBJ]=CL
‘You don't know with what you'll wash it.’ (Txt)

(64) para kẹh=a[**=ya$_i$**]$_{FP}$ {**giti$_i$** ^{h}pendy=a ____ INSTR$_i$}$_{CC}$=ya$_i$
PURP COP.AS=CL=DEM.PL$_{PRO}$ 2.CPL.IRR.ADV wash.clothes.AS[3OBJ]=CL=DEM.PL$_{PRO}$
‘So that it's these things you'll wash it with.’ (Txt)
(Lit. ‘So that THESE ONES are what you'll wash it with.’)

Schachter (1973) is a pioneer in studying the commonalties in structure between clefts and relative clauses, as in the (b) examples in (65) and (66), which are possible answers to the questions in (a), and whose structure is only disambiguated prosodically.

(65) a. –who did you see?
b. –it's [the WOMAN] {that I saw}

(66) a. –who is it (that came)?
b. –it's [the woman {that I SAW}]

Lambrecht (1994:233) explains the fact that the common ground in a cleft like (65b) is encoded with a relative clause appealing to the nature of relative clauses as “a clause construction which is typically (though not necessarily) reserved for the coding of pragmatically presupposed propositions”[34]. For example, (65b) operates with the shared knowledge that there is somebody that the speaker has seen, but whose specific identity

is under contention at the time of the speech act. Similarly, the relative clause in (66b) encodes the shared knowledge that the speaker has seen a specific woman whose identity she takes as known to the hearer. This common ground is evoked in (66b) as a means to help the addressee identify the right woman.[35]

The same happens in Otomi. A sentence like (67) in T-Oto can be ambiguous between an interpretation as a CC and as a restrictive relative clause. But the ambiguity can only happen when the enclitic is omitted. If present, like in example (68), the sentence can only have a cleft interpretation, which as noted in (51), is a property associated with the copular construction.

(67) *ke ni nana tú̲ ñüa*

a. ke [ni nana $_i$]$_{FP}$ {tú̲=ñü=a $\_\__i$ $_{OBJ}$}$_{CC}$
COP DEM.SG woman 1.CPL.R=see=CL
(*–Who did you see?*) –'It was the woman I saw.' (Lit. 'THE WOMAN was who I saw.')

b. ke [ni nana$_i$ {tú̲=ñü=a $\_\__i$ $_{OBJ}$}]$_{SUBJ}$
COP DEM.SG woman 1.CPL.R=see=CL
(*–Who did it?*) – 'It was the woman I saw.'

(68) ke [ni nana$_i$]$_{FP}$ {tú̲=ñü=a $\_\__i$ $_{OBJ}$}$_{CC}$=**ni**$_i$
COP DEM.SG woman 1.CPL.R=see=CL=DEM.SG$_{PRO}$
a. (*–Who did you see?*) – 'It was the woman I saw.' (Lit. 'THE WOMAN was who I saw.')
b. (*–Who did it?*) *– 'It was the woman I saw.'

In T-Oto, the only type of relative clause allowed to serve as CC is the asyndetic type.[36] For example, other types of relatives, such as one introduced by an interrogative pronoun like in (69a) can only be used with the relative clause interpretation, never as a CC, even if the enclitic were to be used to reinforce a possible copular reading. Similarly, the relative pronoun *te* 'what', which can only occur in headless relatives, like (70a), cannot occur in clefts, as shown in (70b).

(69) a. ke [ni nana $_i$ {**to**$_i$ tú̲=ñü=a $\_\__i$ $_{OBJ}$}] $_{SUBJ}$
COP DEM.SG woman who 1.CPL.R=see=CL
(*–Who did it?*) – 'It was the woman I saw.'

b. *ke [ni nana $_i$]$_{FP}$ {**to**$_i$ tú̲=ñü=a $\_\__i$ $_{OBJ}$}$_{CC}$(=ni$_i$)
COP DEM.SG woman who 1.CPL.R=see=CL=DEM.SG$_{PRO}$
(*–Who did you see?*) Intended reading: 'It was the woman who I saw.'

(70) a. gu̲=tsi=hu̲ {**te** bi=münts'i ni mi t'yu̲=ga}
1.CPL.IRR=ingest=PL.INCL what [3]CPL.R=gather DEM.SG 1♀POSS son=1[SG]$_{PRO}$
'We're going to eat what my son gathers.' (Txt)

b. *kẹh=a=[ni$_i$]$_{FP}$ {**te** bi=münts'i ni mi t'yu̲=ga}$_{CC}$=ni$_i$
COP.AS=CL=DEM.SG$_{PRO}$ what [3]CPL.R=gather DEM.SG 1♀POSS son=1[SG]$_{PRO}$=DEM.SG$_{PRO}$
Intended reading: 'That's what my son gathers.'

In contrast, *ko* has the opposite behavior to *to* and *te*, because it can ONLY occur in a cleft. More specifically, in T-Oto it can only occur in a CC. For instance, the fact that it cannot be used as a canonical relative pronoun is shown in the ungrammaticality of (71). In contrast, a construct like (72) is and can only be interpreted as a cleft, even in the absence of other structural cues.

(71) *bi=zoh=a [ni nana$_i$ {**ko**$_i$ tú=ñü=a __$_i$ OBJ}]
[3]CPL.R=talk.to.AS=CL DEM.SG woman REL? 1.CPL.R=see=CL
Intended meaning: 'He talked to the woman I saw.'

(72) [ñü ni nana$_i$]$_{FP}$ {**ko**$_i$ tú=ñü=a}$_{CC}$
PRTCL DEM.SG woman FOC 1.CPL.R=see=CL
'It was the woman I saw.' (Lit. 'THE WOMAN was who I saw.')

Example (72), which is based on Construct (d) in Table 5, is the preferred encoding option when the FP is a fronted pronominal, like in (73).[37]

(73) a. [ñü=ga$_i$]$_{FP}$ {ko$_i$ gwu=testiga}$_{CC}$=ga$_i$
PRTCL=1[SG]$_{PRO}$ FOC 1.CPL.IRR.TRANSLOC=testify=1[SG]$_{PRO}$
'It's me who's going to testify.' (Txt)

b. pe [ñü=k'e$_i$]$_{FP}$ {ko$_i$ giti dispone}$_{CC}$=k'e$_i$='na
but PRTCL=2[SG]$_{PRO}$ FOC 2.CPL.IRR decide=2[SG]$_{PRO}$=QUOT
'But apparently it's you who decides.' (Txt)

5.2. The element *ko* does not occur to the left edge of the CC. In the previous section, I have shown that *ko* is not a canonical relative pronoun. More remarkable is the fact that it occurs at a position lower in the clause than the position commonly taken by complementizers or canonical relative pronouns. This can be seen in the behavior of negation. Negative adverbs in Otomi occur to the left edge of the pre-V zone, in such a way that anything that follows them to the right occupies a lower position in the clause. This is illustrated in the restrictive relative clause in (74), where the negation marker *hiN*≈ clusters with the inflectional formative *bi* of 3rd person completive realis and precedes it, while the interrogative relative pronoun *to* 'who' occurs further to the left at the edge of the clause. Note the ungrammaticality of (75) where negation has been raised over the interrogative/relative pronoun.

(74) tú=ñü [ni kha'ni {**to** **him**≈bi 'nah-ku≈mẹ^{y}ndyu}]
1.CPL.R=see DEM.SG man who NEG≈[3]CPL.R give.to.1/2-1OBJ.AS≈money
'I saw the man who didn't give me money.'

(75) *tú=ñü [ni kha'ni {**hin**≈**to** bi='nah-ku≈mẹ^{y}ndyu}]
1.CPL.R=see DEM.SG man NEG≈who [3]CPL.R=give.to.1/2-1OBJ.AS≈money
Intended reading: 'I saw the man who didn't give me money.'

In contrast to relative pronouns such as *to* 'who', negation always precedes *ko*. This is shown in (76), which are instances of a cleft where semantic negation stays within the CC and has not been raised to the copula.

(76) a. pe {hin≈**ko**$_i$ gu=khüt'u}$_{CC}$=[k'e$_i$]$_{FP}$;
but NEG≈FOC 2CPL.IRR=pay.AS[3OBJ]=2[SG]
ta=küti≈ni di t'yu
[3]CPL.IRR=SS/pay.AS[3OBJ]≈DEM.SG 2POSS son
'But it's not you who will pay for it; your son will pay for it.' (Txt)
(Lit. 'But YOU are the one who isn't going to pay for it...')

b. [ni=r Hwä$_i$]$_{FP}$ {hin≈**ko**$_i$ ru=xah-pu-te}$_{CC}$=ni$_i$
DEM.SG=SG John NEG≈FOC [3]HAB=teach.AS-3DAT.AS-ANTIP=DEM.SG$_{PRO}$
'It's not John who's a teacher.' (Archivo de Lenguas Indígenas de México, 121)
(Lit. 'JOHN is the one who doesn't teach.')

Examples like (76) indicate that *ko* is a pronominal element located in the pre-V zone. This means that *ko* occurs at a lower position in the clause, and not initially.[38]

5.3. The element *ko* as a focus extraction pronoun. We have seen that *ko* is not a canonical relative pronoun and it does not occur to the left edge of the CC. What is *ko* then? To attain a possible analysis we have to consider the fact that it can only occur in clefts, in such a way that its function has to be necessarily interpreted in the context of the semantic-pragmatic structure of clefts. In such a context, I propose that *ko* has a pronominal function which involves standing as a variable in the CC for the referent that has been extracted to a focus position, encoded as the subject of the copula. In this light, *ko* can still be analyzed as a relative pronoun, but as a very special one, because it has the specification that the relativized noun it stands for is in focus. An analysis in terms of a relative pronoun is not fully at odds typologically, because as *ko* occurs in the pre-V zone of the predicate of the CC, it could be treated as a resumptive pronoun. Following Demirdache's (1991; 1997) proposal that resumptive pronouns are true relative pronouns, one can treat *ko* as a resumptive relative pronoun, even though it is not one that occurs *in situ*.[39]

Similarly, although *ko* is facultative, when it is present, it is also overwhelmingly used to refer back to a FP with a human referent. This also means that when the referent of the FP is not human or at least not highly animate, the CC occurs without *ko* as a default. I take such tendencies as significant. In §8, I show that they represent remnants of a historical situation where *ko* was a pronoun for humans only. Nowadays in T-Oto, it may also be used with inanimate referents, although only rarely.[40]

6. The syntax of clefts in Tilapa Otomi.

6.1. Cleft variants. In the previous section, we have seen that clefts are not only based on the copular construction, but that they also have surface variation like the copular construction. To this situation we have to add the facultative character of the focus extraction pronoun *ko*. All logical possibilities that obtain are given in Table 6. The table

also includes Construct (e), which is specific to clefts. All possibilities are attested in my corpus, except one, but they differ much as to usage frequency.

INSERT TABLE 6

The minimal formal expression of a copular construction is Construct (b), where the pronominal enclitic is obligatory,[41] because it is the only element that serves as a cue to the addressee that the speaker is using a copular construction. But clefts allow a step further: because the focus extraction pronoun can only occur in a CC, its presence is sufficient to represent a case of either Construct (b) or Construct (d). For example, (77), which has no pronominal enclitic, is a case of Construct (d). Incidentally, because the CC is a heavy constituent at times, there are more chances with clefts that speakers avoid using an overt enclitic than in counterpart constructs in the copular construction.[42]

(77) ’ne [a mi t’yu̲$_i$]$_{FP}$ {**ko**$_i$ ru̲=’ë̲m-bi ar doktora}$_{CC}$=’na
and DEF.SG 1♀POSS son FOC [3]HAB=say.AS-3DAT SG she.doctor=QUOT
‘And apparently it's my son who tells the doctor.’ (Txt)
(Lit. ‘And apparently MY SON is the one who tells the doctor.’)

Example (78) is an instance of a cleft by way of Construct (b). Such examples are rare but happen when the referent of the FP has been introduced in previous discourse.

(78) xtú̲=phorma=’mbe [**nt’a grupo̲$_i$**]=wa *al DIF*
[3]PRF.R=form=PL.EXCL one group=here at.the.DIF
‘We created a group here at the DIF.’

má=’mb^{w}u̲ nt’a rrehidor=wa {bi=t-’ët’i ’a Nintxi}
[3]IMPF=exist one councilor=here CPL.R=PASS-place[3SO] P ST
‘There was a councilor that was elected in Santiago Tianguistenco.’

’ne {**ko**$_i$ bi=syondy=a}$_{CC}$=**k’u̲$_i$**
and FOC [3]CPL.R=SS/search.AS[3OBJ]=CL=3PL$_{PRO}$
‘And it was them [those at the DIF] that went to find him.’ (Txt)
(Lit: ‘And THEY (were) the ones who went to find him.’)

The cleft in an example like (78) still has the two overt elements of Construct (b). Examples of Construct (b) where the enclitic is omitted count as the minimal expression of a cleft consisting of just a CC with *ko*. Such a situation happens only rarely, mainly because clefts are more commonly encoded by way of other constructs in natural discourse. But examples are attested in my corpus. One such example is (9) above. Another is (79), which comes from a reported conversation between a husband and a wife. The couple has a teenager boy who is visiting his aunt in the US with his mother, and decides to stay on in the US. The mother comes back to Mexico and tells her husband about it. The father is worried as to how to deal with a situation that involves illegal immigration. The wife appeases him saying that her sister is going to tell the police that he, their son (and topic of the conversation), is her (the sister’s) own son.[43]

(79) ta=’ñëndy=a: “–{ko mi t’yu̲}CC”
[3]CPL.IRR=SS/say.AS=CL FOC [INCPL.R]1♀POSS son[3SO]
‘She’s going to say: “–He’s my son”.’ (Txt)
(Lit. ‘(It’s him) who’s my son.’)

The different variants in Table 6 contrast in degree of usage. Construct (a), which is not attested in my corpus in the copular construction, is however very common in clefts. Allegedly, this is because the subject phrase in an Otomi cleft, being in focus, is particularly informative and is thus commonly overt. Construct (d) with *ko* is the default encoding of a cleft with a fronted pronominal FP, like examples in (73) above. But without *ko*, construct (d) is practically unattested in clefts. In fact, example (58b) above is the only example in my corpus. The general trend is to avoid being over informative on the encoding level, while making sure at the same time that the structure is understood as an instance of a cleft. This is accomplished by eliding the copula, but using *ko* or by using the copula while avoiding *ko*. This could explain why Construct (c) with a fronted FP and *ko* is unattested.[44]

The last construct in Table 6 is Construct (e), and it involves a truncated cleft, which following Patten (2012) I take to be a shortened form of a cleft that only includes the copula and the FP, and sometimes the personal enclitic. This variant is only used in answers to questions. I introduce it in the following section.

6.2. Clefts in questions and answers. Questions in Otomi are often framed in cleft form. This is a natural way to place a Wh-element in focus. But they are also used in responses to introduce the participant in narrow focus. This is done by way of truncated clefts, where the CC is elided because it is taken for granted in the context. A particularly good example of this sort of communicative exchange is the textual extract in (80) which features three question-answer couplets in an exchange at the emergency room of a local hospital. The exchange is recounted by a narrator. The actors of the exchange are the narrator's son (S), who took the narrator to the hospital under a hypo-glycemic shock, and the doctor (D) that receives them at the entrance. Note that the equivalent of ‘who is it’ in (80a) in T-Oto is a predicative pronoun resulting from a cleft where the elements have fused phonologically to become one word. The same happens to the equivalent of ‘what is it’ in (80c).

(80) a. D –[ton]FP=kẹ=a=’a? ['toŋgjaʔa]
who.AS=COP=CL=3SGPRO
‘–¿Who is he?’

b. S –ke [mu̲ ʰta]FP– bi=’ñëm-bi
COP 1♂POSS father [3]CPL.R=SS/say.AS-3DAT
‘–It's my father– he said.’

c. D –[tẹʸ]FP=kẹ {rati kha di ʰta}CC?
what.AS=COP [3]INCPL.R do 2POSS father
‘–What is it that happens to your father?’

d. S –h<u>a</u>ndi, xi=ntyü
who.knows [3]PRF.R=TS/die
‘–Who knows; he's dead.’

e. D –’ne [to]$_{FP}$ {xi=ñü}$_{CC}$=’a?
and who [3]PRF.R=watch[3OBJ]=3SG$_{PRO}$
‘–And who is it that watched over him?’

f. S –kẹh=a [m<u>u</u> pariente=’mbe: Gabriel Mendoza]$_{FP}$
COP.AS=CL[3] 1♂POSS family.member=PL.EXCL GM
‘–It was our family member: Gabriel Mendoza.’

In (80) we have three direct questions: (a) and (e) have narrow focus as they ask for the identity of a participant while (c) has broad focus because the speaker asks what happened. In all three questions clefts are used. A truncated cleft is used in the responses in narrow focus.[45]

Up to now, we have seen that in T-Oto we have a cleft construction which can be manifested in different ways at surface level. These variants contrast as for the presence or absence of the copula, as for the presence or absence of the pronominal enclitic that cross-references the subject in focus of the copular clause, and as for the presence or absence of the focus extraction pronoun in the CC. In the next section, I introduce the clefts we observe in Northern Otomi where I show that while the same elements exist, the syntactic behavior of the focus extraction pronoun is rather different. Later on in §8, I explore the historical sources of clefts in Otomi in an attempt to understand their diachronic development.

7. Clefts in Northern Otomi.

In this section I present a very brief overview of the structures we find in Northern Otomi (henceforth N-Oto) to compare them with the situation presented for T-Oto. N-Oto is a large dialectal continuum spoken in various localities which are mainly located in the states of Querétaro and Hidalgo. It comprises two major dialectal areas: On the one hand, (i) the area of the state of Querétaro and surrounding areas that includes Santiago Mezquititlán Otomi (Hekking 1995); SIT-Oto (Palancar, 2009); Tolimán Otomi and Guanajuato Otomi (Lastra, 2001); and on the other hand, (ii) the Otomi of the Mezquital Valley, which is the variant that received the attention of earlier grammatical studies of Otomi (Ecker, 1952, Hess, 1968). All dialects of N-Oto have a very similar syntax and morphosyntax. I exemplify cleft constructions in N-Oto from SIT-Oto and from Mezquital Otomi (henceforth M-Oto).

While T-Oto retained an old V-initial word order as a conservative language, N-Oto has become an SVO language;[46] a change that has consequences for the syntax of clefts. The natural SVO order is illustrated in (81), which represents a sentence under broad focus.

(81) porke [no=r ’b<u>e</u>hño]$_{SUBJ}$ mí=pho’m-a≈[’na=r m<u>a</u>hwi]$_{OBJ}$
SIT-OTO because DEF.SG=SG woman [3]IMPF=place.cloth.on.top-BS≈one=SG shawl
‘Because the woman was putting a shawl over (it).’ (Txt)

But besides basic word order, there are other differences.[47] The most striking one is the fact that the focus extraction pronoun in N-Oto, which is *go* –cognate of T-Oto *ko*– is associated with the copula and not with the predicate of the CC. This is shown in (82a) if compared with the ungrammaticality of (82b), which would resemble the structure we find in T-Oto. Further note that the referent of the FP is inanimate, indicating that *go* is not sensitive to animacy.

(82) a. ’ne [nu=ya ’yothö$_i$]$_{FP}$ **go**$_i$ **ge** {ndí=ñ-’ëñ=he}$_{CC}$=’u$_i$
SIT-OTO and PRTCL=PL tender.maize.cob FOC COP 1.IMPF=M-play.AS=PL.EXCL=3PL$_{PRO}$
‘And it was the tender maize cobs that we played with.’ (Txt)

b. *’ne [nu=ya ’yothö$_i$]$_{FP}$ **ge** {**go**$_i$ ndí=ñ-’ëñ=he}$_{CC}$=’u$_i$
and PRTCL=PL tender.maize.cob COP FOC 1.IMPF=M-play.AS=PL.EXCL=3PL$_{PRO}$
Intended reading: ‘And it was the tender maize cobs that we played with.’

Like in T-Oto, *go* can be omitted, as in (83), where the referent of the FP has been first topicalized, as a way to introduce what is going to be said about it. We can see this because of the occurrence of the cross-reference pronoun on the copula. Example (83) further illustrates that the enclitic can also be omitted in N-Oto.[48]

(83) [yä ndämfri$_i$], **ge**=[’u$_i$]$_{FP}$ {ts’i-p-ä≈rä́ ngo}$_{CC}$
M-OTO PL cattle COP=3PL$_{PRO}$ [INCPL]PASS/eat-3DAT-BS[3S$_O$]≈SG.3POSS meat
‘As for the cattle, it's them that provide the meat to eat.’ (§4/p.39)
(Lit. ‘The cattle, THEY are the ones eaten for their meat.’)

While the language is SVO, the copula phrase in a cleft in N-Oto can also appear in initial position.[49]

(84) a. hont’ä go ge [yoho ya ngü$_i$]$_{FP}$ {mí=’buh=pu}$_{CC}$=’u$_i$
SIT-OTO only FOC COP two PL house [3]IMPF=exist.AS=there=3PL$_{PRO}$
‘It was only those two houses that were there.’ (Txt)

b. ge[=ra... nu=r xöt’ö$_i$]$_{FP}$ {xi=hyots’e}$_{CC}$=’ä$_i$
COP=SG DEF=SG prickly.pear [3]PRF=SS/cook=3SG$_{PRO}$
‘It was the prickly pear what they cooked.’ (Txt)

And like in T-Oto, the copula can also be omitted, as in (85). Note that example (85) does not have the same structure as example (86) in T-Oto, which is an instance of Construct (d). The resemblance is at surface level only, because the ungrammaticality of (82b) shows that *go* is not allowed in the CC in N-Oto.

(85) ’ne [nó=r ta^{h}ta no=r ntxuntsi=’ä$_i$]$_{FP}$
SIT-OTO and DEF.SG.3POSS=SG father DEF.SG=SG girl=DEM.SG
go$_i$ ___ {mí=khux-k=he}$_{CC}$=’ä$_i$
FOC COP [3]IMPF=wake.up-1OBJ=PL.EXCL=3SG$_{PRO}$
‘And it was the father of that girl the one who would wake us up.’ (Txt)

(86) [ñü=a$_i$]$_{FP}$ ___ {**ko**$_i$ ti=^{h}pe̲t'i}$_{CC}$
T-OTO PRTCL=3SG$_{PRO}$ COP FOC [3]CPL.IRR=make.tortillas
'It's her who makes the tortillas.' (Txt)

The same applies to (87), which again is different from (86), because it has the same configuration we observe in (89) where the copula is overt.

(87) [nu=ga$_i$]$_{FP}$ hin≈**go**$_i$ ___ {dä́=nu=se̲}$_{CC}$
M-OTO PRTCL=1SG$_{PRO}$ NEG≈FOC COP 1.CPL=see=alone
'It wasn't me who saw it.' (§1097/p.360)

(88) [ni=r Hwä$_i$]$_{FP}$ ___ {hin≈**ko**$_i$ ru̲=xa̲h-pu̲-te}$_{CC}$=ni$_i$
T-OTO DEM.SG=SG John COP NEG≈FOC [3]HAB=teach.AS-3DAT.AS-ANTIP=DEM.SG$_{PRO}$
'It's not John who's a teacher.' (Archivo de Lenguas Indígenas de México, 121)

(89) bi=bäädi ge hin≈**go**$_i$ **ge**='ä [rä́ ñ'owi]$_{FP}${='ä
M-OTO [3]CPL=SS/know COMP NEG≈FOC COP =3SG$_{PRO}$ SG.3POSS companion=REL.SG
mí='ñääni }$_{CC}$
[3]IMPF=move
'He knew that it wasn't his friend who was moving.' (§848/p.139)

In the same fashion, examples like (90), which are very common in N-Oto, appear to be structural counterparts of the minimal clefts in T-Oto as shown in (91), repeated from (9) above, when in reality they are not.

(90) a. go$_i$ ___ {bi=n-txa̲}$_{CC}$=**'u̲**$_i$
SIT-OTO FOC COP [3]CPL=M-pray.AS=3PL$_{PRO}$
'It's them who prayed.' (Txt)

b. go$_i$ ___ {ñüni xi=ñ-ho}$_{CC}$=**'ä**$_i$
FOC COP [3/INCPL]eat PRF=ST-be.good[3S$_O$]
'It's him who eats well.' (Txt)

c. go ___ {dí=tsü}$_{CC}$
FOC COP 1.INCPL=get.scared
'It's me who gets scared.' (Txt)

(91) ...{ko xu̲ grá=^{h}pödi}$_{CC}$
FOC indeed 2.INCPL.R=know
'...you're indeed the one who knows.' (Txt)

The fact that *go* is associated with the copula in N-Oto indicates that *go* is not a relative pronoun in N-Oto. While it could be simply treated as a focus marker, I propose to analyze it as a pronoun that stands for the focus subject phrase in the matrix clause of a cleft. Like in T-Oto, in the clefts of N-Oto the copula is commonly elided while *go* is retained. But because *go* is associated with the copula in N-Oto, its occurrence in a cleft can be interpreted as a token of the copula. This is why instances like (90) are far more

frequent in N-Oto than their surface counterparts in T-Oto, because the omission of the copula is facilitated precisely because *go* is present.[50]

Up to now, we have seen that T-Oto and N-Oto are two Otomi languages with very similar clefts but which also have an important difference concerning the syntactic position of the focus pronoun *ko/go*. In T-Oto clefts, the pronoun is always associated with the pre-V of the CC, whereas in N-Oto it is always associated with the pre-V of the copula. When the copula is elided, this gives rise to interesting instances that look identical on the surface. In the next section, I investigate how clefts were constructed in the historical documents of Otomi in order to sketch a diachronic path of how the modern situation is likely to have come about.

8. The diachrony of clefts in Otomi: A view from Old Otomi

The first records of Otomi are from the 16th century. An important one is the grammar titled *Arte de la Lengua Othomi* written by Pedro de Cárceres, a Spanish Franciscan Friar who lived and worked at the convent of the Holy Cross in the city of Querétaro. Cárceres' grammar was finished in 1580, and it is the only Otomi grammar from the period to have survived, although only through the edition by Nicolás León, who published it in 1907. Cárceres' grammar was the earliest document on Otomi we knew of until very recently. In 2015, the website ‹primeroslibros.org› made available a facsimile of an Otomi doctrine from 1576, curated at the Benson Latin American Collection of the University of Texas at Austin.[51] This Otomi doctrine was written by Melchor de Vargas, a Spanish Augustinian Friar who was prior of the convent of Actopan in the state of Hidalgo.

The Otomi described in Cárceres' grammar corresponds to a variety spoken in the large northern area under the influence of the political center of Jilotepec, an important Otomi principality in pre-Columbian times. As Actopan was historically under Jilotepec's influence, we can conclude that the two earliest historical sources of Otomi come from the same area, which is also the area where modern N-Oto is now spoken. Because of this, we can also assume that the Otomi language registered in these sources is the direct ancestor of N-Oto. However, the inflectional morphology of T-Oto more closely resembles the one found in these sources, although T-Oto is the southernmost language of the Otomi family, spoken near the border of the state of Morelos. To make sense of this, Palancar (2011; 2012) has proposed that the Otomi spoken in the 16th century, which we can call 'Old Otomi', was still a unified language at the time of the arrival of the Spanish in Mesoamerica, and should be thus considered the ancestor of all the modern languages. The conservatism of T-Oto is also a proof that N-Oto was innovative. Old Otomi was probably spoken until the turn of the 18th century, where its dialectal diversification gave rise to the different languages we observe nowadays. The two sources have some examples of clefts.[52]

The focus extraction pronoun we observe in Otomi languages is descended from Old Otomi *ko*, where it was only used for human referents. This can be seen in the question-answer couplet in (92), where the referent of *ko* in (93b) is the Viceroy in (92a). I take the response to be a truncated cleft like the ones in (80) in T-Oto. Cárceres' own words to explain the relation between these two structures are given in (93).

(92) a. *tongũ ko visoRey?* (Cár./p.55)
to ngũ k'o visorrey?
who be.like HON viceroy
'Who is like the viceroy?'

b. *congũ ko Don Francisco* (Cár./p.55)
ko$_i$ ngũ [k'o Don_Francsico$_i$]$_{FP}$
FOC.HUM be.like HON Mr_Francisco
'It's Mr. Francisco who's like him.'

(93) "To ask for the quality or the quantity of a person, they say: *tongũ,* e.g., *tongũ ko visoRey?,* i.e. 'who is like the viceroy?'. Answer: *congũ*, which is the correlative of *tongũ*, in the same way that *tantus* answers *quantus*, in such a way that saying *tongũ ko visorrey?* with the answer *congũ ko Don Francisco*, they are alike in appearance."[53]

While *ko* was used for humans, for all other non-human referents Old Otomi had the pronoun *ke*. This is seen in the couplet in (94), where *ke* cross-references a horse.[54]

(94) a. *tengũ kãniphani o don Francisco?* (Cár./p.55)
te ngũ k'ä ní phani 'o Don_Francsico?
what be.like DEM.SG DET.SG.3POSS horse HON Mr_Francisco
'What is Mr. Francisco's horse like?'

b. *quẽgũ kãniphani* (Cár./p.55)
ke$_i$ ngũ [k'ä ni phani$_i$]$_{FP}$
FOC.NONHUM be.like DEM.SG 2POSS horse
'It's like your horse.'

Other examples of clefts are also found in Cárceres although they are regrettably scarce. Examples involving human *ko* are given in (95), which in Cárceres are given again as possible answers to a question such as 'who made these?'.[55]

(95) a. *cogueque* (Cár./p.50)[56]
ko$_i$≈geh-[ke$_i$]$_{FP}$
FOC.HUM≈COP.AS-1[SG]$_{SO}$
'It's me.'

b. *coguecãgo cotocha* (Cár./p.50)
ko$_i$≈geh-[kä=ga$_i$]$_{FP}$ {**ko**$_i$ tó=kha}$_{CC}$
FOC.HUM≈COP.AS-1[SG]AS=1[SG] FOC.HUM 1CPL.R=do
'It's me who did it.'

Interestingly, example (97b) has *ko* twice in the sentence: once associated with the copula and another associated with the predicate of the CC like in T-Oto. This indicates that Old Otomi allowed correlative *ko* in both positions, which is something not possible in the modern languages. Further note that when *ko* (also *ke*, further below) is associated

with the copula, the copula undergoes voicing from *ke* to *ge* because they form a phonological word. This is further evidence of how the focus pronoun occurs in the pre-V.[57]

There are also examples like (96), which may instantiate a structure like in T-Oto represented as analysis (a), or a case-like structure in N-Oto represented as (b).

(96) *nũgãgo hnĩgotochago* (Cár./p.50)

a. [nü=gä=ga$_i$]$_{FP}$ ___ {hnï≈**go**$_i$ tó=kha}$_{CC}$=ga
PRTCL=1[SG] =1[SG] COP NEG≈FOC.HUM 1CPL.R=do[3OBJ]=1[SG]$_{PRO}$

b. [nü=gä=ga$_i$]$_{FP}$ hnï≈**go**$_i$ ___ {tó=kha}$_{CC}$=ga
PRTCL=1[SG]=1[SG] NEG≈FOC.HUM COP 1CPL.R=do[3OBJ]=1[SG]$_{PRO}$
'It's not me who did it.'
(Lit. 'I am the one who didn't do it.')

The non-human focus pronoun *ke* is used in the question-answer couplet in (97).

(97) a. *ntãkãqueguegue?* (Cár./p.56)
[^{n}'ta=k'ä]$_{FP}$ **ke**$_i$≈ge[(=ge)]$_{FP}$[58]
which=DEM.SG$_{PRO}$ FOC.NONHUM≈COP(=3S$_O$)
'Which one (non human)?'

b. *quegueã* (Cár./p.51)
ke$_i$≈ge=['ä]$_{FP}$
FOC.NONHUM≈COP=3SG$_{PRO}$
'It's that one.'

The question in (97a) is about a non-human referent. A question about a human referent would be (98) and as expected, it has *ko*.

(98) *ntãk'ocoguegue?* (Cár./p.55)
[^{n}'ta=k'o]$_{FP}$ **ko**$_i$≈ge[(=ge)]$_{FP}$
which=HON.SG$_{PRO}$ FOC.HUM≈COP(=3S$_O$)
'Which one (human)?'

The animacy distinction present in the correlative pair *ko/ke* in Old Otomi was lost in the modern languages, which use the reflexes of old *ko* for all types of referents regardless humanness or animacy.[59] The old marker *ke* got lost in T-Oto. It lived on in N-Oto in the form of *ge* where it became the general complementizer of the language.[60] This can be seen in example (99) where it introduces a complement clause, and in (100) where it introduces a relative clause, restrictive or non-restrictive.

(99) ya de_nkhapu da=phödi [**ge** ya bi=dü nu=ya txi-khö'i]
SIT-OTO ANT like.that IRR=PASS/know[3S$_O$] COMP ANT [3]CPL=SS/die DEF=PL DIM-people
'And like that it's known that the people died.' (Txt)

(100) a. nu ha dí=’bu=he... ’bu=[’na=r txi-ntxumpho
SIT-OTO REL where 1.INCPL=live.AS=PL.EXCL [3/INCPL]live.AS=one=SG DIM-non.indian.woman
{**ge** ndí=kum-p=he}]
REL 1.IMPF=grind.corn-3DAT=PL.EXCL
‘Where we live, also lives a woman for whom we grind corn.’ (Txt)

b. ntonse no=r txi-male=’ä,
then DEF.SG=SG DIM-old.woman=DEM.SG
{**ge** m=ár thühü Tasya ’nehe}, go ___ {bi=nangi}
REL [3]IMPF=SG.3POSS name Anastasia too FOC COP [3]CPL=get.up
‘Then that old woman, whose name was also Anastasia, was the one who got up.’ (Txt)

Even if *ge* was retained in SIT-Oto, it did not survive as a focus pronoun, and cannot be used in clefts, where speakers only use *go*. In M-Oto, however, we still find it sporadically as a free alternative of *go* when the referent in focus is inanimate. This can be seen in the cleft in (101), which is about a certain herb remedy that is sold in markets and that is kept in a glass bottle. But here, *ge* is interchangeable with *go*.

(101) nu’mu gä noonxi tsa≈dä hyand-ä≈n’a
M-OTO when PREP Monday MOD≈[3]IRR SS/see-BS≈one
ge nduunthi yä khä’i di=hñä yä xito,
COMP many PL people [3]IRR=SS/carry PL bottle
ge$_i$ ge[=’ä]$_{FP}$ {ri=maa da=hyoñ}$_{CC}$ =’ä
FOC.INAN COP=3SG$_{PRO}$ [3]IRR.AMBUL=SS/go [3]IRR=SS/look.for.AS=3SG$_{PRO}$
‘When on a Monday one may see that many people are carrying bottles; that's what they are looking for.’ (§640/p.109)

Finally, (102) is found in Vargas' doctrine and it is the only one of its kind. It shows an instance of what seems to be the structure we find in most clefts in T-Oto. The example comes from the first verse of the Lord's Prayer, which was rendered with a cleft. Given the amount of surface variation in the modern languages, it is highly probable that the same happened in the old language and that example (102) could also have been alternatively produced with a second *ko* associated with the copula like in example (95b) above. But I assume that such repetitions were not at all favored in the old dialect of the South, thus giving rise to the cleft we observe in T-Oto.

(102) *matacahœ que copibuinonamahetzi* (Var/f.8)
[o ma Ta^h=ka=hu$_i$]$_{FP}$ **ke** {**ko**$_i$ pi=’b^{w}ui nü=nu má-hëts’i}$_{CC}$
HON 1POSS father=1=PL.INCL COP FOC.HUM [3]CPL.R=live PRTCL=there LOC-above
‘Our Father, who art in heaven.’
(Lit. ‘Our venerable father is the one who is there above.’)

The example is a translation of Spanish *Padre Nuestro, que estás en los cielos*, which, like in English, contains a non-restricted relative clause. Even though the phrase makes sense in a Christian tradition, it may have sounded odd to native Otomi ears and was rendered with a cleft. The next version of the prayer is found two hundred years later in the Otomi Catechism by Francisco de Miranda dating from 1759. This work is

written in an old form of N-Oto; a language that we can call 'Colonial Otomi' and which is no longer the Old Otomi documented in the early sources. The first verse of the prayer is rendered again with a cleft and is given in (103). The cleft has no overt copula. The author placed a comma between the FP and the CC to make the Otomi look more like the Spanish version, perhaps led by the impression that the Otomi CC was a non-restricted relative clause.

(103) *MAtá ahe, goguibᵩi àmahêtzi* (Mir/f.1)

[ma	Ta=he$_i$]$_{FP}$	**go**$_i$	___	{gí='bu̲i	'a	mahëts'i}$_{CC}$
1POSS	father=PL.EXCL	FOC	COP	2.INCPL=live	P	heaven

'Our Father, who art in heaven.'

Example (103) could be a grammatical sentence in N-Oto, but a modern equivalent of (102) could not. If the copula were to be restored in modern N-Oto in an example such as (103), it would have to invariably occur after the focus pronoun *go* and not before it. This was shown in the ungrammaticality of (82b), above. We cannot know for sure if the encoding possibility in (102), which occurred two hundred years earlier, was still available in Colonial Otomi at the time it was registered, but it is unlikely. We also ignore at what precise moment N-Oto stopped having the possibility of retaining *ko* in the CC as we now observe in T-Oto, or when T-Oto lost *ko* in the copula. What we know for sure is that all cleft structures in later Colonial sources give a picture which is identical to the situation we have described for modern N-Oto, so the change must have taken place in the dialectal diversification. With this information in mind, we can reconstruct the situation for Old Otomi in Table 7, where all possibilities existed, but only some were taken up by the daughter languages.

INSERT TABLE 7

In the table, *ko*(≈COP) in (a) and *ke*(≈COP) in (c) indicate a situation where the focus pronoun is associated with the pre-V of the copula, which in turn can be elided. Similarly, *ko*-CC in (b) and *ke*-CC in (d) point to the structures where both *ko* and *ke* are resumptive relative pronouns in the CC. Table 7 shows that the syntax of clefts in the two modern languages is innovative, because both lost constructional material present in Old Otomi, although in different ways. None of the cases in (b) and (d) survived in N-Oto, and with *ke* lost in T-Oto, *ko*-CC was the only structure to remain. N-Oto favored (a) and (c), but as *go*, the reflex of human *ko*, took over all the referential space, it became insensitive to animacy, relegating non-human *ge* to its rare sporadic uses with inanimates in M-Oto; extinct in SIT-Oto. While *ge* lost terrain in its role as a focus marker with respect to *go*, it still flourished in the language because from the structure *ke*-CC, *ke* was reanalyzed as a complementizer in N-Oto, although how this may have happened falls outside the scope of the present paper.

The historical sources of Old Otomi have been invaluable in allowing us to reconstruct this situation. The very different results we obtain in both languages are likely to have emerged from the fact that most material in the original construction was facultative. The syntactic distribution of facultative material is likely to be difficult to learn and be passed on when there is so much optionality at usage level. What we observe is the result of choices taken to deal with the structural ambiguity of examples

like (96) above. We know less of the equivalent of Construct (a) in the sources, and because of the nature of the data (based on translations of religious texts), we know practically nothing about the pronominal enclitics to the left edge of the CC that we observe in the modern languages. However, because they are well attested in the family, I assume that they existed prior to the dialectal diversification. I have also assumed that it is only in T-Oto where these enclitics still reveal the alterations of the old word order in the copular construction from an underlying order COP-COMPL-SUBJ to an altered COP-SUBJ-COMPL. In N-Oto, however, where the order is SVO and where there is no longer a ban on DPs or pronouns based on the presentational particle *ñü* occurring inside the clause, the enclitics have lost their original syntactic motivation and should perhaps be better treated nowadays as a case of frozen morphosyntax associated with the copular construction.[61]

9. Conclusion

Clefts are compelling structures from all points of view. From a syntactic point of view they are biclausal, but they express a semantic proposition that could be otherwise expressed in a single clause. From the pragmatic point of view, they are structures dedicated to the encoding of contrastive focus. In this paper, I have shown that clefts in Otomi are based on the copular construction and that they consist of elements that can be omitted under certain circumstances. For example, focusing on Tilapa Otomi, I have proposed that the cleft construction may occur at surface level in at least eight different surface variants, depending on whether the copula, the focus phrase, the focus pronoun or the personal enclitic are overt or omitted. Furthermore, Otomi also uses truncated clefts (i.e., with no clefted clause) as a standardized way to frame the information in narrow focus in the answers to questions.

I have compared cleft constructions in Tilapa Otomi and Northern Otomi to obtain a comprehensive understanding of how clefts work in this family. The clefts in these two languages are structurally similar in many respects, and both have a focus pronoun (*ko/go*). But while in Tilapa Otomi, this pronoun is associated with the predicate of the clefted clause where it could still be analyzed as a relative pronoun, in Northern Otomi it is associated with the copula, and its function is to index the focus phrase in the matrix.

I have also explored the diachronic development of clefts in modern Otomi by studying the few clefts we find in the historical sources from the Colonial period of the New Spain. Based on the historical data, we can conclude that the syntax of clefts in both languages is innovative because both have lost grammatical material that was present in earlier stages.

The study of the syntax of focus constructions in the languages of Mesoamerica remains an unexplored terrain. While considerable advances have been made in the Mayan languages, much less is known about cleft constructions in other families. For example, no previous work existed on the Oto-Pamean languages. The present study provides a first analysis of clefts in Otomi, but it is also an invitation to fellow Mesoamericanists to contribute to building a body of knowledge on this fascinating area of the syntax of the indigenous languages of Mexico and Central America.

REFERENCES

AISSEN, JUDITH. 1992. Topic and focus in Mayan. Language 68:43–80.

_____. To appear. Information structure. In Judith Aissen, Nora England and Roberto Zavala (eds.), A Handbook of Mayan Languages. Oxford University Press.

BLACK, CHERYL A. 2000. Quiegolani Zapotec Syntax: A Principles and Parameters Account. Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

BOHNEMEYER, JÜRGEN. 2014. Non-identificational clefts in Yucatec Jürgen. Paper read at Semantics of Under-Represented Languages in the Americas 8. University of British Columbia.

BROADWELL, GEORGE AARON. 1999. Focus alignment and optimal order in Zapotec. MS.

CÁRCERES, FRAY PEDRO DE. 1580/1907. Arte de la Lengua Othomi. Edited by Nicolás León, Boletín del Instituto Bibliográfico Mexicano 6, 38–155.

DAVIS, HENRY; LISA MATTHEWSON AND S. SHANK. 2004. Clefts vs. Nominal Predicates in Two Salish Languages. In Donna Gerdts and L. Matthewson (eds.), Studies in Salish Linguistics in Honor of M. Dale Kinkade. Missoula, MT: University of Montana Press

DECLERCK, RENAAT. 1988. Studies on copular sentences, clefts and pseudo-clefts. Walter de Gruyter.

DELIN, JUDY. 1995. Presupposition and shared knowledge in it-clefts, Language and Cognitive Processes 10.2:97–120.

DEMIRDACHE, HAMIDA K. 1991. Resumptive chains in restrictive relatives, appositives and dislocation structures. Ph.D. dissertation, MIT.

_____. 1997. Dislocation, resumption and weakest crossover. In Elena Anagnostopoulou, Henk van Riemsdijk, and Frans Zwarts (eds.), Materials on Left Dislocation, pp. 193–231.Amsterdam: Benjamins.

DEN DIKKEN, MARCEL. 2005. Specificational copular sentences and pseudoclefts. The Blackwell companion to syntax, eds. Martin Everaert, Henk Van Riemsdijk, Rob Goedemans, and Bart Hollebrandse, pp. 292–409. Wiley-Blackwell.

_____. 2006. Relators and Linkers. The Syntax of Predication, Predicate Inversion, and Copulas. Cambridge, MA: MIT Press.

DRUBIG. HANS-BERNHARD. 2003. Toward a typology of focus and focus constructions Linguistics 41.1:1–50.

DRUBIG, HANS-BERNHARD AND W. SCHAFFER. 2001. Focus constructions. In Martin Haspelmath and Ekkehard Koenig and Wulf Osterreicher (eds.), Language typology and language universals. pp. 1079-1104. Berlin: Mouton de Gruyter.

ECKER, LAWRENCE. 1952. Compendio de Gramática Otomí (Introducción a un Diccionario Otomí-Español). Sobretiro del Anales del Instituto Nacional de Antropología e Historia IV, 32, (1949-1950). México DF: Instituto Nacional de Antropología e Historia.

GUTIÉRREZ BRAVO, RODRIGO. 2015. Las cláusulas relativas en maya yucateco. [Estudios de Lingüística XXII]. Mexico City: El Colegio de México.

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-GREEN, NÉSTOR. 2014. *Gramática del yühü (otomí de San Jerónimo Acazulco).* Mexico: CDI-Organización Civil Ndöngü.
_____. 2015. Morfosintaxis verbal del otomí de Acazulco. PhD Dissertation. CIESAS.
_____. 2016. Registration versus applicative constructions in Acazulco Otomi. International Journal of American Linguistics 82.3: 353–383
HESS, H. HARWOOD. 1968. The Syntactic Structure of Mezquital Otomi. [Janua Linguarum. Series Practica 43]. The Hague: Mouton.
HOLLENBACH, BARBARA E. 1995. A preliminary catalogue of focus devices in Mixtecan languages. Summer Institute of Linguistics Working Papers 11: 1–16.
JESPERSEN, OTTO. 1937. Analytic Syntax. London: Allen and Unwin.
_____. 1949. Modern English Grammar on Historical Principles. Part VII. Syntax. London: Allen and Unwin.
LAMBRECHT, KNUD. 1994. Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents. Cambridge Studies in Linguistics 71. Cambridge: Cambridge University Press.
_____. 2001. A Framework for the Analysis of Cleft Constructions. Linguistics 39.3: 463–516.
LEE, FELICIA A. 1997. Focus and Judgment Type in San Lucas Quiaviní Zapotec. Proceedings of the Twenty-3rd Annual Meeting of the Berkeley Linguistics Society [General Session and Parasession on Pragmatics and Grammatical Structure], 234–244.
LYON, JOHN. 2014. Predication and equation in Okanagan Salish. PhD Dissertation, University of British Columbia.
MIKKELSEN, LINE. 2005. Copular Clauses: Specification, Predication and Equation. Amsterdam: John Benjamins.
MIRANDA, FRANCISCO DE. 1759. Catecismo breve en lengua Otomí. Mexico City: Imprenta de la Bibliotheca Mexicana. Accessible at <bibliotecadigital.aecid.es/bibliodig/es>
MORO, ANDREA. 1997. The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure. Cambridge: Cambridge University Press.
OBERLANDER, JON AND JUDY DELIN. 1996. The function and interpretation of reverse wh-clefts in spoken discourse. Language and Speech 39.2–3:185–227.
PALANCAR, ENRIQUE L. 2009. Gramática y Textos del Hñöñhö: Otomí de San Ildefonso Tultepec, Querétaro. Volume I: Gramática. Mexico City: Plaza y Valdés.
_____. 2011. The conjugations of Colonial Otomi. Transactions of the Philosophical Society 109: 1–19.
_____. 2012. The conjugations classes of Tilapa Otomi: An approach from canonical typology. Linguistics 50.4: 783–832.
_____. 2013. The evolution of number in Otomi: The many faces of the dual. Studies in Language 37:1, 94–143.
PATTEN, AMANDA L. 2012. The English it-cleft: A constructional account and a diachronic investigation. [Topics in English Linguistics, 79]. Berlin: DeGruyter Mouton
QUIRK, RANDOLPH, SIDNEY GREENBAUM, GEOFFREY LEECH, AND JAN SVARTVIK. 1985. A Comprehensive Grammar of the English Language. London: Longman.

SALINAS PEDRAZA, JESÚS. 1984. Etnografía del Otomí. Mexico City: Instituto Nacional Indigenista.
SASSE, HANS-JÜRGEN. 1987. The thetic/categorical distinction revisited. Linguistics 25.2: 511–580.
SCHACHTER, PAUL. 1973. Focus and relativization. Language 49.1: 20–46.
TONHAUSER, JUDITH. 2003. F-constructions in Yucatec Maya. Jan Anderssen, Paula Menendez-Benito, and Adam Werle (eds.), The Proceedings of SULA, pp. 203–23. Amherst: GLSA, University of Massachusetts.
_____. 2007. Agent focus and voice in Yucatec Maya. Proceedings of the Thirty-ninth Meeting of the Chicago Linguistic Society, 39.1:540–58. Chicago: Chicago Linguistics Society.
TRECHSEL, FRANK R. 1993. Quiche focus constructions. Lingua 91:33–78.
VARGAS, MELCHOR DE. 1576. Doctrina cristiana muy útil y necesaria en Castellano, Mexicano y Otomi. Mexico City: Pedro Balli. In the Joaquin García Icazbalceta Collection, 1937 at the Benson Latin American Collection of the University of Texas at Austin. Accessible at <primeroslibros.org>
VERHOEVEN, ELISABETH, AND STAVROS SKOPETEAS. 2015. Licensing focus constructions in Yucatec Maya. International Journal of American Linguistics 81.1:1–40.
WARD, MICHAEL. 1987. A focus particle in Quioquitani Zapotec. Summer Institute of Linguistics Working Papers 9: 26–32
WEHR, B. 1984. Diskursstrategien im Romanischen. Tübingen: Narr.
ZIMMERMANN, MALTE. 2008. Contrastive focus and emphasis. Acta Linguistica Hungarica 55:347–60.

TABLES

‘dance’		**T-Oto** REALIS	**T-Oto** IRREALIS		**SIT-Oto** REALIS	**SIT-Oto** IRREALIS
INCPL	1st	tŕá në	gratu̲ në		dí në	
	2nd	grá në	gragu̲ në		gí në	
	3rd	ra në	tra në		në	
HAB	1st	tŕú̲ në	gru̲ në			
	2nd	grú̲ në	gru̲ në			
	3rd	(ru̲) në	tŕu̲ në			
CPL	1st	tú̲dú̲ në	gu̲tu̲ në		dá në	ga në
	2nd	gú̲gú̲ në	gugu̲ në		gá në	gi në
	3rd	bi në	ti në		bi në	da në
PRF	1st	xtú̲dú̲ në	xtigu̲tu̲ në		(x)tá në	(x)ka në
	2nd	xkú̲gú̲ në	xtigugu̲ në		(x)ká në	(x)ki në
	3rd	xpi në	xtigi në		xi në	(x)ta në

TABLE 1. INFLECTIONAL FORMATIVES IN OTOMI

		Singular	Dual		Plural
1st		=ga/ka	EXCL	__	=ga/ka=’mbe
			INCL	=ga/ka=wi	=ga/ka=hu̲
2nd		=k’e	__		=k’e=wi
3rd		=’a/’ä	__		=k’u̲
DEM	PROX	=na	__		=ya
	DISTAL I	=ni	__		=yu̲
	DISTAL II	(=k’a)	__		=k’i

TABLE 2. PRONOMINAL ENCLITICS IN TILAPA OTOMI

		T-Oto	SIT-Oto
xa	‘someone/anyone’	✓	__
to		__	✓
xo	‘something/anything’	✓	__
te		__	✓
hiN≈	NEG	✓	✓
hingi		✓	✓
hingu≈		✓	__
hin≈khoN≈	‘noone’	✓	__
hoN≈		✓	__
hin≈to		__	✓
(hi≈)ho	‘nothing’	✓	__
ho≈te		✓	__
hin≈te		__	✓
hin≈ts’e	‘no longer’	✓	__
nkhon≈ts’e	‘noone no more’	✓	__
haN≈	‘again’	✓	__
khaN≈	‘that way’	✓	__
paN≈	‘in any way’	✓	__
xa/x≈	‘already’ (ANT)	✓	__
xiN˜/xuN≈	‘also’	✓	__
xu	‘then’/‘indeed’	✓	__
xa	‘much’	__	✓

TABLE 3. PRE-V ELEMENTS IN OTOMI

		Omitted
VARIANT 1 & 2	COP [pro$_i$/DP$_i$] SUBJ [DP]COMPL(=pro$_i$)	*none*
VARIANT 3	COP [DP]COMPL(=pro$_i$)	[pro$_i$/DP$_i$] SUBJ
VARIANT 4	[pro$_i$/DP$_i$] SUBJ COP [DP]COMPL(=pro$_i$)	*none*
VARIANT 5	[pro$_i$/DP$_i$] SUBJ [DP]COMPL=pro$_i$	COP
VARIANT 6	[DP]COMPL=pro$_i$	COP & [pro$_i$/DP$_i$] SUBJ

TABLE 4. THE DIFFERENT SURFACE VARIANTS OF THE COPULAR CONSTRUCTION IN TILAPA OTOMI

						Variant	Example
Construct I				COP	[DP]COMPL [DP]SUBJ	__	__
Construct II			[DP$_i$]SUBJ	COP	[DP]COMPL =pro$_i$	__	__
Construct (a)		COP$_j$	[DP$_i$/pro$_i$/Ø$_i$]SUBJ	____$_j$	[DP]COMPL (=pro$_i$)	1, 2, 3	35, 37, 38
Construct (b)		Ø$_{COPj}$	[Ø$_i$]SUBJ	____$_j$	[DP]COMPL =pro$_i$	6	46
Construct (c)	[DP$_i$/pro$_i$] SUBJ	COP$_j$	____$_i$	____$_j$	[DP]COMPL (=pro$_i$)	4	48b
Construct (d)	[DP$_i$/pro$_i$] SUBJ	Ø$_{COPj}$	____$_i$	____$_j$	[DP]COMPL =pro$_i$	5	48a

TABLE 5. THE SYNTAX OF THE COPULAR CONSTRUCTION IN TILAPA OTOMI

						Ex.	Usage
Construct (a)		COP_j	$[DP_i/pro_i/Ø_i]_{FP}$	$\_\_\_\_j$	$[IP]_{CC}$	(=pro_i)	54, 8	Very common
					$[ko_i\ IP]_{CC}$	(=pro_i)	59, *viii*	Very common
Construct (b)		$Ø_{COPj}$	$[Ø_i]_{FP}$	$\_\_\_\_j$	$[IP]_{CC}$	=pro_i	*ix*	Extremely rare
					$[ko_i\ IP]_{CC}$	(=pro_i)	9, 78, 79	Very rare
Construct (c)	$[DP_i/pro_i]_{FP}$	COP_j	$\_\_\_\_i$	$\_\_\_\_j$	$[IP]_{CC}$	(=pro_i)	58a	Very common
					$[ko_i\ IP]_{CC}$	(=pro_i)	Unattested	Not natural
Construct (d)	$[DP_i/pro_i]_{FP}$	$Ø_{COPj}$	$\_\_\_\_i$	$\_\_\_\_j$	$[IP]_{CC}$	=pro_i	51, 58b	Very rare
					$[ko_i\ IP]_{CC}$	(=pro_i)	60, 73	Common with $[pro]_{FP}$
Construct (e)		COP_j	$[DP_i/pro_i]_{FP}$	$\_\_\_\_j$		(=pro_i)	80b/f	Common in responses

TABLE 6. THE SYNTAX OF CLEFTS IN TILAPA OTOMI

	Old Otomi		T-Oto	N-Oto SIT-Oto	M-Oto	Reanalyzed as a relativizer > complementizer
a.	*ko*(≈COP)	→	____	*go*(≈COP)		
b.	*ko*-CC	→	*ko*-CC	____		
c.	*ke*(≈COP)	→	____	____	*ge*≈COP	
d.	*ke*-CC	→	____	____		N-Oto

TABLE 7. DIACHRONIC DEVELOPMENT OF CLEFTS IN OTOMI

FIGURES

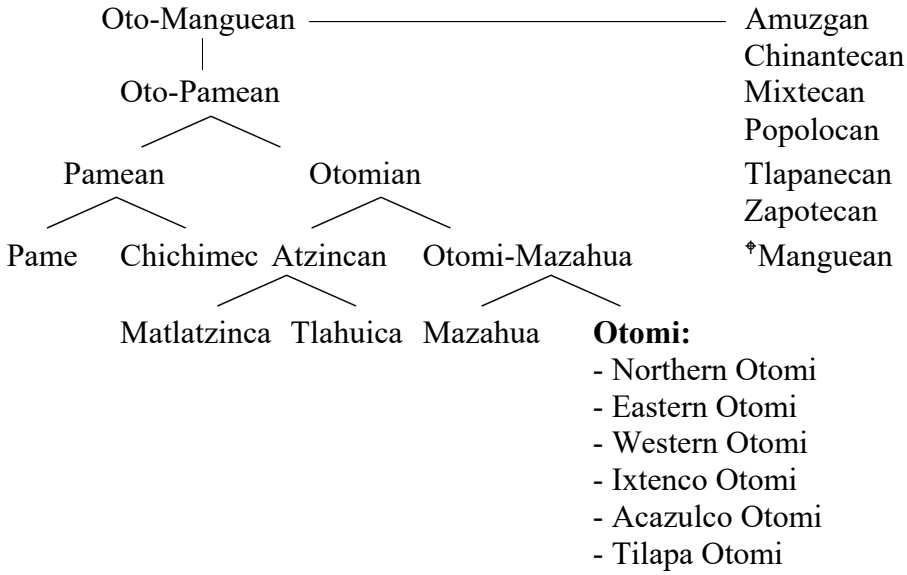

FIGURE 1. OTOMI WITHIN OTO-MANGUEAN

NOTES

[1] This paper was written under the auspices of the 2105-2017 CNRS International Project for Scientific Cooperation "Mesoamerica and the syntax of the relative clause". I want to thank Roberto Zavala, Claudine Chamoreau and Katharina Haude for their comments, and especially Judith Aissen for her teachings and suggestions. I also thank the IJAL anonymous reviewers for their time and wisdom. I am most heartily grateful to one of them, Jürgen Bohnemeyer (who identified himself in his review), for all his insightful comments and suggestions to improve the analysis so that one could get a much better picture of the puzzle. I am also indebted to Tim Feist for proof-reading the English in the manuscript and for spotting areas that needed further attention. All errors and deficiencies remain my own responsibility.

[2] The use of the word 'focus' in Jespersen is non-technical, but in more technical terms, focus is a pragmatic category, which is often hinged on the concept of topic, and is thus interpreted as the new knowledge about the topic. In this light, focus is very close or indistinguishable from comment, if comment is to be taken as the carrier of new information. Lambrecht (1994:207*ff*) is a pioneer in proposing to keep focus and topic independent from each other. His proposal is to treat focus as "the unpredictable or pragmatically non-recoverable element in an utterance". More specifically for Lambrecht, focus is "what makes an utterance into an assertion". In other words, focus is the element of information in the proposition expressed by a sentence that enables an assertion to be made with respect to the presupposition.

[3] The literature on the semantics of copular construction is vast (for a comprehensive review see Declerck, 1988; Mikkelsen, 2005; Den Dikken, 2005; Patten, 2012, among others) mainly because authors do not always agree on how to deal with the semantic subtleties that emerge in such constructions. Generally speaking, specificational constructions are ascriptive predications of identity where the identifying function involves revealing the identity of a given entity. In (5a) the lead actress is a role played by a specific person, whose identity is revealed as being the same person as Ingrid Bergman. In equative situations, it is stated that the referents of both entities is one and the same referent, while both are well identified individuals from the start. In (5b), we both know who *she* is and who *Ingrid Bergman* is; the sentence informs us that they happen to be the same person. In this article, I do not provide any semantic analysis of clefts in Otomi, but neither have I observed semantic differences like these at play in the distribution of cleft variants in Otomi.

[4] Patten (2012:68) rightly points out that the label 'pseudo' in pseudo-clefts (as opposed to IT-clefts, which would be the genuine clefts) originates in the fact that the structure may allow for instances which cannot be paraphrased into a simple clause (e.g. *what I like about it is that it's so sweet*). While this might be significant for some authors, I still consider examples in (4) as clefts in the same fashion as Ward *et al.* (2002:1423) or Lambrecht (2001).

[5] Lambrecht (2001:467) avoids using such terms precisely because he does not want to "address the vexing question of whether the focal constituent in such sentences is a subject or not."

[6] It is also known that clefts can be used as presentational devices (i.e. thetic structures) when the CC lacks presuppositional content (Declerck, 1988; Delin, 1995; den Dikken, 2005, Lambrecht, 1994; Oberlander and Delin, 1996, etc.). In such cases, clefts are used

as rhetorical tropes to present pieces of new information as if they were old. This is done to capture the desired attention of the addressee; attention being obtained as an effect of having to exercise pragmatic accommodation required to understand them. Used in this way, clefts may produce an effect of emphasis, which is provided by their inherent contrastive frame. This is a phenomenon that has been observed in French clefts such as in *j'ai la tête qui tourne* (Lit. 'It's my head that turns.') 'I'm dizzy' (from Wehr, 1984 *apud* Sasse, 1987: 538-539), where a contrastive focus interpretation is out of the question.

[7] ORTHOGRAPHY: Deviations from the IPA. Consonants: C' /Cˀ/ (ejective); *ʰC* (pre-aspirated); *f* /pʰ/ (but also as *ph*); *ġ* [ʒ]; *j* /kʰ/ (but also as *kh*); *'* /ʔ/; *r* /ɾ/; *x* /ʃ/; *tx* /ʧ/; *y* /j/; and allophones *ñ* [ɲ]; *tŕ* [ʈ͡ʂ]; *dŕ* [ɻ]; and *ndy* [nɟ]. Vowels: *a̲* /ɔ/ [ɒ]; *e̲* /ɛ/; *o̲* /ə/; *u̲* /ɨ/; *ẹ* [ẹ]; *ọ* [ọ]; and diaeresis (¨) /nasal/. High tone is represented by an acute accent in elements where it involves a grammatical contrast. ABBREVIATIONS: ♀: female speaker; ♂: male speaker; ADLAT: adlative; ADV: adverbial inflection; AMBUL: ambulative; ANT: anterior ('already'); ANTI: antipassive; AS: morphophonologically adjusted stem (also applied to affixes); BS: bound shape; CC: clefted clause; CISLOC: cislocative; CL: phrase-final clitic; CLF: classifier; COMP: complementizer; COP: copula; CPL: completive; DAT: dative; DEF: definite; DEL: delimitative ('only'); DEM: demonstrative; DIM: diminutive; DTR: ditransitive stem; DU: dual; EXCL: exclusive; FOC: focus; FP: focus phrase; HAB: habitual; HES: hesitative; HON: honorific; HUM: human; IMPF: imperfect; INCL: inclusive; INCPL: incompletive; INSTR: instrument; IRR: irrealis; LIG: ligature; M: middle; NEG: negative; NMLZ: nominalizer; N-Oto: Northern Otomi; NONHUM: non-human; M-Oto: Mezquital Otomi; OBJ: object; PASS: passive; PL: plural; POSS: possessive; P: preposition; PRF: perfect; PRO: pronominal; PROX: proximative; PRTCL: particle; PURP: purposive; QUOT: quotative; R: realis; REL: relativizer; SG: singular; ST: stative; SIT-Oto: San Ildefonso Tultepec Otomi; So: S of patientive predicate; SS: secondary stem; SUBJ: subject; TS: tertiary stem; T-Oto: Tilapa Otomi; TRANSLOC: translocative.

[8] This verb belongs to an inflectional class that only holds intransitive verbs denoting activities. Different markers used for the same categories may be found in the examples in this paper. When this happens, the verb in question belongs to a different class, whose paradigm is not shown here.

[9] For the purposes of this paper, I have treated Otomi inflectional formatives as operating like morphological atoms with no internal structure. For further details see Palancar (2012).

[10] From a morphological perspective, the status of inflectional formatives as words and the fact that the whole paradigm of a verb involves such formatives makes the inflection of Otomi entirely periphrastic.

[11] There are also locative pronominals indicating deixis (e.g. *=g/kwa* 'here'; *=ni* 'there'; *=k'u* 'yonder').

[12] In many other instances, however, it is more difficult to pinpoint the actual pragmatic nature of the emphatic force (see Zimmermann, 2008 for the relation between contrastive focus and emphasis). The context to interpret (14) is the following: The fence at the speaker's house had been damaged by a drunk driver who promised to show up the next day to raise it. Before (14), the speaker says: '*And we were waiting and waiting for him, and when does he show up? Nobody at all was showing up. It was already 10:30 am and there was nobody there, so my son says:*'

[13] The enclitic *=a* is most commonly used as a discourse particle associated with the predicate where it adds a native flavor to the Otomi discourse; delimiting the right edge of its phonological phrase. It may also occur at the right edge of a clause. Associated with conjunctions, it can be translated as 'now'.
[14] In example (18), the verb occurs in a liaison form associated with the first free word of the following phrase. I also indicate this juncture with the approximation sign '≈'.
[15] But also in the conjunction *ñü=mb^{w}u̲* 'when', which is based on the free word *mb^{w}u̲* 'when/if/then'.
[16] The verb *^{h}to^{h}karu̲* 'be someone's turn' is transitive. It is a loanword from Spanish *tocar* 'touch' from the calque of the impersonal expression "it touches someone" with the meaning 'it's someone's turn'.
[17] The particle *ñü* is not used with phrases headed by a possessive marker (44b) or with proper nouns (88).
[18] Most Otomi languages lack adjectives, but in the ones that have them like Acazulco Otomi (Hernández-Green, 2015), adjectives never occur as complements of a copula in predication.
[19] Age property concepts (e.g. 'old', 'young', etc.) and human propensity concepts (e.g. 'silly', 'clever', etc.) are encoded as nouns and treated as instances of class predication (i.e. predications such as 'be clever' or 'be old' are construed as 'be a clever person' and 'be an old person', etc.).
[20] For the irrealis, speakers use the same formatives for the completive realis of intransitive verbs.
[21] The native word for 'mole' in Otomi is the noun *^{h}ts'u̲bi*, but the person telling the story also uses the word *^{h}tusa*, an adapted loanword from Mexican Spanish *tusa*.
[22] Example (38) designates a specific folk character on a float in a parade, not the property of being mad.
[23] The example has a specificational reading. In the context, what was under contention was the identity of a specific woman who is the doctor of the village clinic.
[24] The copula can only be elided when it is a positive statement about a state of affairs that applies either always or at the time of the speech act situation. In other situations, the copula is required, like for example in the negative statement in (41a).
[25] There are also cases like example (i) –an instance of Variant 3– where the pronominal enclitic is identical to the determiner in the complement of the copula. Because of the evidence of cases like (48), and many others like them, I take such cases as simple coincidences.

i. kẹh=a (**ni$_{j}$** sku̲ ba̲htsy)$_{COMPL}$=a=**ni$_{i}$**, ñü!
COP.AS=CL[3$_{i}$] DEM.SG DIM child.AS=CL=DEM.SG$_{PRO}$ see
'That is the child, look!' (Txt)

[26] Copular complements can be categorized as objects, for example in case marking, *It's him* but not **it's he*, or Spanish *Juan es médico* 'John is a doctor' which can be pronominalized as *Juan lo es*.
[27] However even with an overt copula, the default still is to use the enclitic.

ii. porke [ñü ni rú̲ ngopho]$_{FP}$ **kẹh=a**
because PRTCL DEM.SG SG.3POSS brain COP.AS=CL[3]
{bi=kokhi}$_{CC}$=**ni$_{i}$**
[3]CPL.R=bleed=DEM.SG$_{PRO}$

'Because it was her brains that bled.' (Txt)

[28] The presentational determiners of Salish languages, which occur in initial position of clefts, are particularly intriguing in this respect, (see Lyon, 2014; Davis *et al.*, 2004; among others).

[29] But as pointed out in Note 4, there are exceptions, such as (iii).

iii. kẹh=a [{tátu$\underline{\text{u}}$ $^{\text{h}}$tyü='mbe=a}]$_{\text{SUBJ}}$
COP.AS=CL[3] 1.INCL.AMBUL carry[3OBJ]=PL.EXCL=CL
[{g$\underline{\text{u}}$=$^{\text{h}}$p$\underline{\text{a}}$='mbe}]$_{\text{COMPL}}$='a
1.CPL.IRR=sell[3OBJ]=PL.EXCL=3SG$_{\text{PRO}}$
'What we are carrying along is what we are going to sell.' (Txt)

[30] I am grateful to Jürgen Bohnemeyer for suggesting the label 'focus extraction pronoun' to me.

[31] In (61), the FP is not a DP. This suggests that a bare NP is also allowed as a FP in clefts.

[32] Furthermore, both relative clauses and CCs have a restricted V-initial word order in Otomi. This is shown by the ungrammaticality of the (b) examples in (iv) and in the cleft in (v).

iv.a. má=kha [nt'a kwentọ$_i$ {mádi mbehti-gi='mbe __OBJ$_i$
[3]IMPF=exist one story [3]IMPF.HAB SS/tell.DTR-1DAT=PL.EXCL
mi sk$\underline{\text{u}}$ $^{\text{h}}$tö$^{\text{h}}$tsu}]
1♀POSS DIM granny
'There was a story that my granny used to tell us.' (Txt)

b. *má=kha [nt'a kwentọ$_i$ {**mi sk$\underline{\text{u}}$ $^{\text{h}}$tö$^{\text{h}}$tsu** mádi
[3]IMPF=exist one story 1♀POSS DIM granny [3]IMPF.HAB
mbehti-gi='mbe __OBJ$_i$}]
SS/tell.DTR-1DAT=PL.EXCL
Intended reading: 'There was a story that my granny used to tell us.'

v.a. kẹh=a {taga$_i$ tsihme **i ti p$\underline{\text{e}}$phi** __INSTR$_i$}$_{\text{CC}}$
COP.AS=CL[3] [3]IRR.ADV have.a.meal PL NMLZ work
'It was that what the workers were going to eat with.' (Txt)

b. *kẹh=a {**i ti p$\underline{\text{e}}$phi** taga$_i$ tsihme __INSTR$_i$}$_{\text{CC}}$
COP.AS=CL[3] PL NMLZ work [3]IRR.ADV have.a.meal
Intended reading: 'It was that what the workers were going to eat with.'

[33] Instead of *giti*, the inflectional formative to realize {2.CPL.R} for the basic paradigm of first conjugation verbs like $^{\text{h}}$*peni* 'wash clothes' and *x$\underline{\text{u}}$t'i* 'wash dishes' would be *g$\underline{\text{ú}}$*.

[34] Here Lambrecht uses the term 'presupposed proposition' as an equivalent to common ground knowledge

[35] As mentioned, prosody plays a fundamental role in disambiguating the syntax of the two constructions. In the cleft, the FP is stressed.

[36] Except with a locative pronoun, as shown in (56c) above.This restriction is necessarily in connection with the analysis I propose that *ko* does not occur in initial position in a relative clause.

[37] But instances like (vi), which are based on Construct (c) (i.e., with no overt copula), are also attested.

vi. [ñü=ga$_i$]$_{\text{FP}}$ kẹ$^{\text{y}}$n=ga {ko$_i$ t$\underline{\text{ú}}$='ün$\underline{\text{u}}$}$_{\text{CC}}$=ga$_i$
PRTCL=1SG COP.AS=1[SG] FOC 1.CPL.R=give.to.3.AS[3OBJ]=1[SG]$_{\text{PRO}}$

'It's me who gave it to him.' (Txt)

[38] *Ko* occupies a slot in the pre-V zone that precedes the indefinite pronouns *xo* 'something' and *xa* 'someone' (given in Table 2). This is shown in the collocation in (vii).

vii. ntonse [ñü=k'e=^{g}wi=a$_i$]$_{FP}$ {**ko**$_i$ **xọ** gi=mön=^{g}wi=a}$_{CC}$
then PRTCL=2$_{PRO}$=PL=CL FOC something 2.CPL.IRR=say.AS=PL=CL
'Then it's you (PL) who need to say something.' (Txt)

[39] For an insightful analysis of the focus marker in Akan as a relative pronoun, see Drubig (2003).

[40] Example (viii) illustrates an inanimate FP.

viii. pwes kẹ [zunga-yo$_i$]$_{FP}$ {ko$_i$ bi='ëh=a}$_{CC}$
well COP o'clock-two FOC [3]CPL.R=come.AS=CL
'Well, it was at two o'clock that he came.' (Txt)

[41] Example (ix) is the only attestation of this type of option in my corpus.

ix. {tŕúti kha='mbe}$_{CC}$='a
1.HAB make=PL.EXCL=3SG$_{PRO}$
'It (a specific type of food) is what we make.' (Txt)

[42] Although fronted DPs are often marked with *ñü*, example (77) is a case where this does not happen.

[43] Possessed predicates are patientive and they inflect for TAM differently than other nominal predicates.

[44] Instead, what is common is to topicalize the referent to be in focus, and then say something about it while placing it in focus. A common example is (x), which I would treat as an instance of Construct (a) with *ko*.

x. [ñü=a=ya mi ba^{h}tsi$_i$]$_{TOP}$, kêh=a[=ya$_i$]$_{FP}$
PRTCL=CL=DEM.PL 1♀POSS child COP=CL=DEM.PL$_{PRO}$
{ko$_i$ xu handi ti=kha}$_{CC}$=ya$_i$
FOC so perhaps [3]CPL.IRR=do=DEM.PL$_{PRO}$
'As for my sons here, it's them that will perhaps have to do it.' (Txt)

[45] This is an old exchange pattern of Otomi. It is consistently found in the exchanges portrayed in the doctrines written in the oldest attested version of the language. This may be seen in the typical question-answer couplet in (xi) from Vargas (1576: f.11). Note the structural resemblance to (80).

xi. Priest *–daca quecomã?*
[^{n}'ta=k'ä]$_{FP}$ ke {kó=mä}$_{CC}$?
which=DEM.SG$_{PRO}$ COP 2.CPL.R=say
'What have you said?' (Lit. 'Which one (of the prayers) is what you said?')
Believer *–quekã Ave maria*
ke [k'ä Ave_Maria]$_{FP}$
COP DEM.SG Ave.Maria
'The Ave Maria' (Lit. 'It's the Ave Maria.')

[46] Whether this development can be understood as having arisen from contact with Spanish or whether it should be seen as an independent development is a question that I shall leave out of the present paper.

[47] Another important difference involves the CC. In contrast to T-Oto, N-Oto allows for other relative clause types besides the asyndetic one. This is shown in the first cleft of

(xii). The second cleft shows a very rare instance of a pseudo-cleft where the FP occurs in final position.

xii.	pe=ge	nu'mu̲	nguu	[rä	daada$_i$]$_{FP}$	{nu$_i$='ä	bi=du}$_{CC}$	
M-OTO	but=COMP	if	like	SG	father	PRTCL=3SG$_{PRO}$	[3]CPL=SS/die	
	nu'mu̲ go$_i$	___	{dä=njut='ä}$_{CC}$		[yä	bätsi	{bi=gohi}$_i$]$_{FP}$	
	then FOC	COP	[3]IRR=pay=3SG$_{PRO}$		PL	child	[3]CPL=SS/remain	

'But if it's the father who dies, then the one who pays are the children than remain.' (§538/p.106)

[48] Pronominal enclitics in N-Oto may work as demonstratives or occur in clefts associated with the FP.

[49] The order VSO is not at all common in N-Oto, which suggests that the clefts in (84) may be instances of another cleft construction more similar in type to an IT-cleft.

[50] Like in T-Oto, clefts with no copula and no focus pronoun are rare in natural discourse in N-Oto.

[51] I am grateful to George Aaron Broadwell for telling me about the existence of this document.

[52] But while Cárceres' grammar and Vargas' doctrine are based on the same language and the same dialect, Cárceres registers an Otomi which is a little more innovative. This can be seen in the phonology. For example, N-Oto voiced old non-preaspirated stops /p,t,k/ → /b,d,g/, e.g. T-Oto *pödi* 'SS/know', *tehe* 'water', *koho* 'four' vs. SIT-Oto *bödi* 'SS/know', *dehe* 'water', *goho* 'four'. Vargas' doctrine does not have a record of this change, but Cárceres' grammar registers voicing word-internally, especially affecting elements in the pre-V zone, e.g. *to**b**ixohnãbãte* 'I teach elsewhere' [tʊ́≈**b**í xohnabate {1.INCPL.R≈DIST teach}] *vs.* ***p**ixohnãbãte* 'he teaches elsewhere' [**p**í=xohnabate {[3/INCPL.R]DIST=teach}].

[53] "Para preguntar por alguna persona, s. de su calidad o cantidad. dizen, *Tongũ*, vt. *Tongũ ko visoRey*, i. que persona tiene El visorrey. RP, *congũ*, que es correllatiuo, de *tongũ*, asi como a*tantus*. Responde quantus, de manera que diziendo. *Tongũ ko visorrey*. RP, *congũ ko Don Francisco*. s. son semejantes en la persona." (p.55)

[54] "*Tengũ*, este vsan para preguntar por todo animado y no animado excepto personas. Su corresponsivo es *quengũ* Exº *Tengũ kãniphani o don Francisco quengũ*. RP, *quẽgũ kãniphani. i. asi como el tuyo.'*. "*Tengũ*, this one they use it for all animates and inanimates except people. Its corresponsive is *quengũ*, e.g. *Tengũ kãniphani o don Francisco?* Answer: *quẽgũ kãniphani*, i.e. 'like yours'." (p.55).

[55] "Exº si me preguntan quien hizo Estos rrespondo; cogueque, vel. coguecãgo cotocha, I. yo lo hize, pero si rrespondo dubitatiue denotando negacion dire, nũgãgo hnĩgotochago, Como quien dize no lo hize yo." 'E.g. If I'm being asked 'who made those?', I answer: *cogueque* or *coguecãgo cotocha*, i.e. 'I did it', but if I answer with hesitation implying negation, I'd say: *nũgãgo hnĩgotochago*, as someone who says 'it's not me who did it'." (p.50)

[56] This truncated cleft shows that the copula is a patientive predicate. This property has been preserved in the modern languages, as shown in (xii.a) in N-Oto. In T-Oto, it's more difficult to see it because the copula commonly receives just pronominal enclitics unmarked for case, as in (xiii.b), but instances like (xiii.c) reveal its true nature as a patientive predicate.

xiii.a. go geh=ka=**gi**
SIT-OTO FOC COP.AS=1[SG]$_{PRO}$=1[SG]S$_{oPRO}$
'It's me.' (Txt)
b. kẹ^{y}n=**ga**
T-OTO COP.AS=1[SG]$_{PRO}$
'It's me' (Txt)
c. ken-**gu**=ts'e=ga='mbe...
COP.AS-1S$_{o}$.AS=DEL=1$_{PRO}$=PL.EXCL
'It's just us...' (Txt)

[57] Further note how *ko* is registered as *go* in (96) when hosted on the negation marker.

[58] Example (98) has the facultative pronominal enclitic *=ke* [ge] of 3rd person. The pronoun is insensitive to animacy. It was only used with inactive predicates, such as the copula. It has no reflexes in the modern languages.

[59] This is especially the case in N-Oto where *go* is used for plants (83a) or for houses (84a); a pattern which is not that common in T-Oto.

[60] This is an innovation. Old Otomi did not have an (overt) complementizer; a situation that is still observed in T-Oto. The phonological similarity between Spanish complementizer *que* /ke/ and Otomi *ge* is accidental. It may well have been the case that the functional range of *ge* in N-Oto has been enlarged by the influence of Spanish syntax, but *ge* had fully developed functions as a complementizer by the 18th century. I also discard the possibility that *ge* is a loanword from Spanish *que* /ke/. To get modern /ge/, Otomi speakers would have had to borrow Spanish *que* /ke/ before the voicing sound change took place. The general voicing of onset consonants took place in the first half of the 17th century, and that is too early a date for syntactic borrowings from Spanish because bilingualism was scarce or non-existent at that time.

[61] The loss of syntactic motivation may be behind their recruitment in N-Oto as demonstratives, which is a function not found in T-Oto; examples in (85) and (100b) above.