



Lexical and morphological prenasalization in Sà'án Sàvĩ ñà Ñuù Xnúvíkó

Guillem Belmar & Jeremías Salazar

This presentation



Language: Sa'an Savi ña Ñuu Xnuviko



Prenasalized segments in Mixtepec Mixtec



Lexical vs. Morphological
prenasalization: who do I mean?

Perfect aspect



The study



Results



Discussion



Side note: experimental phonetics & Fieldwork



What is next? Prenasalization (?) in Mixtec (and Otomanguean)

Acknowledgments

- Big thank you to the Mixtepec Mixtec speakers who agreed to do the task and to being recorded: Alejandra Garcia, Claudia Salazar, Francisco Bautista, Miguel Martinez, Silverio Garcia and Yuridia Garcia.
- Big thank you to everyone in the Phonetics Lab (Argyro's Lab) in Fall 2022, our discussions helped shape much of the experimental design and the analysis
- Big thank you to Sherry especially for her help discussing some of the stats!



Sà'an Sàvī ñà Ñuù Xnúvíkó (Mixtepec Mixtec)



- Mixtec variety spoken in the municipality of Mixtepec (district of Juchitahuaca, Oaxaca, Mexico)
- About 9,000 speakers (?)
- One of the main branches of Mixtec (Josserand 1983)
- Little information on this variety (one *Phonological sketch of the Yucunany dialect of Mixtepec Mixtec* (Paster & Beam de Azcona 2004))

Consonants of Mixtepec Mixtec





		Bilabial	Dento-alveolar	Alveolar	Palatal	Velar	Labio-velar	Glottal
Plosive	Plain	p	t			k	k ^w	ʔ
	Prenasalized	^m p	ⁿ t			^ŋ k	^ŋ k ^w	
Nasal								
Affricate	Plain			[̄] ts	[̄] tʃ			
	Prenasalized			ⁿ [̄] ts	ⁿ [̄] tʃ			
Fricative				s	ʃ	x		(h)
Approximant		β			j		(w)	
Tap				r				
Trill				r				
Lateral approximant				l				

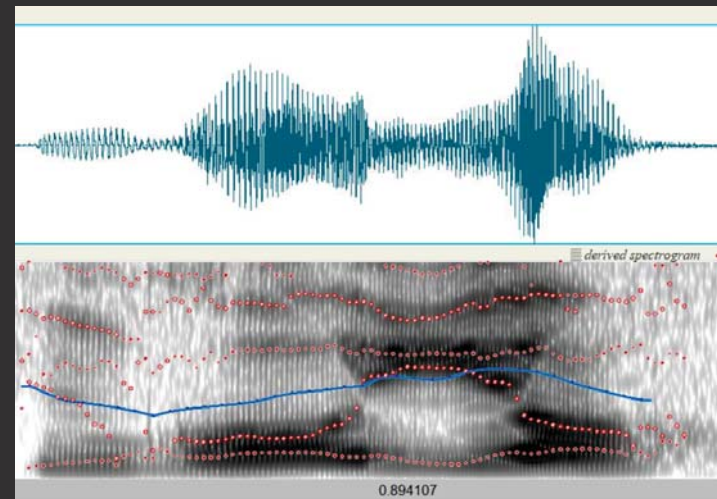
Consonants of Mixtepec Mixtec

		Bilabial	Dento-alveolar	Alveolar	Palatal	Velar	Labio-velar	Glottal
Plosive	Plain	p	t			k	k ^w	ʔ
	Prenasalized	^m p	ⁿ t			^ŋ k	^ŋ k ^w	
Nasal								
Affricate	Plain			^h ts	^h tʃ			
	Prenasalized			ⁿ ts	ⁿ tʃ			
Fricative				s	ʃ	x		(h)
Approximant		β			j		(w)	
Tap				r				
Trill				r				
Lateral approximant				l				

Prenasalized stops and affricates in Mixtepec Mixtec



- *nkuii* [ŋg^wi:²³] ‘fox’ 
- *ncho'o* [n^dʒo³ʔo³] ‘hummingbird’ 
- *ntintsitsă* [n^di³ⁿdzi¹t^sa¹⁴] ‘turtle’ 
- *ntuchĭ* [n^du³tʃi¹⁴] ‘bean’ 
- *Nköyô* [ŋgo¹³jo⁴¹] ‘Mexico’



- *páà* [pa:⁴¹] ‘bread’  vs. *mpáà* [m^bba:⁴¹] ‘co-father’ 
- The most frequent of these are ⁿt and ⁿtʃ

Lexical vs. Morphological prenasalization in Mixtepec Mixtec

- Phonetically, these are often voiced, or at least **partially voiced**
- Prenasalization in Mixtepec Mixtec can be:
 - **Lexical:** can be traced back (for the most part) to Proto-Mixtec (Josserand 1983)
 - present-day *ntuchĩ* [n^{du}3tʃi¹⁴] ‘bean’ from proto-Mixtec **nduti*?
 - Not followed by nasal vowels
 - **Morphological:** due to processes of segmental erosion (Heine & Reh 1984) triggered by grammaticalization = currently widespread in Mixtepec Mixtec
 - *ntĩvì* [n^{di}12βi¹] ‘PFV.blow’ (compare with *tívì* [ti⁴βi¹] ‘IPFV.blow’)
 - Nasal vowels after these segments are possible: *ntàan* [n^{dã}:13] ‘PFV.quake’ (compare with *tàan* [tã:413] ‘IPFV.quake’)

Morphological prenasalization in Mixtepec Mixtec

- Prospective aspect

- *kítsáá* [ki⁴tʂa:⁴⁴] ‘IPFV.start’ vs. *kú nkítsáá* [ku^{4ŋ}gi¹tʂa:⁴⁴] ~ [ũ^{4ŋ}gi¹tʂa:⁴⁴] ‘PROSP.start’

- Perfective aspect

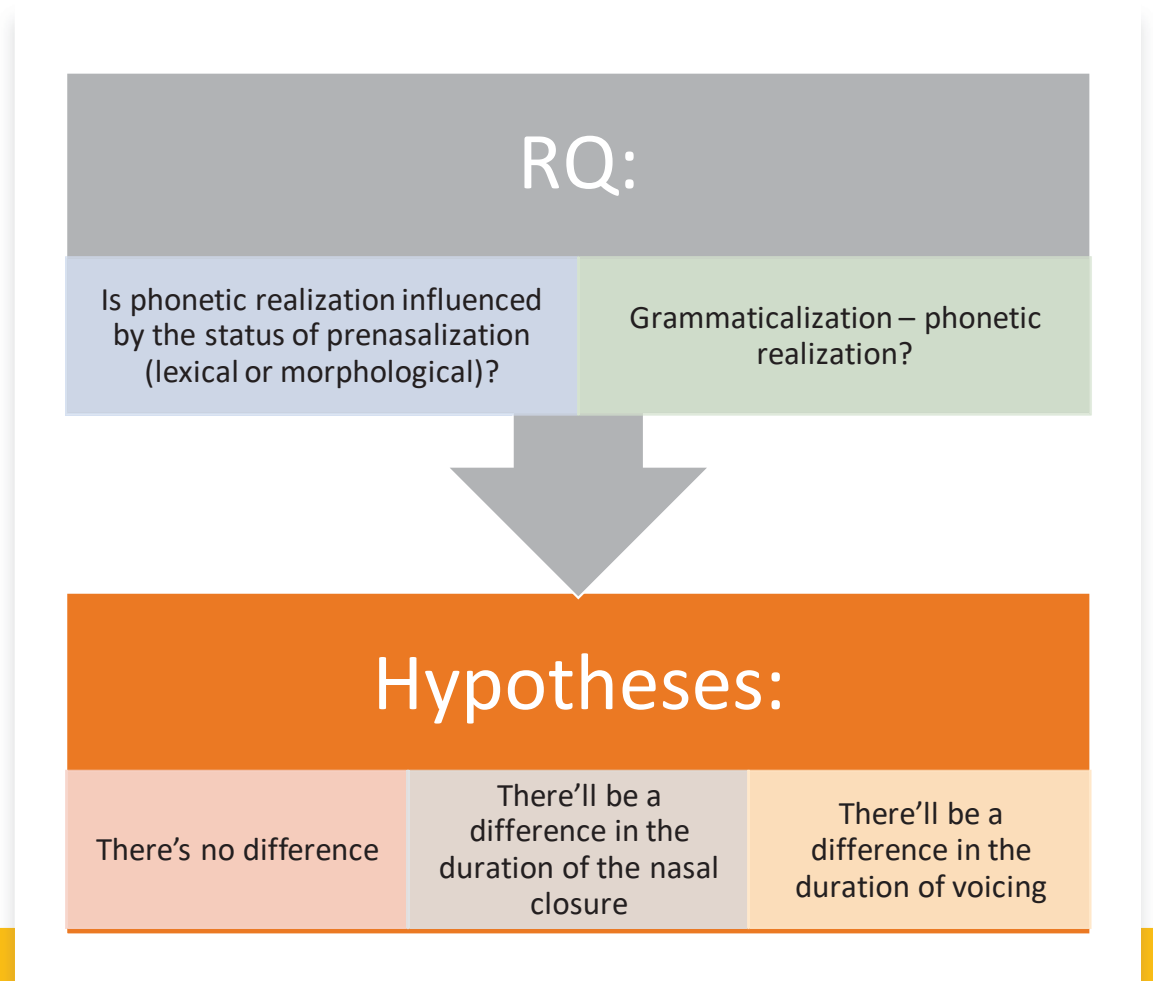
- *cháa* [tʂa:⁴³] ‘IPFV.write’ vs. *nchàa* [ŋ¹³dʒa:¹³] ‘PFV.write’

- Negation

- *katsí* [ka³tʂi⁴] ‘POT.eat’ vs. *nkătsí* [ŋ¹³ga¹³tʂi⁴] ‘NEG.POT.eat’

- Diminutive (?)

Study



First thought....

- Let's take a look at our corpus of unplanned speech, selecting recording from one speaker
- But...

	Plain		Pre-nasalized (lexical)		Pre-nasalized (morphological)	
	Texts	Wordlist	Texts	Wordlist	Texts	Wordlist
/t/ - / ⁿ t/	387	34	231	24	9	0
/ts/ - / ⁿ ts/	298	16	21	4	87	0
/tʃ/ - / ⁿ tʃ/	357	26	76	12	8	0
/k/ - / ⁿ k/	589	42	24	2	10	0

Number of tokens

So, experimental study

- Elicitation task
- 6 participants (recorded using a Tascam Dr-40X and a Shure WH20XLR Dynamic Headset microphone)
 - 3 males and 3 females
 - between the ages of 20 to 60
 - Self-identified as native speakers of Mixtepec Mixtec
 - Due to background noise (and breathiness) data for one male speaker was dropped
- Focusing on ⁿt and ⁿtʃ (by far the most common in lexical items)
- Similar (minimal pairs, near minimal pairs when possible) words that present lexical and morphological prenasalization:
 - 20 nt-words lexical
 - 20 nt-words morphological
 - 20 nch-words lexical
 - 20 nch-words morphological

Side note: of course, this required us to update our IRB protocols and our consent forms!

Morphological prenasalization in Mixtepec Mixtec

- **Prospective aspecto**

- *kítsáá* [ki⁴t^{sa:44}] 'IPFV.start' vs. *kú nkítsáá* [ku⁴ŋgi¹t^{sa:44}] ~ [ũ⁴ŋgi¹t^{sa:44}] 'PROSP.start'

- **Perfective aspect**

- *cháa* [tʃa:⁴³] 'IPFV.write' vs. *nchàa* [n¹dʒa:¹³] 'PFV.write'

- **Negation:**

- *katsí* [ka³t^{si4}] 'POT.eat' vs. *nkătsí* [ŋga¹³t^{si4}] 'NEG.POT.eat'

- **Diminutive (?)**

Since the prospective forms of verbs include a prefix that may be variably realized as [ku⁴], [u⁴], or [ũ⁴], thus presenting potential difficulties for acoustic measurements, the perfective form of verbs is taken to represent morphological prenasalization in this study.

Perfective aspect
in Mixtepec
Mixtec

IPFV	PFV	NEG.PFV	English
<i>kixi</i> [ki ⁴ ʃi ¹]	<i><u>n</u>ikixi</i> [ni ¹ ki ¹ ʃi ¹]	<i>kuě <u>n</u>ikixi</i> [k ^w e ¹⁴ ni ⁴ ki ¹ ʃi ¹]	Sleep
<i>tsika</i> [tsi ⁴ ka ³]	<i><u>n</u>tsika</i> [n ^d zi ¹ ka ³]	<i>kuě <u>n</u>itsika</i> [k ^w e ¹⁴ ni ⁴ tsi ¹ ka ³]	Walk
<i>né'ě</i> [ne ⁴ ʔe ¹⁴]	<i><u>n</u>é'ě</i> [ne ¹ ʔe ¹⁴]	<i>kuě <u>n</u>iné'ě</i> [k ^w e ¹⁴ ni ⁴ ne ¹ ʔe ¹⁴]	Get

Stimuli – Video

- Jeremías Salazar (speaker) recorded the audio for the stimuli sentences.
- Participants watch a video presenting the stimuli sentences and see the image (illustrating the meaning of the target words).
- Participants then say the carrier sentence twice
vàtsi tù'un kávi-rà___ sàtä iin líbrù
βa¹tʃi³tũ¹ʔũ³ka¹βi³ra¹ | __ | sa¹ta¹⁴ʔi:³³li⁴βru¹
X appears in the words he is reading in this presentation
- 80 target words per participant

Takuni ntà'vĩ-rà kò'ö



Vàtsi tù'un ká'vi-rà _____ sàtä iin líbrù



Audio: Claudia Salazar

Procedure

- Explaining task + given an example
 - We developed a script for Jeremías, for presenting the experiment and for explaining it further after recording
- The same video was played for all participants
 - 20-second-long pauses in between stimuli so they had time to repeat the carrier sentences

Ejemplo:

Ntsà'àn-yù ntákuaàn **chakǎ** takuni

Vàtsi tù'un ká'vi-rà ***chakǎ*** sàǎ iin líbrù

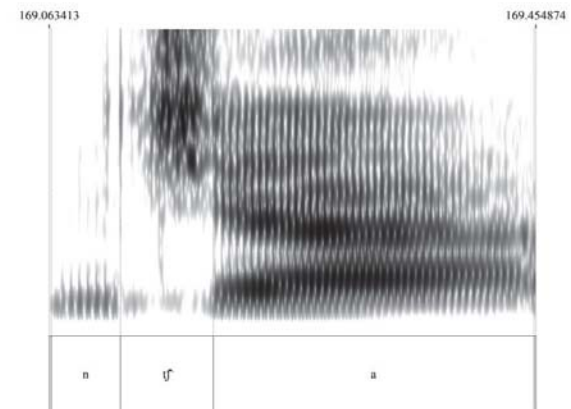
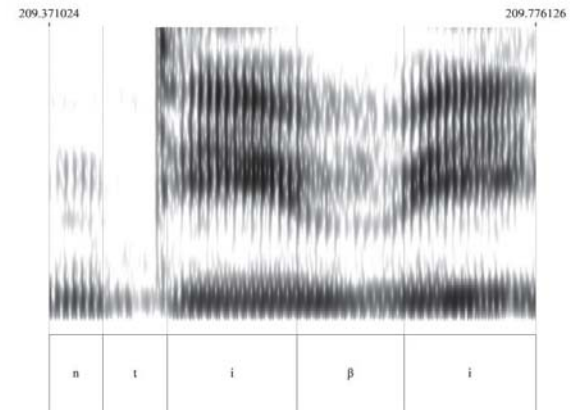
Target words & Stimuli sentences

- We have glossed and translated versions of all these sentences, in case anyone is interested
- We should publish this material somewhere to accompany our paper, but where?

Word (L)	Meaning	Word (M.pfv)	Meaning
Ntusù	'voice'	Ntǔ'un (tú'un)	'get pulled out'
Ntii	'Dead'	Ntĩin (tíin)	'grab'
Ntìvǐ	'egg'	Ntìvì (tìvì)	'blow'
Ntá'vì	'poor; humble'	Ntà'vǐ (tá'vǐ)	'break'
Ntàà	'flat'	Ntàan (tâan)	'quake'
Ntùchǎ	'sea; ocean'	Ntǔu (tûu)	'rise(sun)'
Ntuchǐ	'bean'	Ntùtsǐ (tútsǐ)	'get hurt'
Ntàva	'wood'	Ntavǎ (távǎ)	'take out'
Ntakuǐ	'strong'	Ntàvi (távi)	'toast'
Nta'á	'hand'	Ntǎ'àn inì (tá'àn inì)	'like'

Measurements

- Duration of the nasal closure (%)
- Duration of the oral closure (%)
- Duration of voicing in the oral closure %
 - I included the burst and the fricative in this “closure”
- Additional coding for:
 - Speaker
 - Order (first or second time uttering the carrier sentence)
 - Vowel following the prenasalized segment
 - Number of syllables of the Word

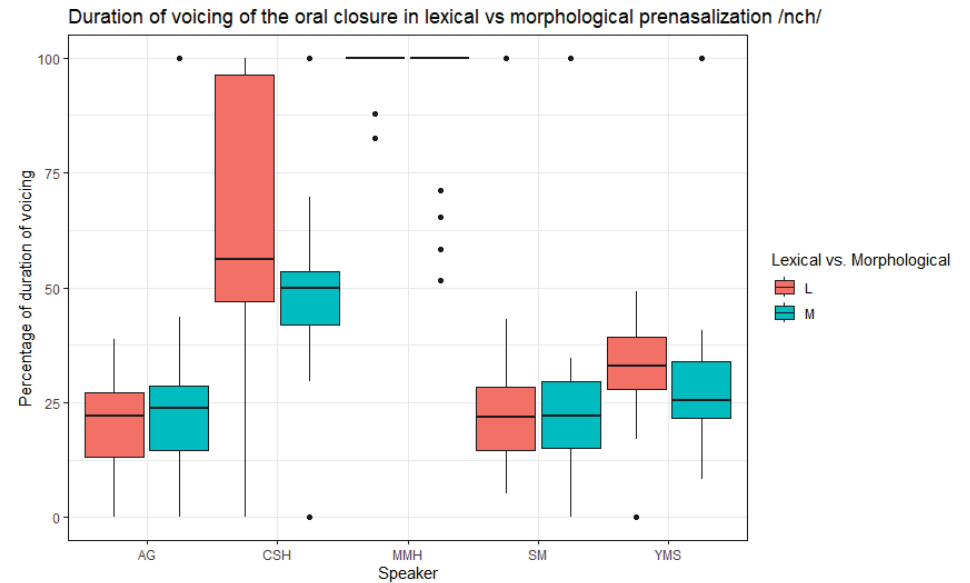
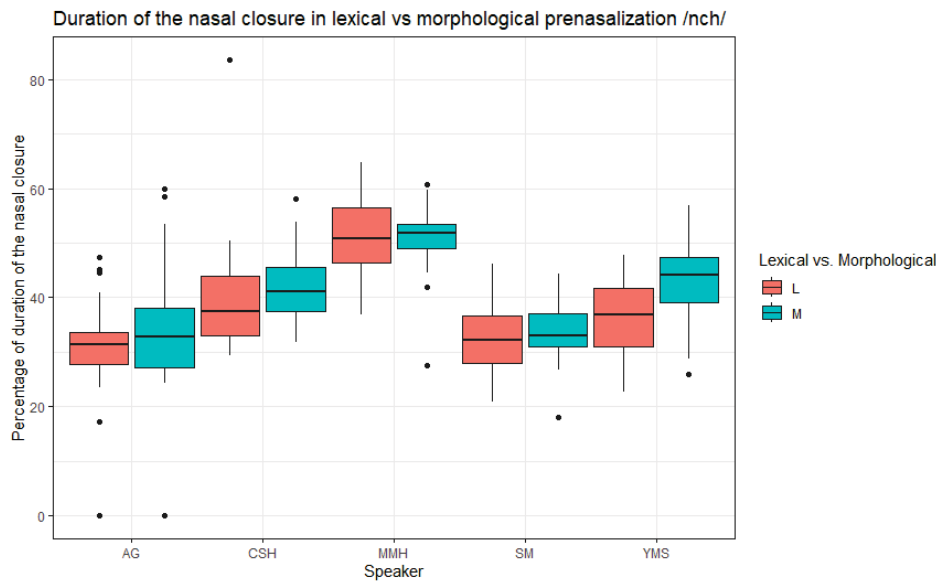


Analysis

- ImerTests determined that Speaker was the only random effect that was significant
- /ⁿt/ and /ⁿtʃ / behave very differently, so we conducted separate analyses
- 4 linear models (2 per segment)
 - DV: duration of the nasal closure OR Duration of voicing in the oral closure
 - IV: L_M*Speaker

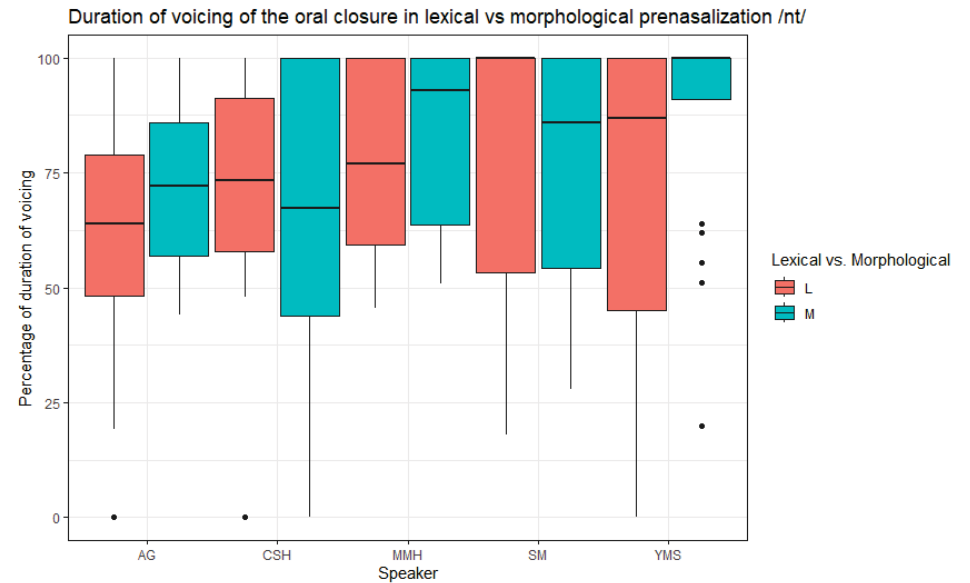
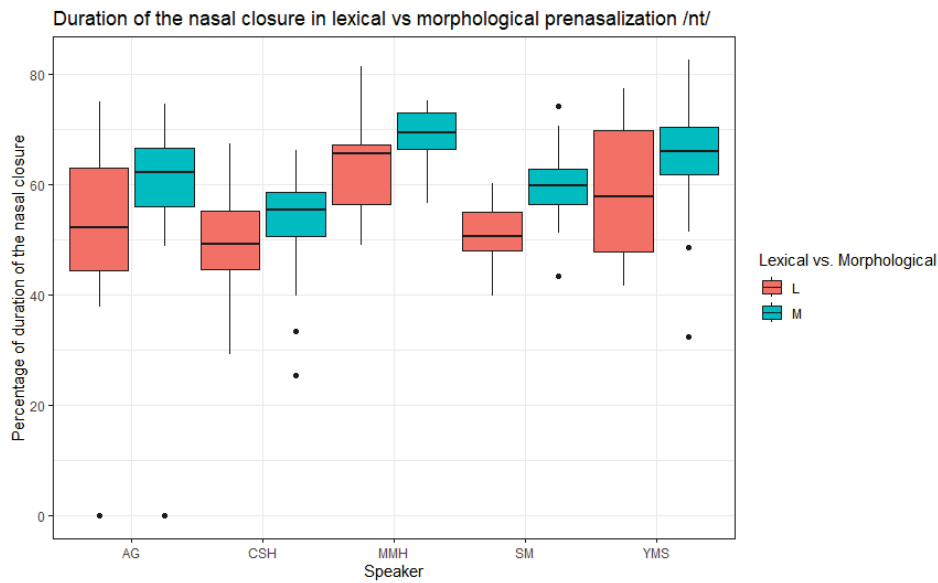
Results /ⁿtʃ/

- **No significant difference** in the relative duration of the nasal closure OR the duration of voicing in the oral closure as a function of L_M, and no interaction with Speaker
- Huge variability among Speakers



Results /ⁿt/

- **Significant difference** in the duration of the nasal closure between lexical and morphological pre-nasalization ($\beta = 10.86$, $p < 0.005$), and no interaction with Speaker.
- No significant model for voicing.
- Less variability among Speakers



Discussion – Lexical vs. Morphological?

- For /ⁿt/ the duration of the nasal closure was **significantly longer** (61.2%) for morphological pre-nasalization than for lexical pre-nasalization (54.15%), although they're both over 50%.

Caveats: 5 speakers, 400 tokens per segment. This was not found for the affricate /ⁿtʃ/.

But: clear trend for /ⁿt/

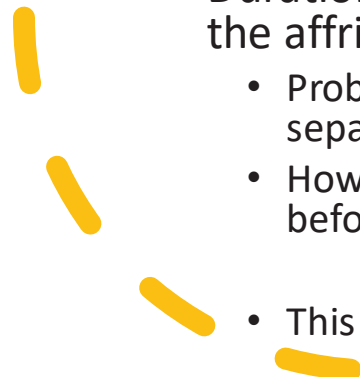
Segmental erosion (grammaticalization process):

- from *nì* to *n+C* to *pre-nasalizedC*?
- Markedness / Information load
 - Arguably, although pre-nasalization is not the only marker of aspect in these form (TONE!)
 - *Tíin* [tĩ:⁴⁴] 'IPFV.grab' vs. *ntiin* [nđĩ:¹³³] 'PFV.grab'



Discussion – Lexical vs. Morphological?

- For /ⁿtʃ/ there is no significant difference in the duration of the nasal closure between morphological pre-nasalization (40.05%) and lexical prenasalization (38.1%), but below 50%
- But why not /ⁿtʃ/?
 - Duration of voicing: It is harder to maintain voicing through an affricate (Ohala & Solé 2008; Zygis et al. 2012)
 - Duration of the nasal stop: Perhaps because of the already complex articulation of the affricate: less time to do more things
 - Problem with the measurements (?): I treated /tʃ/ as a unit. If /t/ and /ʃ/ were segmented separately, the nasal closure would be overwhelmingly longer than the oral closure.
 - However, we know that lexical /ⁿtʃ/ in Mixtepec Mixtec actually comes from Proto-Mixtec *ⁿd before front vowels and before the glide [j]
 - This would not affect the results of the study



Side note: Experimental phonetics & Fieldwork

- Experimental paradigm →
 - (almost) non-negotiable importance of tightly controlled settings, procedures, stimuli...
 - (often) heavy reliance on written representations of data
 - (often) complex procedures
- Unrealistic ideal when working with speakers of minoritized and Indigenous languages
 - Especially in areas where access to research institutions is nearly non-existent

Experimental phonetics & Fieldwork

“Maybe we can get 5 repetitions from each target word”



“Two is already a lot, if we ask for more we’ll end up having a recording in which they ask us why so much, and they will quite before finishing the task”

Experimental phonetics & Fieldwork

“We should control
for word
frequency”



We don't have
large corpora to
calculate word
frequency

Experimental phonetics & Fieldwork

“Stimuli sentences should have the same type of focus”



“We don’t know much about focus in this language yet”



If you don’t speak the language, some of these “limitations” can be very hard to explain to your colleague(s) who will be the one(s) coming up with the sentences.

Vàtsi tù'un ká'vi-rà _____ sàǎ iin líbrù

Experimental phonetics & Fieldwork

“Ideally sessions will take place in the same environment”



In reality, you will be recording each of these in different days, likely at the speaker's house. You cannot control that no distraction will occur.

Experimental phonetics & Fieldwork

“Wait, so what is X?”



When working with these languages you cannot presuppose ANY previous knowledge. You need to EXPLAIN EVERYTHING, and sometimes you simply don't have the space or the time.

It can be done (and we should do it!), but we also need to train our fellow lab experimental people to engage with our research differently than they would engage with their own!



What about now?

- We are working on a JIPA Illustration for Mixtepec Mixtec (with Eric)
 - And a big shout-out to my awesome team RAs!: Alice, Christopher, Katie, Lily, Marinah & Miriam
- There's only one (very recent) illustration of a Mixtec variety (Cortés et al. 2023)
- Thinking about how to include this (and more) and engaging in broader discussions about these segments

Prenasalization in Mixtec (and other Otomanguean languages)

- Commonly described as prenasalized, often voiced, segments (e.g., Cortés et al. 2023; Iverson & Salmons 1996; Rueda Chavez 2021; Salazar et al. 2020)
 - Allophones of nasal consonants? (Marlett 1992) = post-oralized nasal stops?
 - No nasalized vowels after these segments
 - Hypervoicing? (Iversons & Salmons 1996)
 - Not clusters
 - Pointing at post-stopped nasals (Di Canio et al. 2019, on Yoloxóchitl Mixtec)
 - Same duration as other consonants, longer duration of the nasal closure, sometimes no oral duration but still release
 - Also evidence from resyllabification in play language in Zenzontepec Chatino (Campbell 2020)

Implications

- If treating these as **post-oralized nasal stops** (Marlett 1992) → are we saying they are phonologically different in lexical items than as a result of morphological process?
 - Main argument: Mixtec distinguishes between oral and nasal vowels; Vowels after nasal stops are nasal; to maintain a difference between oral and nasal vowels = post-oralization of the nasal stops
 - This holds for Mixtepec Mixtec as it seemed to hold for Yoloxóchitl Mixtec (Di Canio et al. 2019), BUT ONLY in lexical prenasalization.
 - Words with morphological prenasalization do not have this limitation:
 - *Tiin* [tʰiː⁴⁴] ‘IPFV.grab’ vs. *ntiin* [n^htʰiː¹³³] ‘PFV.grab’
 - Synchronically, however, these are not perceived as different sounds by speakers, and there seems to be no reason to overcomplicate the analysis.

Implications

- If treating this prenasalization as **hypervoicing** (Iversons & Salmons 1996) → there is a clear difference between the stops (average voicing over 71% in our study) and the affricates (average voicing below 50% in our study, but a lot of interspeaker variation)
 - This could be explained articulatorily
 - However, just as pointed out for Yoloxóchitl Mixtec (Di Canio et al. 2019), for the stop, the nasal closure is longer than the oral closure
- If treating them as **clusters** → Are they longer than other consonants? + there were some tokens that had no nasal closure
 - In Yoloxóchitl Mixtec (Di Canio et al. 2019) these were found to be equally long or even shorter (in word-medial position) than other consonants. I suspect the same is true for Mixtepec Mixtec.

Implications

- Regardless,
 - Stops and affricates behave very differently
 - Not really surprising
 - Significant difference in the duration of the nasal closure in the prenasalized alveolar stop
 - How do we explain this?
 - Is this because of historical reasons (aka, morphological prenasalization likely coming from a cluster)?
 - Is this an instance of morphology influencing phonetic realization?



¡Tá tsà'vǐ-kue-ní ñàà cháa
sò'o-ní!
