

On variation in descriptions of lesser-known languages:

A case study of verbal enclitics and
auxiliary verbs in Shuhi (Tibeto-Burman)

Katia Chirkova

CNRS-CRLAO

katia.chirkova@gmail.com

2022 年博雅语言学论坛

5 November 2022

- “Language Variation and Change”
 - variationist sociolinguistics
 - the cognitive and social factors in language variation
 - synchronic variation as a reflection of diachronic change
- variation in descriptions
- extensive differences in description of lesser-known languages
- methodology

- The luxury of methodology first becomes a necessity when continued investigations produce several competing theories, and we have to find out which one is right. [...] the highest goal of linguistics *the possibility of being right*.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1.1: 97-120. P. 98.

- principles that govern the gathering of empirical data

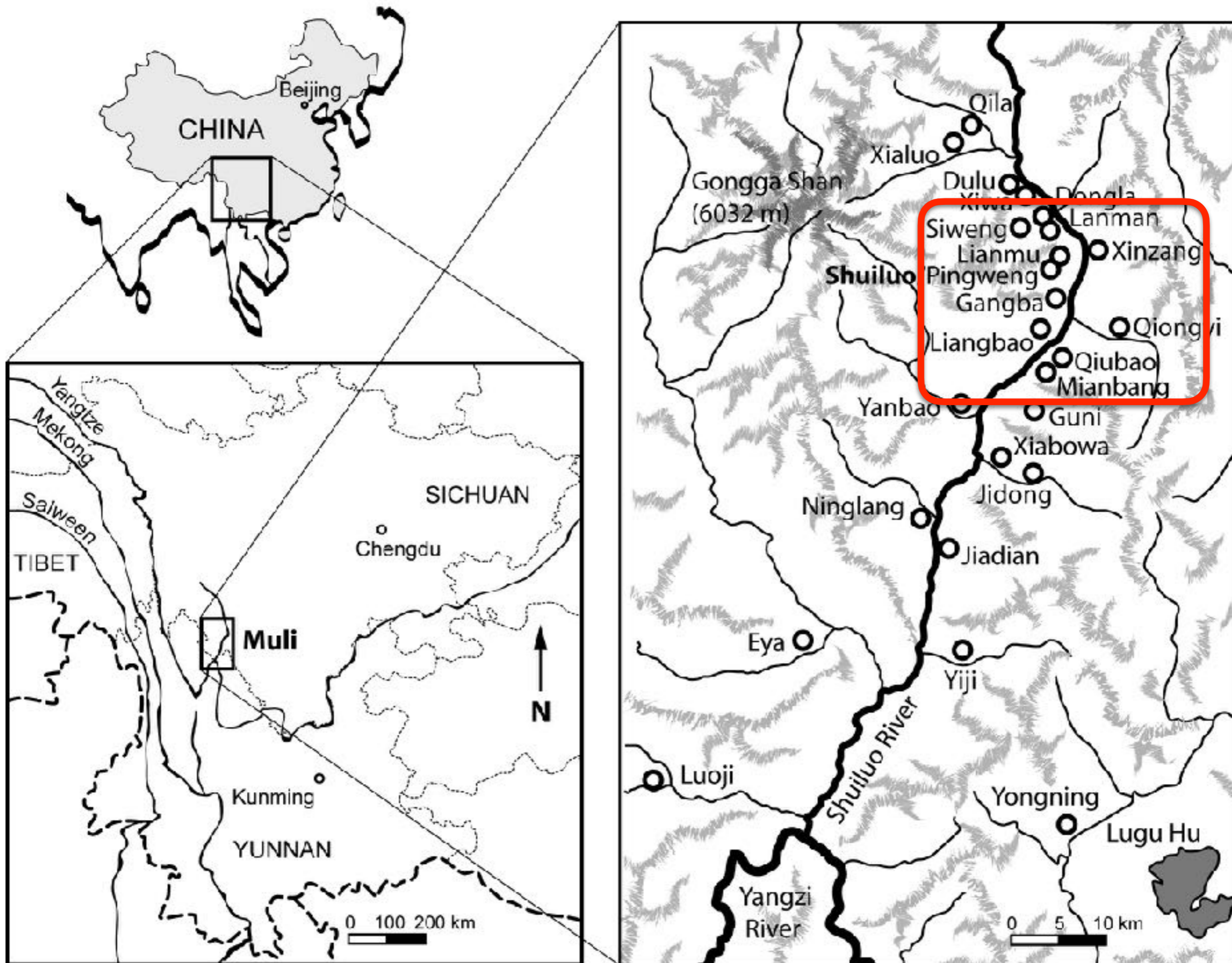
A case study

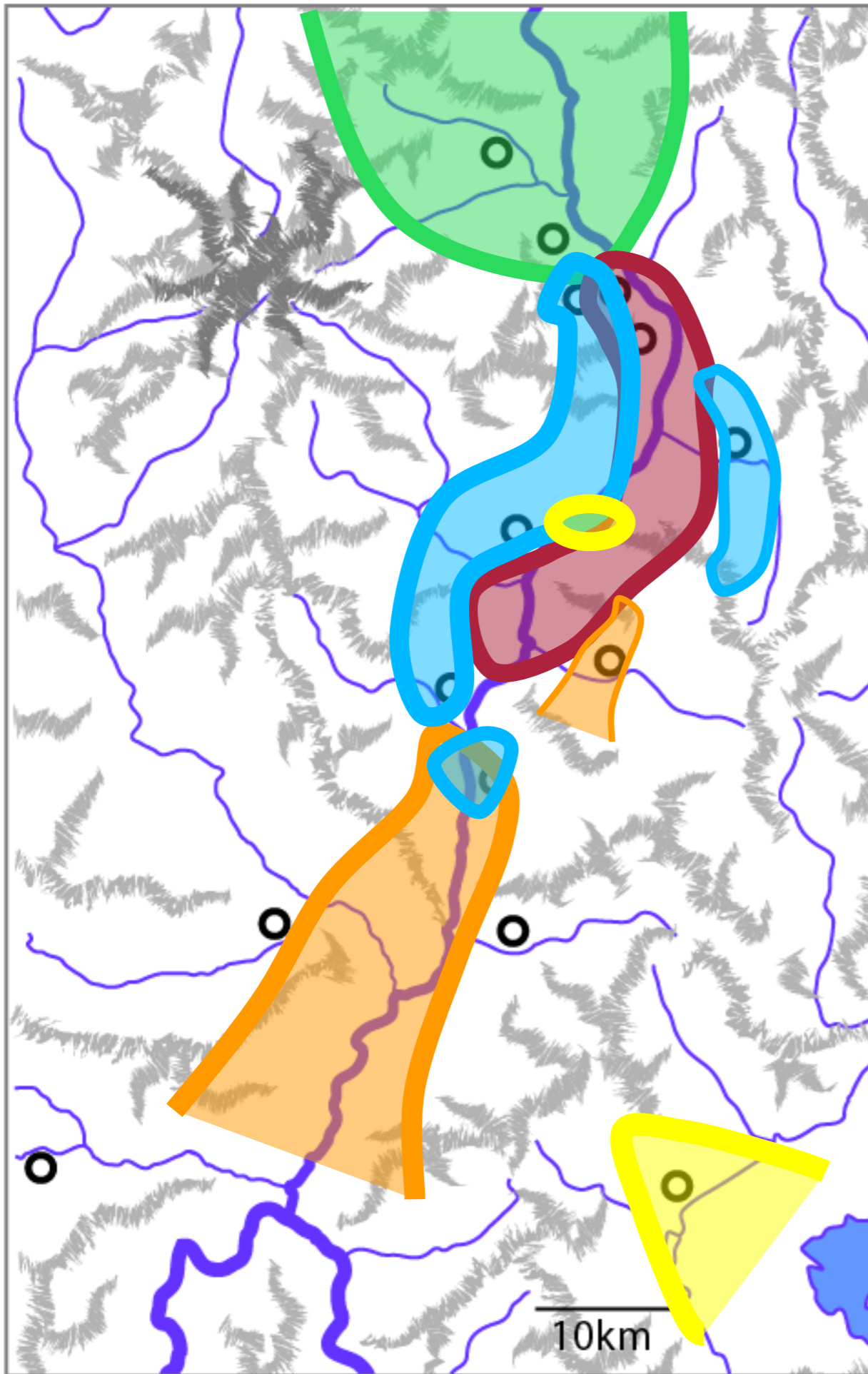
$\int u^{55} \tilde{h} i^{55} k e^{55} t \zeta 3^{53}$ 书亨语

Xumi 旭咪

Shixing 史兴

Shuhi, ISO 639-3 *sxg*





■ Kami Tibetans

■ Shuhi

■ Pumi

■ Naxi

■ Mosuo

Previous Studies

- Sun Hongkai 孙宏开 (1983)
- Xu Dan 徐丹 (2009)
- Sun Hongkai 孙宏开, Xu Dan 徐丹, Liu Guangkun 刘光坤, and Lurong Duoding 鲁绒多丁 (2014) *Shixingyu Yanjiu* 《史兴语研究》 [A Grammar of Shixing]
- Huang Bufan 黄布凡 & Renzeng Wangmu 仁增旺姆 (1991)
- Chirkova (2009, 2017) A grammar sketch

Verb template

Directional prefix /
Perfective prefix

Delimitative prefix
dʒi- ‘one’

Verb stem

Enclitic /
auxiliary
verb

Reported evidential
*pɛ*⁵³ ‘speak’ +
aspect marker

Prefix	Meaning	Example
<i>dʒi-</i>	upward	/dʒi ³³ -tɕ ^h ɛ ³³ / ‘lift up’
<i>mjæ-</i>	downward	/mjæ ³³ -ʒɛ ⁵³ / ‘put down’
<i>k^hu-</i>	inward	/k ^h u ⁵⁵ -lɛ ⁵⁵ / ‘come in’
<i>bə-</i>	outward	/bə ⁵⁵ -lɛ ⁵⁵ / ‘come out’
<i>lɛ-</i>	PFV	/lɛ ³³ -ʁo ³⁵ / ‘have broken’ (intr.)

*xɛ*³³ = *no*⁵⁵

*bə*³³-*dʒi*³³-*ts*^h*ʃ*⁵⁵

*dʒ*³³ = *ʁʃ*⁵⁵

*pʒ*⁵⁵ = *gje*⁵³

this=inside

outward-one-dance

do=IRR

speak=PROG.?

‘Would you agree to dance?’

Verbal enclitics and auxiliary verbs

Category	Sun (1983), Sun et al. (2014)		Huang & Renzeng (1991)		Chirkova (2017)
	1 or 2 person	3 person			
Prospective (将行体)	we ⁵⁵	we ⁵⁵ dzɿ ³⁵	gɜ ⁵⁵		gɜ, ɿẽ
Progressive (进行体)	ji ⁵⁵		ji ⁵⁵		ji (dynamic verbs)
Perfective (已行体)	tɛɛ ⁵³ sɿ ³³ , sɿ ³³	tɛyɛ ⁵³	sɿ ^{33/55}		lɛ-
Direct (亲见体/亲验语气)	hɛ ⁵⁵ dzan ³¹ , dzan ³¹ , ɛan ³¹		process	result	dzɿ̃ (stative verbs)
Durative			tɛæ ³³	(wu ³³)dzɿ̃ ³³	
Perfect (完成体)	ts ^h a ⁵⁵ ha ³⁵ sɿ ³³		ts ^h a ³³ wu ³³ sɿ ³³		sɿ
Terminative or telic					ts ^h ɛ
Experiential (经验体)	dze ⁵³		dzɜ ⁵³		dzɜ

Table 1. Verbal enclitics and auxiliary verbs in Shuhi, as presented in Sun (1983) and Sun et al. (2014), Huang & Renzeng (1991), and Chirkova (2017)

Verbal enclitics and auxiliary verbs

Category	Sun (1983), Sun et al. (2014)		Huang & Renzeng (1991)	Chirkova (2017)
	1 or 2 person	3 person		
Prospective (将行体)	we ⁵⁵	we ⁵⁵ dzɿ ³⁵	gɜ ⁵⁵	gɜ, ɿẽ
Progressive (进行体)	ji ⁵⁵		ji ⁵⁵	ji (dynamic verbs)
Perfective (已行体)	tɕɛ ⁵³ sɿ ³³ , sɿ ³³	tɕɛyɛ ⁵³	sɿ ^{33/55}	lɛ-
Direct (亲见体/亲验语气)	hɛ ⁵⁵ dzan ³¹ , dzan ³¹ , ɕan ³¹		process	result
Durative			tɕæ ³³	(wu ³³)dzõ ³³
Perfect (完成体)	ts ^h a ⁵⁵ ha ³⁵ sɿ ³³		ts ^h a ³³ wu ³³ sɿ ³³	
Terminative or telic				
Experiential (经验体)	dze ⁵³		dzɜ ⁵³	dzɜ

Table 1. Verbal enclitics and auxiliary verbs in Shuhi, as presented in Sun (1983) and Sun et al. (2014), Huang & Renzeng (1991), and Chirkova (2017)

- different markers
- different analysis of the same markers

Reasons for differences

(i) differences in the number of speakers: individual idiosyncrasies

(ii) differences between the sub-varieties of Shuhi, possibly reflecting contact influence from different languages:

- Sun (1983) and Sun et al. (2014): the lower and middle reaches
- Huang & Renzeng (1991) and Chirkova (2009, 2017): the upper reaches of Shuiluo River

(iii) differences in speech genres

- Sun (1983), Sun et al. (2014), Huang & Renzeng (1991): greater dependence on translation / elicitation
- Chirkova (2009, 2017): greater dependence on folk stories

New corpus

- an audio and video corpus of 73 hours (of which over 18 hours are annotated) (80 interlinear texts)
- over twenty speakers from the three main sub-varieties of the language of the upper, middle, and lower reaches of Shuiluo River
 - a few large-scale studies with over 100 speakers
- various genres (translated texts, traditional stories, personal narratives, procedural texts, songs, elicitation with non-verbal stimuli, ritual speech events, proverbs)
- various methods (elicitation, experiments, observation)

XUMI

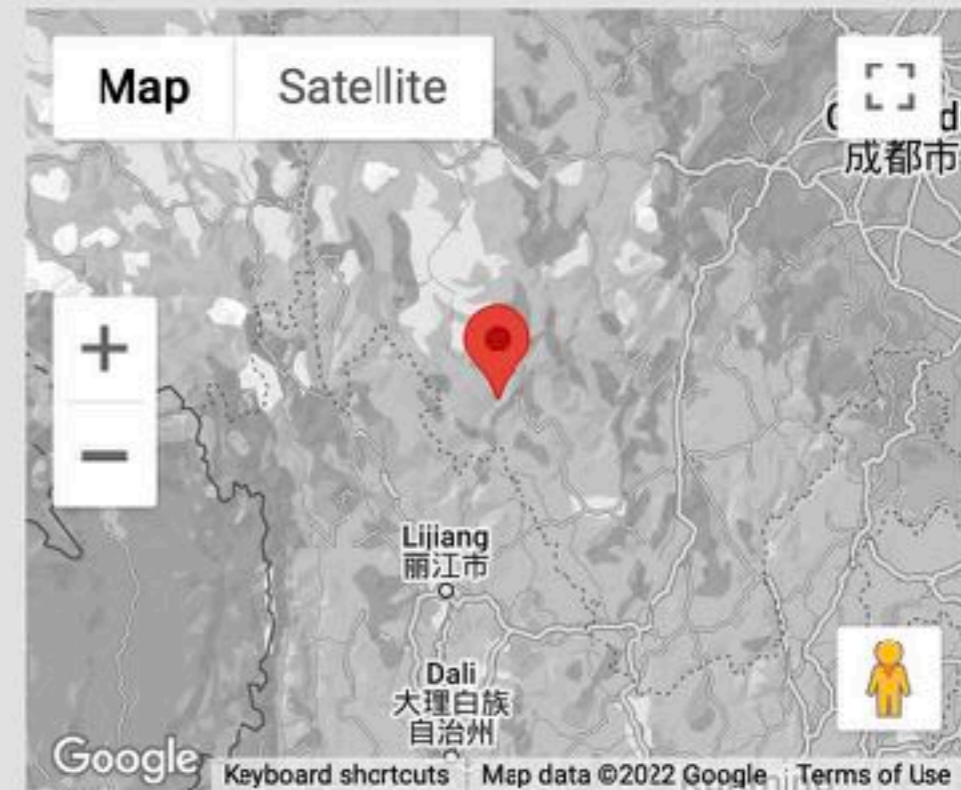
Depositor: Ekaterina Chirkova

Collection: Xumi: Documentation of a Highly Endangered Language of South-West China

The Xumi or Shuhi language (a.k.a. Shixing, ISO 639-3 sxx) is a little studied Tibeto-Burman language, comprising two sub-varieties with restricted mutual intelligibility (Upper Xumi and Lower Xumi). This deposit is based on Xumi data collected by the research team of the project "Ersu and Xumi: Comparative and Cross-Varietal Documentation of Highly Endangered Languages of South-West China" (MDP0257, 2013-2017), funded by the Endangered Languages Documentation Programme (ELDP) of SOAS, University of London. The collection comprises audio and video recordings of personal narratives, traditional stories, traditional songs, conversations, elicitations from Mandarin Chinese (using both written and non-written stimuli), and translations from Mandarin Chinese collected in the Lizu-speaking areas in Southwest China. Some recordings are accompanied by annotations including the following layers of analysis:

Access Collection here:

Link: <http://hdl.handle.net/2196/b115c458-4687-4183-8313-b134b86f41f8>



Verbal enclitics and auxiliary verbs

- the corpus encompasses most markers reported in previous studies
- the corpus contains additional markers, not mentioned in previous work (prospective events: *bo*, *gwei*, *ku*, etc.)

Category	Sun (1983), Sun et al. (2014)		Huang & Renzeng (1991)	Chirkova (2017)
	1 or 2 person	3 person		
Prospective (将行体)	we ⁵⁵	wɛ ⁵⁵ dzɿ ³⁵	gɜ ⁵⁵	gɜ, ɿẽ
Progressive (进行体)	ji ⁵⁵		ji ⁵⁵	ji (dynamic verbs)
Perfective (已行体)	tɕɛ ⁵³ sɿ ³³ , sɿ ³³	tɕyɛ ⁵³	sɿ ^{33/55}	lɛ-
Direct (亲见体/亲验语气)	hɛ ⁵³ dzan ³¹ , dzan ³¹ , ɕan ³¹		process	result
Durative			tɕæ ³³	wu ³³ lzõ ³³
Perfect (完成体)	ts ^h a ⁵³ ha ³⁵ sɿ ³³	ts ^h a ³³ wu ³³ ɿ ³³		sɿ
Terminative or telic				ts ^h ɛ
Experiential (经验体)	dzɛ ⁵³		dzɜ ⁵³	dzɜ

Table 1. Verbal enclitics and auxiliary verbs in Shuhi, as presented in Sun (1983) and Sun et al. (2014), Huang & Renzeng (1991), and Chirkova (2017)

(i) differences in the number of speakers: individual idiosyncrasies

(ii) differences between the sub-varieties of Shuhi

(iii) differences in speech genres

- distribution in different genres (tɕyɛ is restricted to elicitation with non-verbal stimuli)



zã³⁵ dzĩ tje⁵⁵ k^hu⁵⁵-lje⁵⁵ se sɛ³³ɹɛ⁵⁵ la mje³³-zũ⁵⁵ tɕye
son one there inward-come CONJ chair LOC downward-sit

[The speaker watches a video clip describing what she sees:] ‘A man has come in and sat on the chair.’ (Staged events, 079ET_chairsit)



zā³³ dʒī³³ mɯ⁵⁵zæ⁵⁵ dʒī³³ ke³³ku⁵⁵ bɛ sei
 son one girl one sit.in.lotus.posture make LNK

dʒā³³hō³³ nɔ⁵⁵ mje³³-zū⁵⁵ dʒā [...]
 house inside downward-sit

‘A man and a woman are sitting in lotus posture.’(001ET_handtohand)



mɯ⁵⁵ zæ⁵⁵ dʒĩ³³ ɿ³³ tē⁵⁵ ʧē⁵⁵ nɔ⁵⁵ bə³³-tsa⁵⁵ tɕyɐ

girl one TOP window inside outward-jump

‘A girl climbed out of the window.’ (003M_inoutin)

ŋɐ⁵⁵ tē⁵⁵ ʧē⁵⁵ nɔ⁵⁵ bə³³-dʒĩ³³ sə⁵⁵

1SG window inside outward-exist

‘I climbed out of the window.’



... kə⁵⁵də⁵⁵ɹə⁵⁵ tje⁵⁵ mje³³-zɿe⁵⁵ wu sei
tray there downward-drop RES LNK

‘...the tray fell down and broke.’ (005P_collidebreak)

lje³³-ko³⁵ xɛ = sɔ
PFV-break.INTR

Summary

- different sources of information about the reported events
 - tɕyɐ: observe the event
 - dzã: observe the result
- sə (tɕɛ = sə): person-related differences (1st person)
 - xɛ = sə: the speaker's access to information
- the lexical semantics of the verb (controllable and non-controllable; internal and observable state verbs)

Controllable verbs, e.g. dzue³⁵ 'hit, beat, strike'

	1st	2nd	3rd
sə	dzue ³³ = sə ⁵³		
tɕye		dzue ³³ = tɕye ⁵³	
tɕe	dzue ³³ = tɕe ³³ = sə ⁵³	dzue ³³ = tɕe ³³ = dzã ⁵³	
dzã	--	dzue ³³ = dzã ⁵³	

Non-controllable verbs, e.g. ɲe³³bɰ⁵⁵ 'laugh', le⁵⁵yɯ⁵⁵ 'rejoice'

ɲe⁵⁵ / t^hi⁵⁵ ɲe³³bɰ⁵⁵ = xɛ = sə
 1SG / 3SG laugh=go.PST=PRF
 'I / he laughed.'

t^hi⁵⁵-ɲe³¹ le⁵⁵-yɯ⁵⁵ = xɛ = dzã pe⁵⁵ = i = dzã
 that-P PFV-be.happy=go.PST=exist speak=PROG=exist
 'They were happy.'

Overview

- a complex interplay between the lexical semantics of the verb (controllable and non-controllable, internal and observable state verbs), aspect, and the egophoricity-evidentiality parameter
- also common in the neighboring languages (e.g. Sun 1993; Lidz 2010; Tournadre 2011; Tournadre & LaPolla 2014; Daudey 2014)

Verbal enclitics and auxiliary verbs

Category	Sun (1983), Sun et al. (2014)		Huang & Renzeng (1991)		Chirkova (2017)
	1 or 2 person	3 person			
Prospective (将行体)	we ⁵⁵	we ⁵⁵ dzɿ ³⁵	gɜ ⁵⁵		gɜ, ɿẽ
Progressive (进行体)	ji ⁵⁵		ji ⁵⁵		ji (dynamic verbs)
Perfective (已行体)	tɛɛ ⁵³ sɿ ³³ , sɿ ³³	tɛyɛ ⁵³	sɿ ^{33/55}		lɛ-
Direct (亲见体/亲验语气)	hɛ ⁵⁵ dzan ³¹ , dzan ³¹ , ɛan ³¹		process	result	dzõ (stative verbs)
Durative			tɛæ ³³	(wu ³³)dzõ ³³	
Perfect (完成体)	ts ^h a ⁵⁵ ha ³⁵ sɿ ³³		ts ^h a ³³ wu ³³ sɿ ³³		sɿ
Terminative or telic					ts ^h ɛ
Experiential (经验体)	dze ⁵³		dzɜ ⁵³		dzɜ

Category	Marker
Aspect	
Prospective	gɜ, ɿẽ
Progressive	ji ⁵⁵
Experiential	dzɜ
Perfective	directional prefix
Evidentiality / egophoricity	
Direct (Visual)	tɛyɛ
Indirect	dzõ ³³
Egophoric / non-egophoric	sɿ / xɛ

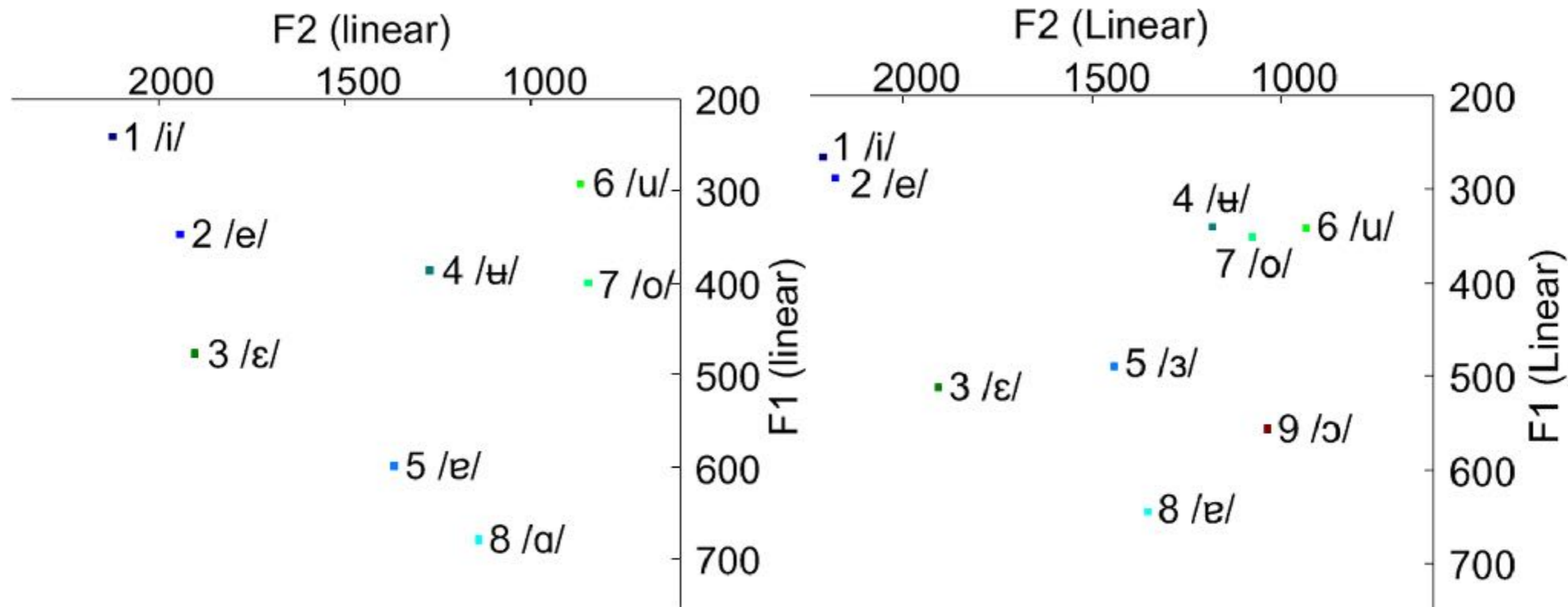
Summary

- reasons for differences in previous studies
 - differences in the size and the composition of the corpus underlying the analysis
- which one is right?
 - no right or wrong, but rather more or less complete descriptions

Desideratum

- in publications, to provide a detailed description of the set of data used for analysis and the methods of data collection
- principles of gathering of empirical data
 - diversity of genres and methods
 - need to combine elicited and non-elicited data, non-verbal stimuli (of particular importance when long stay in the area where the language is spoken is not possible)

Studies of variation

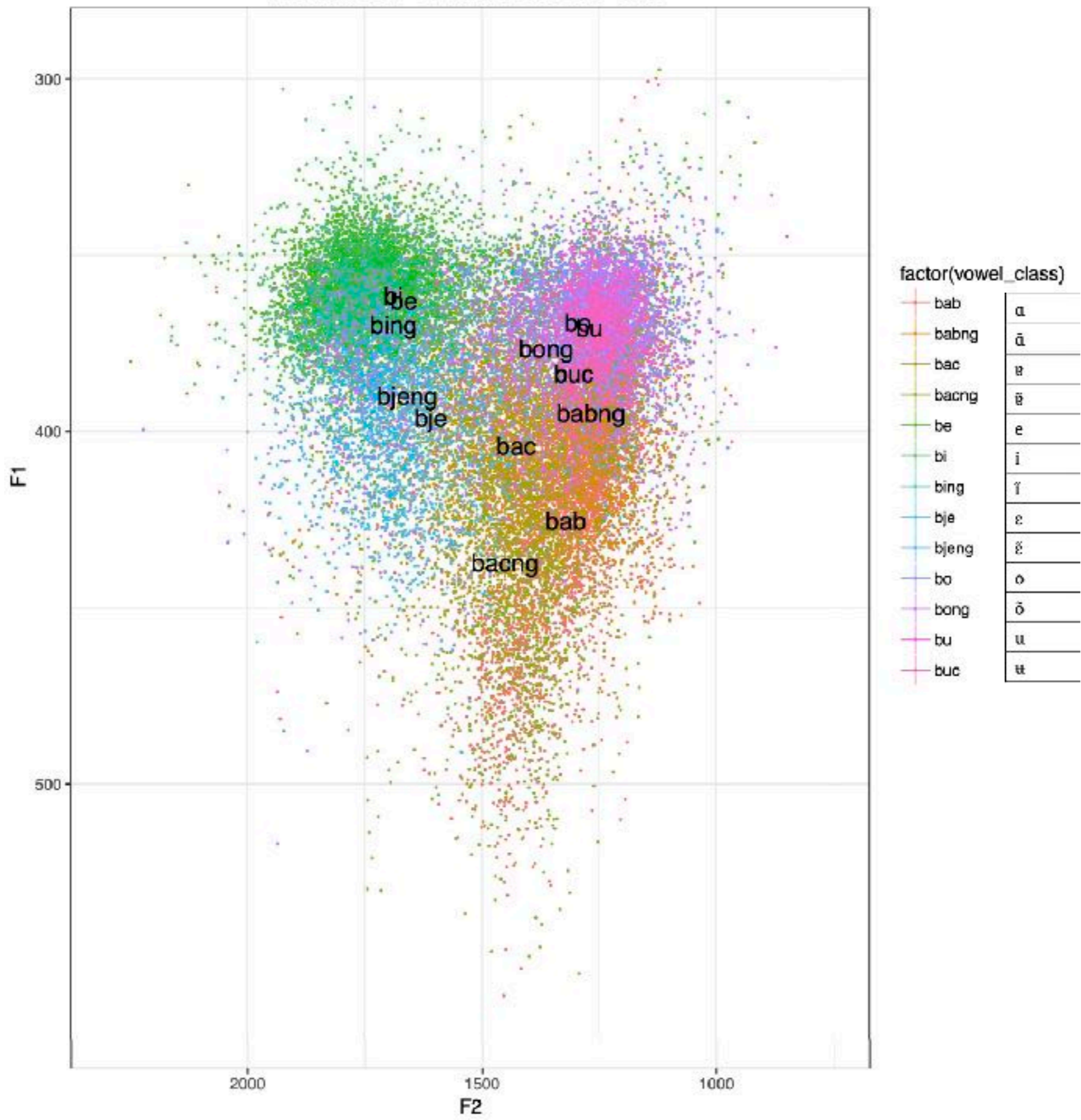


Lower Reaches

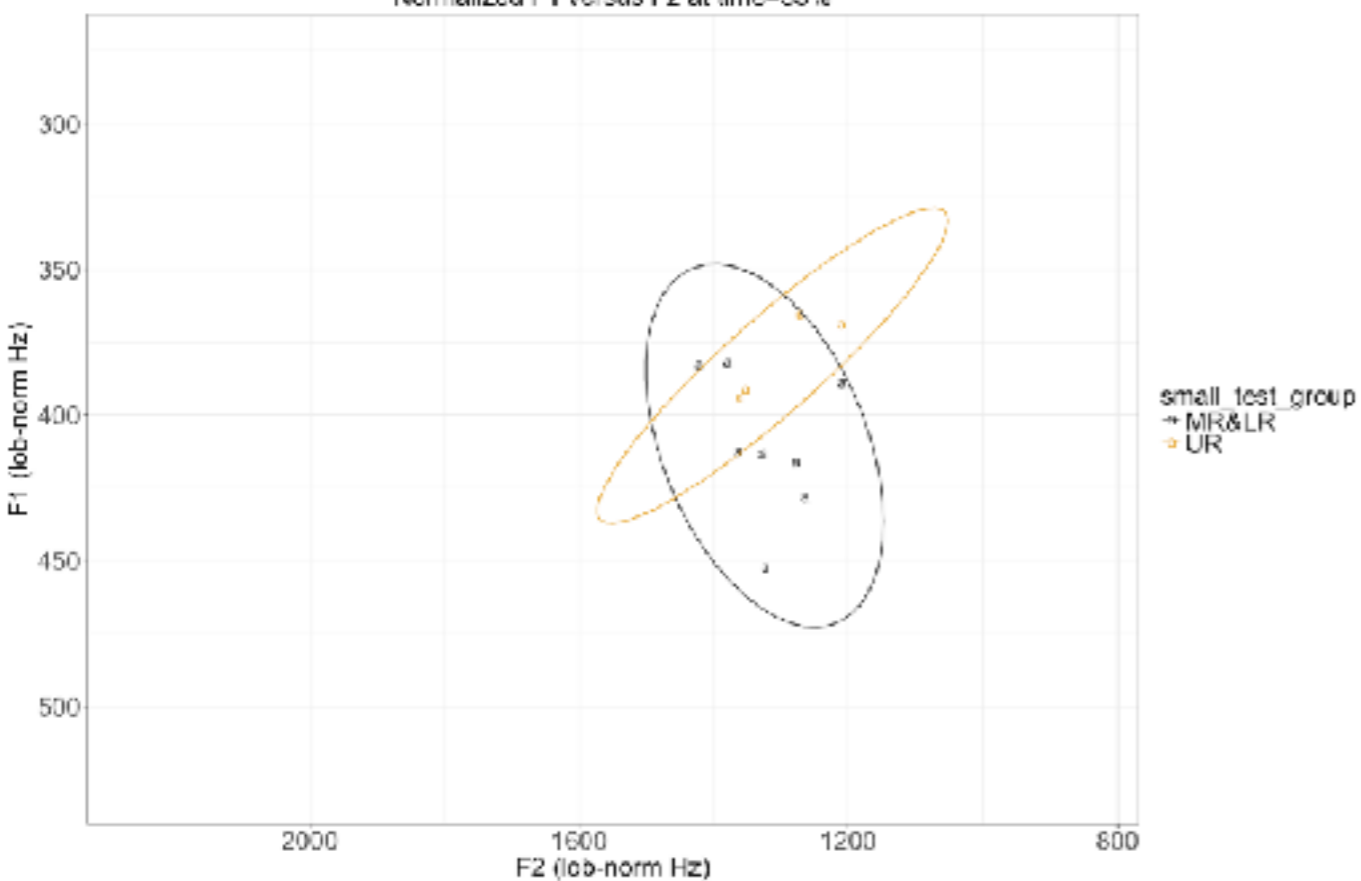
Upper Reaches



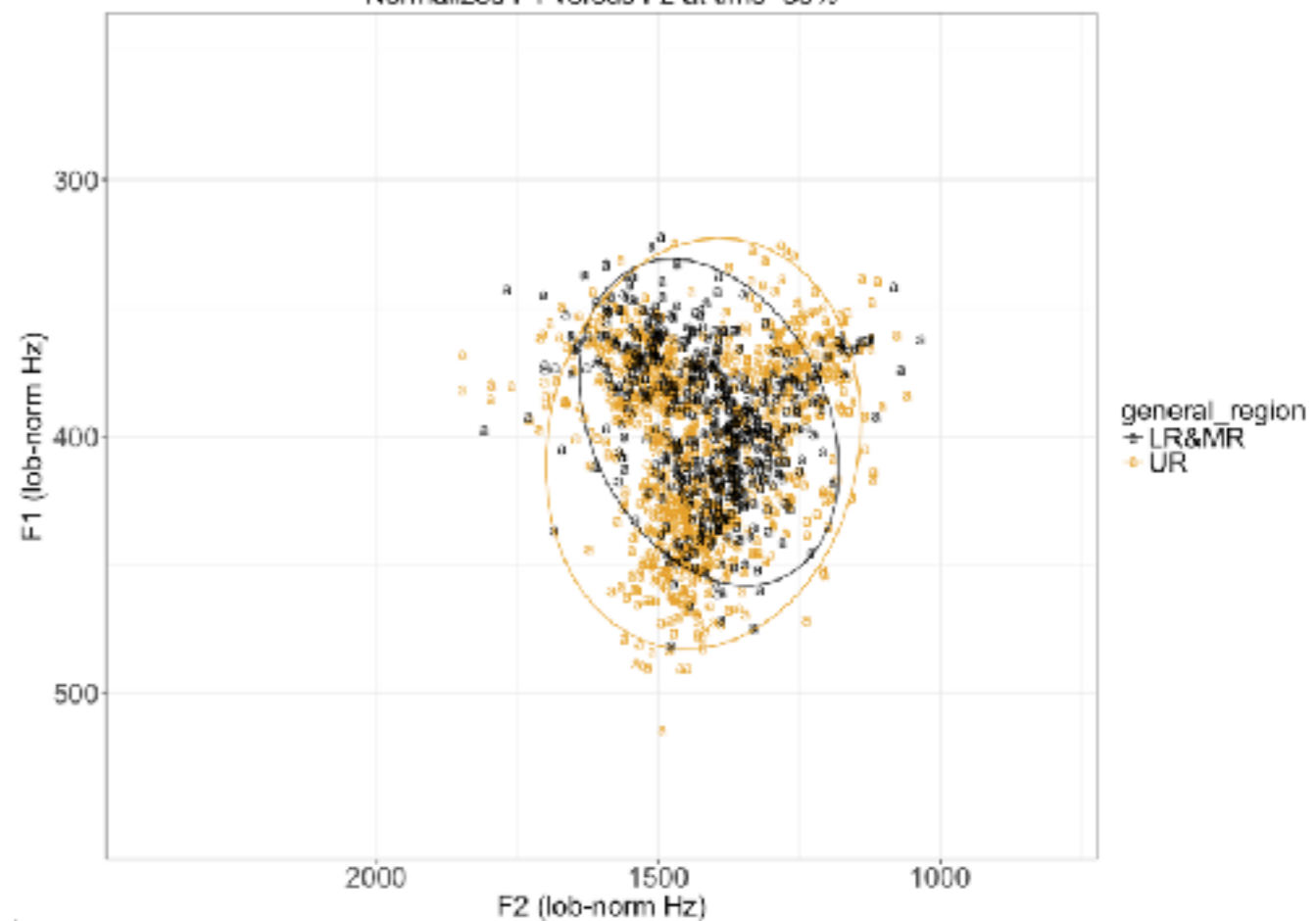
Means of vowel classes
100 speakers, 29,200 tokens
Normalized F1 versus F2 at time=35%



Jackal&Crow
 Speakers 5, 8, 36 (UR group) versus speakers 77, 86, 102 (MR & LR group)
 All tokens of bab vowel (#8 non-nasal) in words 35a and 76a
 Normalized F1 versus F2 at time=35%



Jackal&Crow
 All UR speakers versus all MR/LR speakers
 All tokens of bab vowel in all relevant words:
 82b 78a 94a 90a 35a 87a 32a 34b
 Normalized F1 versus F2 at time=35%



Conclusion

- Labov (1972: 118): The most effective way [...] is to approach a single problem with different methods, with complementary sources of error.
- Data from a variety of distinct sources and methods, properly interpreted, can be used to converge on right answers to hard questions.

Thank you for your attention!

Sound files available at:

<http://elararchive.org>

<https://pangloss.cnrs.fr>

<https://cocoon.huma-num.fr>