LIMITATIONS OF INTELLIGIBILITY: DYNAMIC RELATIONS BETWEEN MIXTEC COMMUNITIES

Sandra Auderset (MPI for Evolutionary Anthropology)
Simon L. Peters (UC Santa Barbara)
Inî G. Mendoza (U Chicago)

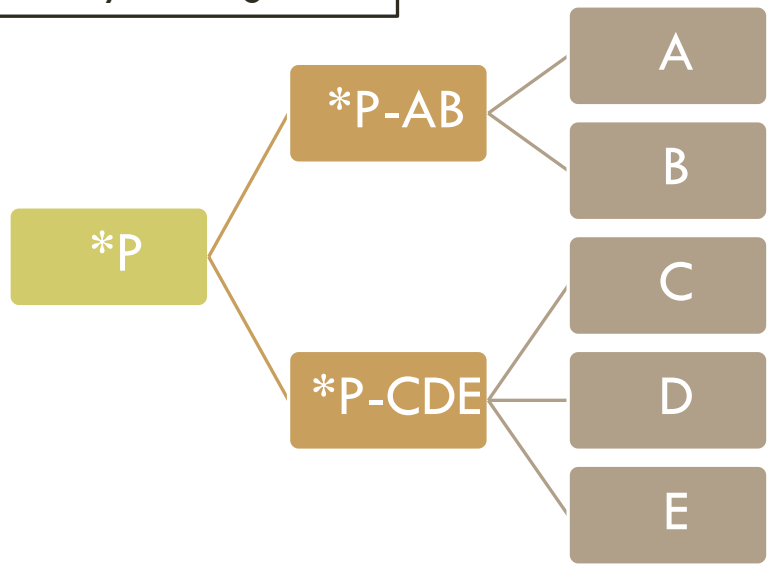
2023 SSILA Annual meeting - online
January 22 2023

ROADMAP

1. Intelligibility and the language/dialect distinction
2. Introduction to the communities (San Martín Peras and San Martín Duraznos)
3. Measures of intelligibility and proxy measures
4. Linguistic similarities and differences between SMP and SMD
5. Perceived intelligibility in the origin and diaspora communities
6. Consequence of the language/dialect classification
7. Conclusion

INTELLIGIBILITY AND THE LANGUAGE/DIALECT DISTINCTION

Languages
not mutually intelligible



Schematic representation of a language family tree with clear splits

Dialects
some mutually intelligible

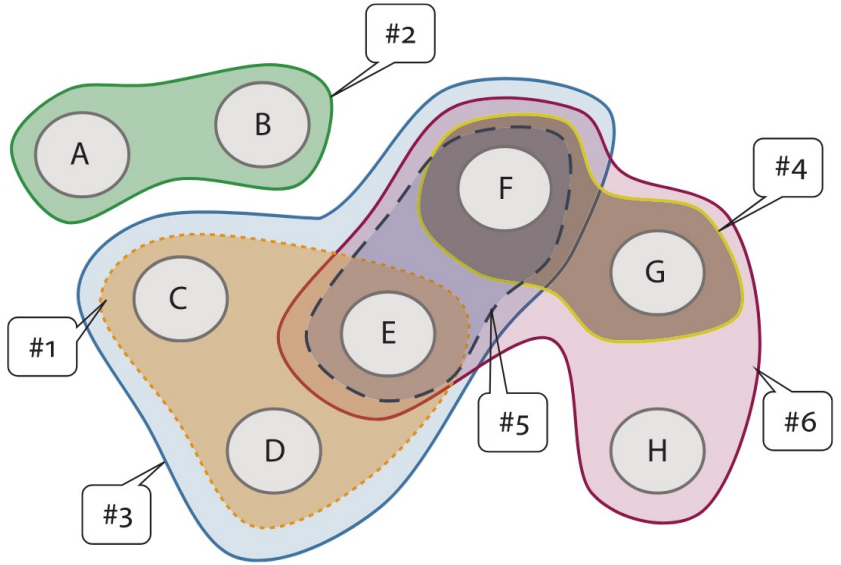


Figure 1. Intersecting isoglosses in a dialect continuum, or a "linkage"

Source: Kalyan, François, Hammarström. 2019. Journal of Historical Linguistics 9(1)

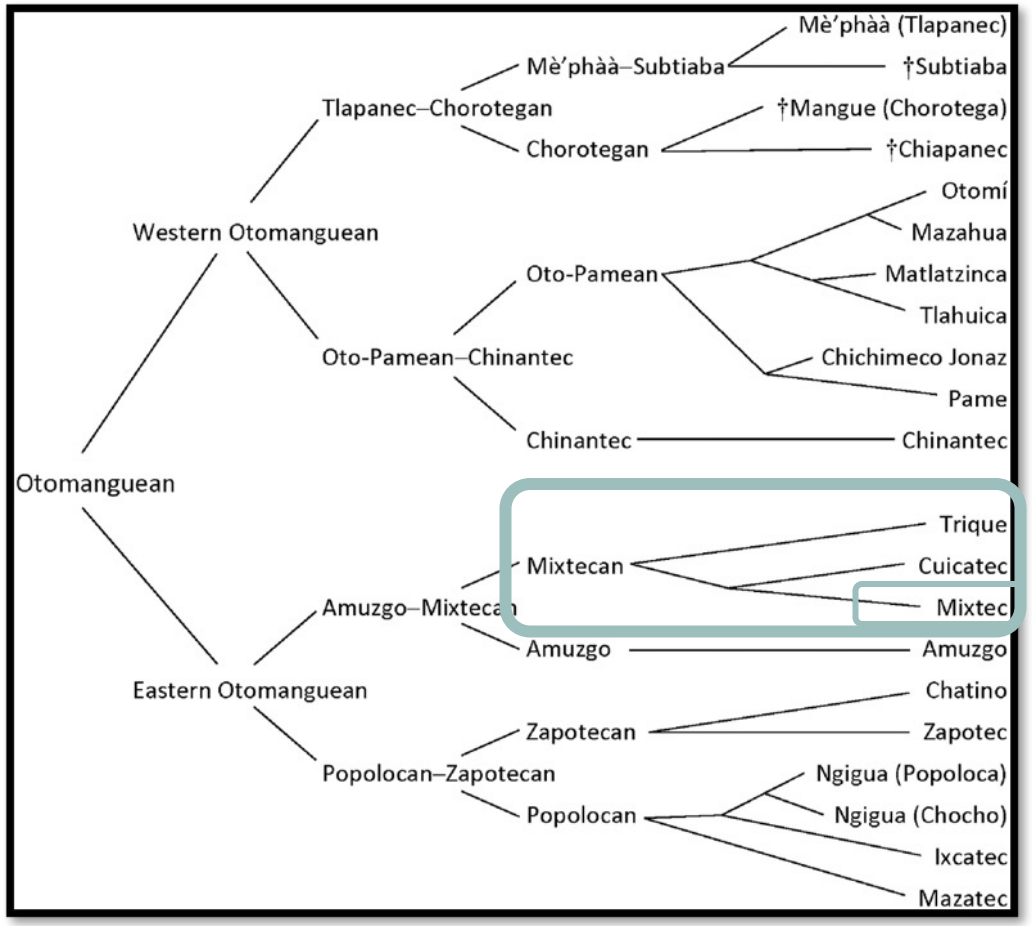
INTELLIGIBILITY AND THE LANGUAGE/DIALECT DISTINCTION

Known issues:

- Idealized models
  - Gradation
  - Difficult to measure
  - Arbitrary cut-offs

- No linguistic basis
  - Determined by extra-linguistic factors
  - Mismatches between proxy measures and perceived intelligibility

INTELLIGIBILITY AND THE LANGUAGE/DIALECT DISTINCTION IN MIXTEC



TU'UN SAVI (MIXTEC):

- characterized as a dialect area/dialect continuum
- unclear how many 'varieties' there are
  - 52 languages with dialect distinctions (Egland 1983; Simons & Fennig 2017)
    => reflected in Ethnologue and Glottolog
  - 81 varieties (INALI 2008)
  - 12 primary dialect areas (Josserand 1983) with classification based on isoglosses of vowel correspondences

COMMUNITIES: LOCATIONS

San Martín Peras (SMP)
- 11,616 speakers in Mexico (INEGI 2020)
~several thousand speakers in the CA diaspora

San Martín Duraznos (SMD)
~400 speakers in Mexico
~50 speakers in the CA diaspora

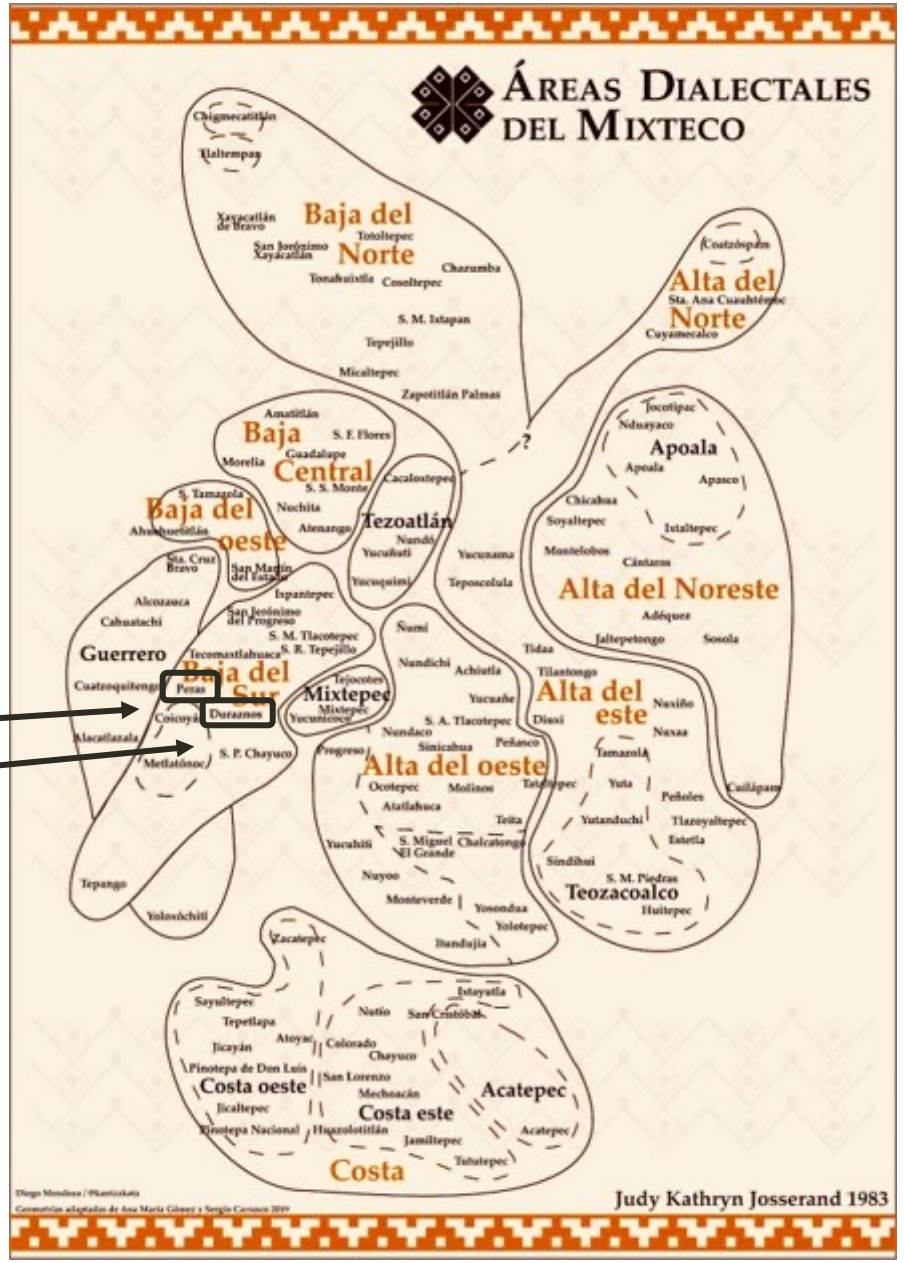
political map of California diaspora communities in Ventura and Santa Barbara Counties

origin communities in Oaxaca (Aguilar Sánchez, 2020:24)

COMMUNITIES: CLASSIFICATION

- Otomanguean > Mixtecan > Mixtec
- Southern Baja group in Josserand 1983
- Glottolog & Ethnologue:
  - San Martín Peras < Western Juxtlahuaca Mixtec
  - San Martín Duraznos < Juxtlahuaca Mixtec

https://glottolog.org/resource/languoid/id/sout3179



Adapted by Diego Mendoza

PROPOSED MEASURES OF INTELLIGIBILITY

Directly quantifying intelligibility:

- Recording/interview method (Voegelin & Harris 1951, Bradley 1968, Kirk 1970, among others)
- Sociolinguistic questionnaire (Padgett 2017, Trentman & Shiri 2020)

Linguistic proxy measures:

- Phonological distance: Normalized Levenstein distance (Wichmann 2019)
- Lexical distance: Number of shared cognates (Swadesh 1950, Padgett 2017)
- Grammatical distance: comparing equivalent sentences (Szeto 2000)

ISSUES WITH MEASURES

Intelligibility determined largely by extra-linguistic factors:

- Asymmetric, non-reciprocal (Olmstead 1954)
- Language attitudes (Wolff 1959, Gooskens & van Heuven 2020)
- Exposure, instruction (Gooskens & van Heuven 2020)
- 'unnatural' task/situation (Wolff 1959)

Proxy measures:

- Focus on lexical items in isolation, not actual language use (Szeto 2000)
- Does not necessarily match with perceived intelligibility

SHARED COGNATES IN SMP AND SMD

Cognate vocabulary: 83% cognate in 209-item basic vocabulary list

	FISH	BIG.SG	WAX	SISTER.F	STOOL
SMP	tsʲ a ¹ k a ⁵	k a ⁵ ʔ n u ³	i ¹ h m a ³⁵	k u ³ ʔ u ¹	tsʲ a ⁵ j i ³
SMD	tɕ a ³ k a ⁵	k a ⁵ ʔ n u ³	ɲ u ¹ m a ¹	k i ¹ ʔ β i ¹	tɕ e ³ i ¹
NLD-t	0.33	0	0.5	0.43	0.83

Partially cognate basic vocabulary

	BIRD	NOPAL	URINE
SMP	s a ³ a ¹	β i ³ ʔ n tsʲ a ³⁵	l a ³⁵ʃ a ³
SMD	l a ³ a ³	t u ¹ + β i ¹ ʔ n tɕ a ⁵	tɕ i ¹ + ʃ a ⁵ ʃ a ⁵
NLD-t	0.4	0.55	0.66

Non-cognate basic vocabulary

	BONE	CORN STALK	DUST
SMP	i ¹ k i ³⁵	ʃ e ¹ ʔ e ¹	j a ⁵ k a ³⁵
SMD	k o ³ ʔ n t o ³	t a ³ k a ¹	m a ³ tʃ i ³
NLD-t	1	0.83	0.83

TONAL INVENTORIES AND MELODIES

SM Duraznos		
Low (L/1)	Mid (M/2)	High (H/5)

SM Peras	
Low (L/1)	Rising (R/35)
Mid (M/3)	Falling (F/31)
High (H/5)	

Attested combinations on bimoraic lexical elements

L-L ʃi¹ta¹ 'tortilla'	M-L ntu³tʃi¹ 'beans'	H-L la⁵ʔβa¹ 'frog'	L-L i¹βi¹ 'two'	M-L ɲu³ʔu¹ 'fire'	H-L sa⁵ʔβa¹ 'frog'	--
L-M ku¹ɲu³ 'meat'	M-M βe³ʔe³ 'house'	H-M ta⁵ʔã³ 'relative'	L-M ki¹si³ 'pot'	M-M βe³ʔe³ 'house'	H-M kʷi⁵tsi³ 'short'	R-M la³⁵ʃa³ 'urine'
L-H nu¹tɕi⁵ 'sand'	M-H tɕa³ka⁵ 'fish'	H-H ʒi⁵ʒi⁵ 'dew'	L-H ju¹tsi⁵ 'sand'	M-H ntʃu³ʃi⁵ 'hen'	H-H ju⁵ju⁵ 'dew'	--
			L-R ʃi¹ta³⁵ 'tortilla'	M-R ntu³tʃi³⁵ 'beans'	H-R ni⁵hĩ̃³⁵ 'thin'	R-R tsʲo³⁵to³⁵ 'rat'
			--	M-F ko³ɲu³¹ 'meat'	--	R-F mi³⁵ni³¹ 'lake'

TONAL INVENTORIES AND MELODIES

SM Duraznos		
Low (L/1)	Mid (M/2)	High (H/5)

SM Peras	
Low (L/1)	Rising (R/35)
Mid (M/3)	Falling (F/31)
High (H/5)	

Attested combinations on bimoraic lexical elements

L-L ʃi¹ta¹ 'tortilla'	M-L ntu³tʃi¹ 'beans'	H-L la⁵ʔβa¹ 'frog'	L-L i¹βi¹ 'two'	M-L ɲu³ʔu¹ 'fire'	H-L sa⁵ʔβa¹ 'frog'	--
L-M ku¹ɲu³ 'meat'	M-M βe³ʔe³ 'house'	H-M ta⁵ʔã³ 'relative'	L-M ki¹si³ 'pot'	M-M βe³ʔe³ 'house'	H-M kʷi⁵tsi³ 'short'	R-M la³⁵ʃa³ 'urine'
L-H nu¹tɕi⁵ 'sand'	M-H tɕa³ka⁵ 'fish'	H-H ʒi⁵ʒi⁵ 'dew'	L-H ju¹tsi⁵ 'sand'	M-H ntʃu³ʃi⁵ 'hen'	H-H ju⁵ju⁵ 'dew'	--
			L-R ʃi¹ta³⁵ 'tortilla'	M-R ntu³tʃi³⁵ 'beans'	H-R ni⁵hĩ³⁵ 'thin'	R-R tsʲo³⁵to³⁵ 'rat'
			--	M-F ko³ɲu³¹ 'meat'	--	R-F mi³⁵ni³¹ 'lake'

SEGMENTAL AND TONAL PROCESSES

SM Peras	Base form	Combined form	Gloss
2SG.NHON ũ^{5} - fuses with final tone of preceding element - overwrites vowel or reduces it to a glide	ja^{5}ti^{3}	ja^{5}tj-ũ^{35}	'your (SG) net'
	tsi^{35}na^{31}	tsi^{35}n-ũ^{15} or tsi^{35}na^{31} sa^{1}n-ũ^{15}	'your (SG) dog'
	kʷi^{1}ka^{1}	kʷi^{1}k-ũ^{15}	'your (SG) comb'

SM Duraznos	Base form	Combined form	Gloss
2SG.NHON ũ assumes final tone of preceding element	ʒa^{5}ti^{5}	ʒa^{5}ti^{5}-ũ^{5}	'your (SG) net'
	tɕi^{1}na^{3}	tɕi^{1}na^{3}-ũ^{3}	'your (SG) dog'
	kʷi^{1}ka^{1}	kʷi^{1}ka^{1}-ũ^{1}	'your (SG) comb'

GRAMMATICAL DISTANCE

VARIETY	POTENTIAL	INCOMPLETIVE	COMPLETIVE	GLOSS
SMP	ku¹si¹	ki⁵ʃi¹	ni¹-ki¹ʃi¹	'sleep'
SMD	ku¹si¹	ki⁵ʃi¹	i¹-ki¹ʃi¹	
SMP	ka³ʃi¹⁵	ʃa⁵ʃi¹⁵	ʃa¹⁵ʃi¹⁵	'eat'
SMD	ka³ʃi³	ʃa⁵ʃi⁵	i¹-ʃa³ʃi³	
SMP	ka³sa³	ki⁵ʃa³	ki¹ʃa³	'do, make'
SMD	sa⁵ʔa³	sa⁵ʔa³	i¹-sa⁵ʔa³	
SMP	ku³u³	ku⁵u³	ku¹u³	'be able to'
SMD	ko³ã³	ki³βi³	i¹-ki³βi³	

LANGUAGE IN USE

Greeting (to a woman, respectful)

SMP		
i⁵jo³	ũ⁵	ʃi³ʃi³
IPFV.exist	2SG.NHON	aunt

SMD		
tʃa⁵a³	ni⁵	ʃi³ʃi³
greeting	2SG.HON	aunt

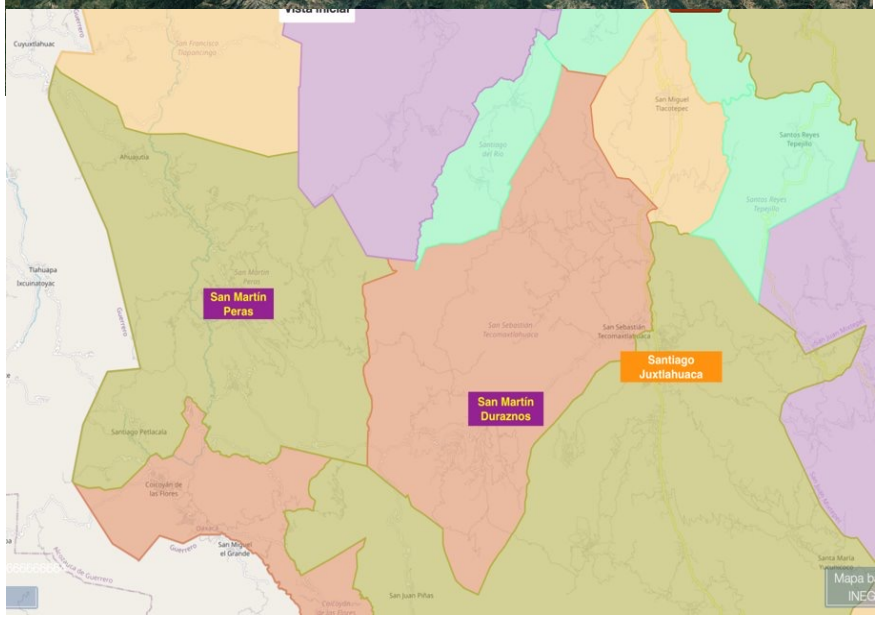
Thanks (to a peer, informal)

SMP	
ta³ʃa¹ʔβi³	ũ⁵
thanks	2SG.NHON

SMD		
ʃi¹nte¹e⁵	i³ni³	ũ³
thanks	inside	2SG.NHON

PERCEIVED INTELLIGIBILITY IN THE ORIGIN COMMUNITIES

- The varieties are closely related (Josserand 1983)
- but not (readily) mutually intelligible (Egland 1983)
- Geographic proximity: 25km apart
- Orient toward different commercial centers
- Belong to separate political districts



PERCEIVED INTELLIGIBILITY IN THE DIASPORA

Higher degree of social proximity in diaspora (Hernández Martínez et al. 2021)

- SMP speakers constitute a plurality
- Many SMD speakers report acquiring SMP with relative ease
- Asymmetrical

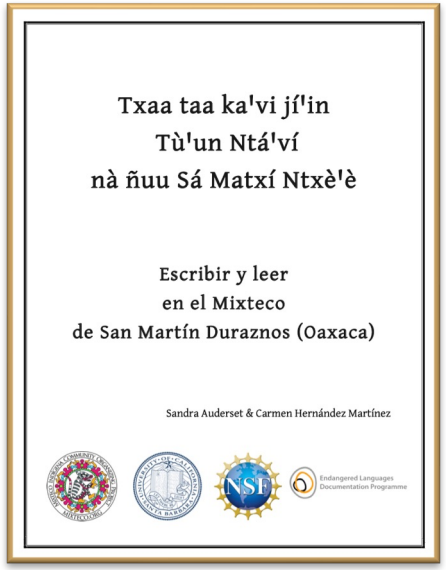
Differing degrees of **social proximity**—not linguistic structure—influence perceptions of linguistic distance

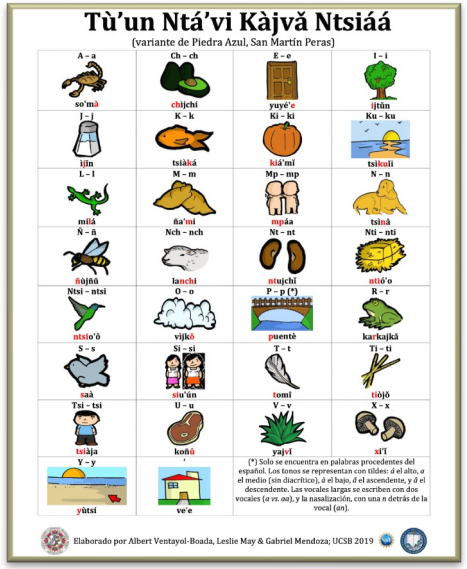
Mismatch between perceptions in the origin vs. diaspora communities can be explained by differing spatial and social relations in each setting

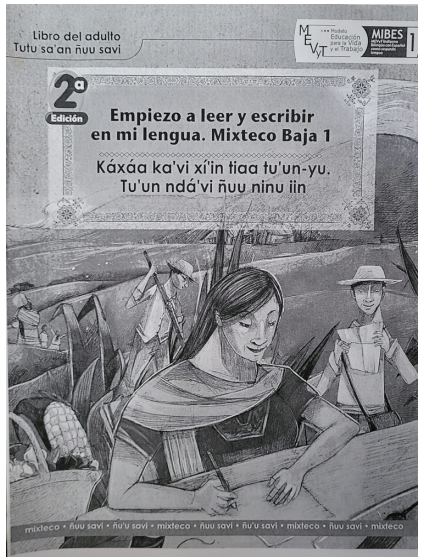
CONSEQUENCES: CATALOGUING AND LANGUAGE ACCESS

- Cataloguing matters for
  - Funding allocation (documentation grants, research grants, etc.)
  - Representation (dialects are more 'hidden', can feel like less important)
  - Community access to information about their own language

- Language Access
  - Pedagogical materials
  - Narrative or oral stories

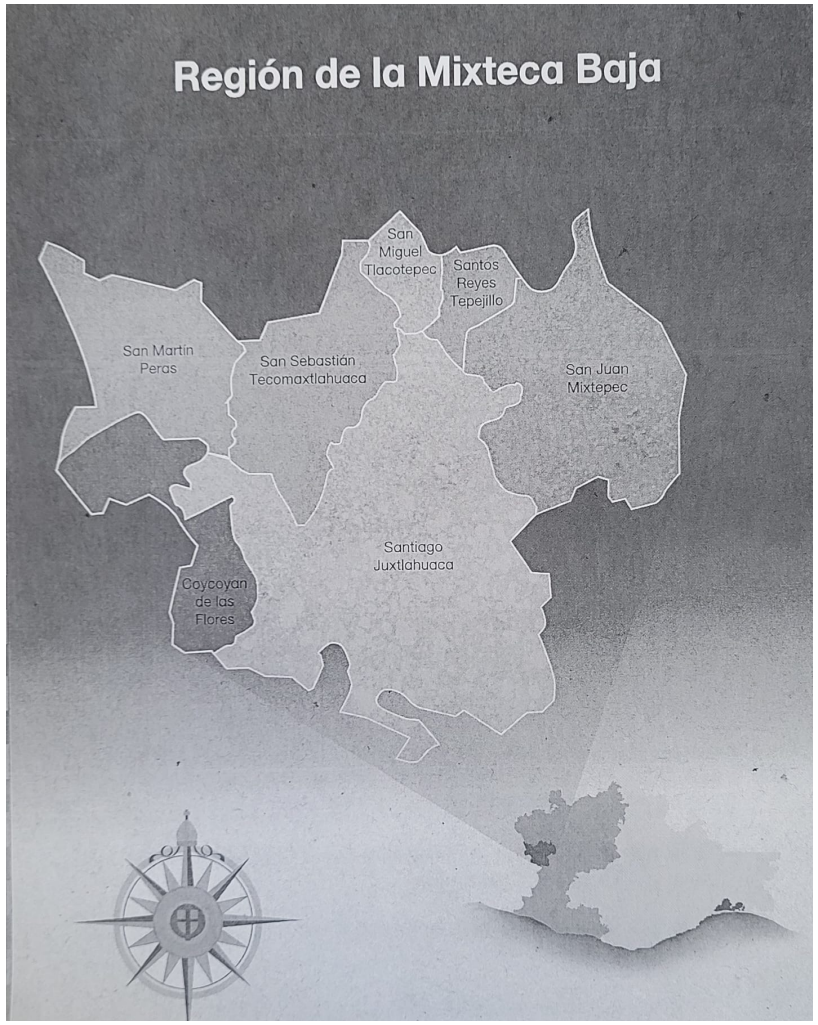






CONSEQUENCES: LANGUAGE JUSTICE AND SOCIAL IMPLICATIONS

- Language justice
  - Interpretation services (courts, medical issues, school, etc.)
  - Translation services (flyers, court documents, and different levels of literacy in Mixtec, etc.)
- Further social Implications
  - Calling Indigenous languages *"dialecto"* as an erasure of Indigenous people and their language
  - History of oppression of indigenous languages in Mexico.



CONCLUSIONS

"Most linguists would agree that it is difficult and often controversial to distinguish languages from dialects. Many, however, would also agree that the notions of language and dialect are still useful, even for the linguist who is aware of the problems of definition that they entail" (Wichmann 2019: 823, emphasis ours)

Measuring 'intelligibility' to classify speech varieties into languages and dialects

- Evades objective assessment independently of social factors
- Provides no knowledge gain – especially not in dynamic situations
- Can have real-life negative impacts for already minoritized communities

Shìntèé ini-ntò!

Taxà'vi-ntó!

THANK YOU!

ACKNOWLEDGEMENTS

Carmen Hernández Martínez

Catalina Martínez Ramírez

Griselda Reyes Basurto

Eric W. Campbell

Mixteco/Indígena Community Organizing Project (MICOP)

ELDP (SG5066)

NSF award #1660355 to UCSB

REFERENCES

Aguilar Sánchez, Omar. 2020. *Ñuu Savi: Pasado, presente y futuro. Descolonización, continuidad cultural y re-apropiación de los códices mixtecos en el Pueblo de la Lluvia*. Leiden University, Doctoral thesis.

Bradley, C. Henry. 1968. A method for determining dialectal boundaries and relationships. *América Indígena* 28(3): 751-760.

Egland, Steven T. 1983. *La inteligibilidad interdialectal en México: Resultados de algunos sondeos*. México, D.F.: Instituto Lingüístico de Verano, A.C.

Gooskens, Charlotte & Vincent van Heuven. 2020. How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables? *Linguistic Approaches to Bilingualism* 10(3):351-379.

Hernández Martínez et al. 2021 Hernández Martínez, Carmen; Eric W. Campbell & Griselda Reyes Basurto. 2021. MILPA (Mexican Indigenous Language Promotion and Advocacy): A community-centered linguistic collaboration supporting Indigenous Mexican languages in California. In Justyna Olko & Julia Sallabank (eds.), *Revitalizing endangered languages: A practical guide*, 216–217. Cambridge: Cambridge University Press.

Instituto Nacional de Estadística Geografía e Informática (INEGI). 2020a. Principales resultados por localidad (ITER). *Censo de Población y Vivienda 2020*. CSV file for state of Oaxaca (ITER_20CSV20.csv), retrieved from https://www.inegi.org.mx/app/scitel/Default?ev=9#, accessed 4 November 2021.

Josserand, Judy Kathryn. 1983. *Mixtec dialect history*. Tulane University, PhD dissertation.

Kalyan, Siva, Alexandre François & Harald Hammarström (eds.). 2019. *Understanding language genealogy: Alternatives to the tree model*. Special issue of the Journal of Historical Linguistics 9(1).

REFERENCES

Kirk, Paul L. 1970. Dialect intelligibility testing: The Mazatec study. *International Journal of American Linguistics* 36(3): 205-211.

Olmstead, David L. 1954. Achumawi-Atsugewi non-reciprocal intelligibility. *International Journal of American Linguistics* 20(3): 181-184.

Padgett, Erin. 2017. *Tools for assessing relatedness in understudied language varieties: A survey of Mixtec varieties in Western Oaxaca, Mexico*. Grand Forks, ND University of North Dakota, MA thesis.

Simons, Gary F. & Charles D. Fennig. 2017. Ethnologue: Languages of the World. 20th edn. Dallas: SIL International

Swadesh, Morris. 1950. Salish Internal Relationships. *International Journal of American Linguistics* 16(4): 157-167

Szeto, Cecilia. 2000. Testing intelligibility among Sinitic dialects. *Proceedings of the Australian Linguistic Society.*

Trentman, Emma & Sonia Shiri. 2020. The mutual intelligibility of Arabic dialects: Implications for the language classroom. *Critical Multilingualism Studies* 8(1): 104-134.

Voegelin, C. F. & Zellig S. Harris. 1951. Methods for determining intelligibility among dialects of natural languages. *Proceedings of the American Philosophical Society 95(3): 322-329.*

Wichmann, Søren. 2019. How to distinguish languages and dialects. Association for Computational Linguistics 45(4): 823-831

Wolff, Hans. 1959. Intelligibility and inter-ethnic attitudes. *Anthropological Linguistics* 1(3): 34-41.

Variety	WAX	BIRD	STOOL	BONE	NOPAL	STONE
SMP	i ^{1} h m a $_{1}{}^{5}$	s a ^{3} a ^{1}	tsʲ a ^{5} ɟ i $_{3}$	i ^{1} k i $^{1}{}^{5}$	β i ^{3} ʔ ^{n}tsʲ a $^{3}{}^{5}$	ʒ u ^{1} ʔ u ^{5}
SMD	ɲ u ^{1} m a $_{1}$	l a ^{3} a ^{3}	tɕ e ^{3} i $_{1}$	k o ^{3} ʔ ^{n}d o $_{3}$	t u ^{1} + β i ^{1} ʔ ^{n}dʑ a ^{5}	ʒ i ^{1} i ^{1}
	MAN	FISH	CAT			
SMP	tsʲ a ^{1} ^{h} a $_{3}$	tsʲ + a ^{1} k a ^{5}	tʃu ^{5} ^{h} tu $_{3}$			
SMD	tɕ a ^{1} a ^{3}	tɕ + a ^{3} k a ^{5}	tʃ i ^{5} t i $_{5}$			

OUTLINE FOR TALK (TO BE DELETED FOR PRESENTING)

SANDRA [5-6min]

- Introducing the problem: why the concept of intelligibility is problematic
- Background on communities, speakers, and languages
- Linguistic proxy measures
- SIMON [7-8min]
- linguistic similarities and differences between SMD and SMP
- Intelligibility in the origin communities vs. the diaspora

INI [4-5min]

- Consequences of the language/dialect split: cataloguing, funding, language access, pedagogical materials, etc.

SANDRA [1-2min]

- Conclusion: there is no 'objective' measure of intelligibility, no way to define language vs dialect (as we've known since forever...)

PHONEME INVENTORY?

[consolidated table]