San Martín Peras Mixtec

Ben Eischens
University of California Los Angeles
beischens@ucla.edu

Andrew A. Hedding
University of Washington
ahedding@uw.edu

San Martín Peras Mixtec

San Martín Peras Mixtec (autonym: *Tu'un Nta'vi* [tṵ̀ʔṹ ⁿtáʔβī] or *Tu'un Savi* [tṵ̀ʔṹ sàʰβǐ]) is an Otomanguean language spoken by roughly 11,500 people in the municipality of San Martín Peras, in Oaxaca, Mexico (Instituto Nacional de Estadística y Geografía, 2020). The municipality is in the district of Juxtlahuaca, bordering the state of Guerrero. As of 2020, approximately 97% of the population of the municipality over three years old is a speaker of an Indigenous language. Of those that speak an Indigenous language, approximately 60% also speak Spanish, meaning that around 37% of the total population is monolingual in Mixtec (Instituto Nacional de Estadística y Geografía, 2020). Despite these figures, it is difficult to estimate the total number of native speakers of the language, as many community members have migrated to other parts of Mexico and the United States, especially to several towns in California (Mendoza, 2020).

San Martín Peras Mixtec is part of the Otomanguean language family. It forms part of the Eastern Otomanguean branch, the Amuzgo-Mixtecan subgroup and the Mixtecan major subgroup (Campbell, 2017). There is no consensus on the number of distinct varieties of Mixtec languages. San Martín Peras Mixtec is classified by Josserand (1983) as part of the Southern Baja dialect region, one of the twelve major dialect groups that she defines. Ethnologue considers San Martín Peras Mixtec to be part of the Western Juxtlahuaca variety (ISO 639-3: JMX), one of 52 distinct varieties that has been assigned an ISO code (Eberhard, Simons & Fenning, 2020). A recent Bayesian phylogenetic analysis of Mixtecan languages identifies 23 distinct subgroups and classifies San Martín Peras Mixtec as being a part of group 7.3 (Auderset et al, 2023). Finally, the Mexican government recognizes 80 varieties of Mixtec and considers residents of San Martín Peras and some neighboring municipalities to speak To'on Savi del Oeste (Instituto Nacional de Lenguas Indígenas, 2008). According to INALI, given that: (i) more than 25% of the speakers of To'on Savi del Oeste are between the ages of 5 and 14; (ii) there are more than 1000 total speakers of To'on Savi del Oeste; (iii) To'on Savi del Oeste is spoken in more than 50 communities; the language is not considered to be at immediate risk of language loss (Embriz Osorio & Zamora Alarcón, 2012). However, there is clear phonological variation across the 83 different communities that speak To'on Savi del Oeste. For instance, the Mixtec spoken San Martín Peras has contrastive breathy phonation (see section 3), while the Mixtec spoken in neighboring Coicoyán de las Flores does not (Beatham and Beatham, 2019). Moreover, as migration and increased connectedness with other communities have expanded the number of

young people who primarily speak Spanish in San Martín Peras, there is reason to be concerned about the long term longevity of the specific variety of the language spoken in the municipality.

Suffice it to say that the term "San Martin Peras Mixtec" should be interpreted as an umbrella term that primarily provides a geographic description of where most speakers reside. Throughout this Illustration, we abbreviate the language name as "SMPM."

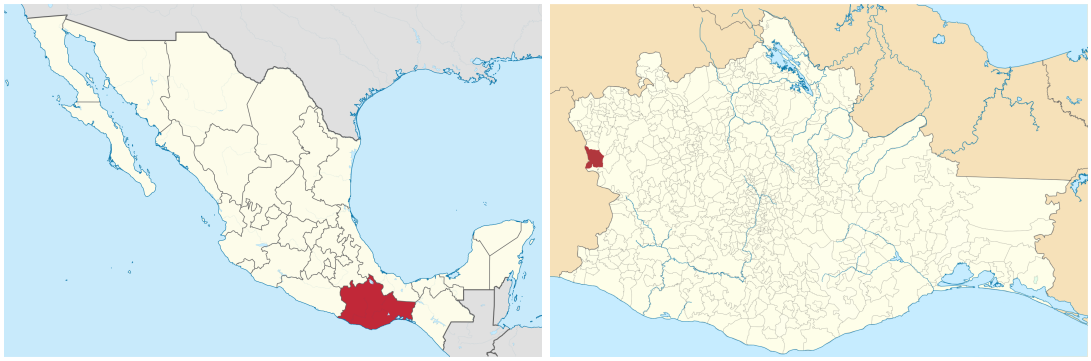


Figure 1: The State of Oaxaca (left) and the municipality of San Martín Peras (right)

This work adds to a substantial list of phonological and phonetic studies of the sound systems of Mixtec languages, going back to the mid-20th century (Pankratz and Pike, 1967; Pike and Small, 1974; North and Shields, 1977; Josserand, 1983; Marlett, 1992; Macaulay and Salmons, 1995; Iverson and Salmons, 1996; Gerfen and Baker, 2005; Daly and Hyman, 2007; Gerfen, 2013; McKendry, 2013; Herrera Zendejas, 2014; DiCanio et al., 2014; Becerra Roldán, 2015; Mendoza, 2016; León-Vázquez, 2017; DiCanio et al., 2018; Peters, 2018; Becerra Roldán, 2019; Penner, 2019; Rueda Chaves, 2019; DiCanio et al., 2020; Peters and Mendoza, 2020; DiCanio et al., 2021; Cortés et al., 2022; Eischens, 2022; Stremel, 2022; Uchihara and Mendoza, 2022; Eischens, 2023; Caballero et al., to appear; among others).

The data in this Illustration were collected from 3 speakers of the language. All audio recordings illustrating the phones of the language come from one 71-year-old male speaker (JGO) born in the community of Ahuejutla, where he has lived his whole life. Ahuejutla, a town of roughly 1200 inhabitants, is approximately 10 miles from the municipal center of San Martín Peras. The recording of the retelling of the North Wind and the Sun story is from a 52-year-old woman (NGC) who is originally from Ahuejutla and who has lived in California for approximately 20 years. The data used for fast Fourier transforms and center of gravity measurements for fricatives, voice onset time (VOT) in plain and prenasalized consonants, vowel

formants, strength of excitation (SoE) for non-modal vowels, and tone plots come from a task in which NGC and one additional female speaker in her 50s (RDC) produced target words in a carrier sentence at normal and slow rates of speech. This data is also accompanied by representative audio recordings. RDC is originally from the town of San Martín Peras and has lived in California for approximately 20 years. To our knowledge, the varieties of Mixtec spoken in Ahuejutla and San Martín Peras are almost identical, though some small lexical differences may exist. We know of no tonal distinctions between the two towns, though NGC and RDC occasionally have minor differences in their pronunciations of affricates (e.g., [t͡sìkūʃí] vs. [t͡ʃìkūʃí] for 'worm').

Recordings were made using a Zoom H5 Handy Recorder (16-bit quantization rate and 44.1 kHz sampling frequency) and a Nady HM-45U headset microphone. Recordings were spliced using Audacity. Wherever possible, individual examples of words were spliced out of carrier phrases in which the target word was immediately preceded and followed by a mid-tone to avoid known effects of tone sandhi. The nature of the elicitation task with JGO required the use of a number of distinct carrier phrases, and the carrier phrase corresponding to each example in the manuscript is listed in the appendix. In addition to the audio files for each individual word produced, audio files are also available for each target word in its carrier phrase, with the filenames for these recordings ending in '_CP'. Carrier phrases for audio recordings from NGC and RDC are not provided in the appendix because the same carrier phrase was used in all cases.

1 Consonants

The following table illustrates the basic phonemic contrasts in the language (Peters, 2018; Mendoza, 2020; Eischens, 2022).[1]

[1] Note that [β], the voiced bilabial fricative, is listed as an approximant due to its phonological patterning in the language. Additionally, pre-nasalization is indicated by a superscript [n] following Keating et al. (2019). As we note below, there is some disagreement between these references on the phonological status of prenasalized obstruents. However, these sources broadly agree on the basic sounds of the language. See discussion under the "obstruents" heading below.

	Bilabial	Alveolar	Palatalized Alveolar	Postalveolar	Palatal	Velar	Labialized velar	Palatalized Velar
Plosive	p / ᵐp	t / ⁿt	tʲ / ⁿtʲ			k / ᵑk	kʷ	kʲ
Nasal	m	n			ɲ			
Tap		ɾ						
Fricative		s	sʲ	ʃ				
Affricate			t͡sʲ / ⁿt͡sʲ	t͡ʃ / ᶮt͡ʃ				
Approximant	β				j			
Lateral approximant		l						

[páʰtò]	*pátò*	‘duck’ (cf. Spanish *pato*)[2]
[ᵐpáà]	*mpáà*	‘godfather of one’s child, or father of one’s godchild’ (cf. Spanish *compadre*)
[táʰtà]	*tátà*	‘father’ or ‘sir’
[ⁿtáʔβī]	*ntá’vi*	‘poor’
[tʲàʰtá]	*tiàtá*	‘a type of oak tree’
[ⁿtʲaʔmĭ̌]	*ntia’mĭ*	‘radish’
[kàʰkū]	*kàku*	‘was born’
[ĩĩᵑ̃kà]	*iinkà*	‘other’
[kʷáʔà]	*kuá’à*	‘red’
[kʲáʔmĭ]	*kiá’mĭ*	‘(a type of) squash’

[2] We provide examples in both the IPA and the Mixtec orthography advanced by Ve’e Tu’un Savi (The Mixtec Language Academy) (Norma de escritura del Tu’un Savi (idioma mixteco), 2022), with minor alterations to the way that tones are marked. Specifically, in the orthographic examples in this paper, we represent low tones with a grave accent (ò), rising tones with a caron (ǒ), and falling tones with a circumflex (ô). We follow the other orthographic recommendations of Ve’e Tu’un Savi. Important among these are that vowel nasalization, usually marked with a coda [n], is not marked in vowels following nasal consonants because these vowels are always nasalized. It should also be noted that this orthography is still not in wide use in Ahuejutla. In fact, there is no standardized orthography in use in the town, so there remains considerable variation in the ways that speakers write the language. For example, some speakers prefer to write /ʃ/ using *sh* and some prefer to write /t͡s/ using *z*. In the IPA examples, we represent bimoraic vowels with a series of identical vowels rather than the standard long vowel marking. Because the tone-bearing unit in Mixtec is the mora, this allows a more straightforward representation of the tonal contrasts available on bimoraic vowels.

[málì]	*málì*	‘godmother of one’s child, or mother of one’s godchild’ (cf. Spanish *comadre*)
[nánà]	*nánà*	‘mother’ or ‘madam’
[ɲānī]	*ñani*	‘brother (of a male)’[3]
[ɾà lōˀō]	*rà lo’o*	‘boy’ (literally "small male")
[sâˀmǎ̋]	*sâ’mă*	‘tortilla cloth’
[sʲâˀă]	*siâ’ă*	‘Tecomaxtlahuaca’ (a town)
[ʃàà]	*xàà*	‘chin’
[t͡sʲă̋ˀà]	*tsiă’à*	‘jug/pitcher’
[ⁿt͡sʲàˀβă]	*ntsià’vă*	‘toothless’
[t͡ʃáákà]	*tyáakà*	‘more/most’
[ⁿt͡ʃī̋ʰʃǐ]	*ntyixǐ*	‘corncob’
[βálí]	*válí*	‘small (pl)’
[jāā]	*yaa*	‘white’
[lāⁿt͡ʃī]	*lantyi*	‘lamb’

Obstruents

San Martín Peras Mixtec has 3 phonemic plosive consonants: /p/, /t/, and /k/. However, /p/ is restricted to the loan vocabulary of the language and is not found in non-loans. In some environments, /k/ is pronounced allophonically as [ɡ]. The environment that most commonly seems to license this type of allophonic variation is non-root-medial position between two vowels. For example, in the word /tʃáa=ka/ (`more, most’), the /k/ is usually pronounced as a [ɡ] or a [ɣ]. This voicing process seems to be subject to both interspeaker and inter-utterance variation. We note, however, that this allophonic voicing never seems to happen root-internally and only seems to occur in multi-morphemic words.

Within plosives, SMPM also has contrastive secondary articulations. For instance, both /t/ and /k/ can be contrastively palatalized, e.g., *tiàtá* [tʲàʰtá] ‘type of oak tree’ and *kiá’mĭ* [kʲáˀmĭ] ‘(a type of) squash’. In addition, /k/ can be contrastively labialized, e.g. *kuá’à* [kʷáˀà] ‘red’. There are two principal reasons to consider palatalization and labialization secondary articulations, rather than consonants in and of themselves. The first is that [w] is not an

[3] Some speakers pronounce the second nasal in this word as palatal.

independent consonant in SMPM and only appears in conjunction with [k]. The second is that there are distributional differences between palatalized consonants and the palatal glide [j]. Specifically, [j] may precede the high front vowel [i], as in *nà yivĭ* [nǎ jīβĭ] ‘person’ (20d), but palatalized consonants never precede the vowel [i]. That is, a hypothetical word like [kʲīβĭ] does not exist. Because palatalized consonants are more distributionally restricted than consonant [j], it is unlikely that palatalization is actually an instance of the consonant [j].

San Martín Peras Mixtec has 2 voiceless fricatives: /s/ and /ʃ/. In a small number of words, /s/ is contrastively palatalized, e.g., *siâ’ă* [sʲâʔă] ‘Tecomaxtlahuaca (a town)’. Fast Fourier transforms (FFTs) of individual tokens of [s] and [ʃ] from NGC are shown in Figure 2 below, along with average center of gravity (CoG) measurements for both fricatives for NGC and RDC. In all cases, measurements were taken from a 50ms window centered on the peak of noise intensity for the fricative. The data were measured and visualized in Praat (Boersma and Weenink, 2020).

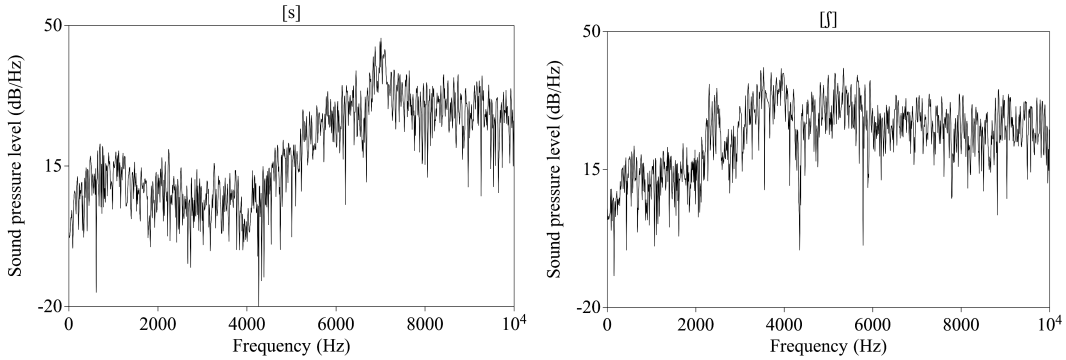


Figure 2: FFTs for [s] (left) and [ʃ] (right) produced by NGC. Examples taken from the words [sânǎ] (‘crazy’) and [ʃânǔ] (‘cigarette’)

	RDC	NGC
CoG mean (SD) [s]	7669 (596) n = 34	8258 (792) n = 34
CoG mean (SD) [ʃ]	4610 (550) n = 21	5394 (527) n = 18

Table 1: Mean and standard deviation CoG (Hz) measurements by speaker and fricative.

In addition to fricatives, SMPM has 2 voiceless affricate consonants, /t͡s/ and /t͡ʃ/, which each contrast with a prenasalized counterpart. The contrast between /t͡s/ and /ⁿt͡s/ can be seen in the consonant word-list, and the contrast between /t͡ʃ/ and /ʲⁿt͡ʃ/ can be seen in (1) below.

(1) a. [t͡ʃiʰkí] tyìkí ‘prickly pear fruit’
b. [ʲⁿt͡ʃɪ̃ʰʃĭ] ntyixĭ ‘corncob’

Unlike the plosive series, these affricates are not contrastively palatalized. /t͡ʃ/ and /ʲⁿt͡ʃ/ never occur with palatalization, and [t͡s] and [ⁿt͡s] are predictably palatalized before all vowels except /i/ (Stremel, 2022), as shown in example (2) below. Given this predictability, we assume that /t͡sʲ/ and /ⁿt͡sʲ/ are underlyingly palatalized and are allophonically depalatalized before a high front vowel. Impressionistically, /t͡s/ sounds like palatalized [t͡ɕ] on some productions before [i]. We leave for future research whether this is a phonological alternation or a coarticulatory effect.

(2) a. [ⁿt͡sîʰkă] *ntsîkă* ‘banana’
b. [ⁿt͡sʲèʰě] *ntsièjě* ‘strength’
c. [ⁿt͡sʲàˀβă] *ntsià’vă* ‘toothless’
d. [ⁿt͡sʲôʰǒ] *ntsiôjǒ* ‘moonlight’
e. [tsʲṵ̀ˀṵ́] *tsiù’ún* ‘venomous spider’

There is no plain voicing distinction in San Martín Peras Mixtec. However, all stops and affricates except /kʲ/ and /kʷ/ contrast with prenasalized versions. Throughout the paper, we will refer to this as a contrast between *plain* and *prenasalized* consonants. The phonological status of prenasalized consonants in the language is a point of debate. Peters (2018:13) analyzes them as sequences of a nasal consonant and a stop, noting that they syllabify as a coda in words like *lantyi* [lāⁿt͡ʃɪ̃] ‘lamb’. On the other hand, Eischens (2022) analyzes them as complex segments, since onset consonant clusters are banned in the language, but pre-nasalized stops and affricates may occur word-initially (see consonant examples above). We adopt Eischens’ (2022) analysis, given the general ban on consonant clusters and the fact that a word-initial nasal + obstruent cluster would violate the sonority sequencing principle (e.g., Kiparsky, 1979; also noted in Iverson and Salmons, 1996:166), making it an unlikely onset. We also consider phonotactic restrictions on the distribution of prenasalized consonants as evidence that they are single segments. That is, they may only be followed by oral vowels, never nasal vowels (see section 2

below). This is unexpected if they are, in fact, a sequence of a nasal and a plain consonant, since plain consonants may be followed by both oral and nasal vowels.

The unary analysis of prenasalized consonants is somewhat complicated, though, by the distribution of the prenasalized consonant [ᵑk] in the language, which only occurs root-medially. We know of only two monomorphemic words with [ᵑk],[4] both shown in (3) below.

(3) a. [t͡ʃĩ²ᵑkì] *tyi'nkì* 'acorn'
b. [līᵑkō] *linko* 'bud of the flower of the maguey cactus'

There are three possible ways to account for this restricted distribution. First, one could adopt Peters' (2018:13) analysis of pre-nasalized consonants as a bi-consonantal series of a nasal + plosive. Under this analysis, the [ᵑ] in the preceding examples would be a nasal coda of the first syllable. However, given that there are no other codas in the language, and that we analyze pre-nasalized stops in other cases as complex segments, we do not adopt this analysis. The second possible analysis is to suppose that [ᵑk] may only occur in loan words, as is the case for [p] and [ᵐp]. Indeed, *linko* is plausibly related to the Mexican Spanish word *gualumbos* (also spelled *golumbos* or *hualumbos*), which also refers to the edible flowers of the maguey cactus (Piedra-Malagón et al., 2022).[5] Notably, these terms all involve a lateral followed by a nasal + stop sequence, like the word in SMPM. Given the rarity of pre-nasalized labial obstruents in San Martín Peras Mixtec, it is possible that the labial place of articulation was borrowed as a dorsal. However, we currently have no evidence that *tyi'nkì* is a loan word. The third possible analysis—and the one that we tentatively adopt here—is that [ᵑk] is restricted phonotactically to root-medial position. This possibility is bolstered by the fact that [ᵑk] only occurs in root-medial position in other Mixtec varieties as well, such as Chalcatongo (Iverson and Salmons, 1996), Alcozauca, (Mendoza Ruiz, 2016) and Yucuquimi de Ocampo Mixtec (Leon-Vázquez, 2017).

In what remains of this section, we qualitatively and quantitatively illustrate the phonetic characteristics of plain and prenasalized stops and affricates, focusing on voice onset time (VOT) and the internal structure of prenasalized stops and affricates. For plain stops, VOT was

[4] A third word with a prenasalized word-medial [ᵑk] (ĩĩᵑkà 'other') is likely derived from the numeral one (ĩĩ) and an "additive" clitic (=ka) (though see Mantenuto, 2020 for a more nuanced discussion of its meaning).
[5] *Gualumbo* is itself a borrowing from the Otomí word *uadombo* (Bravo Vargas, 2014).

measured from the release burst until the beginning of periodic voicing associated with the following vowel. For plain affricates, VOT was measured from the offset of frication associated with the affricate, marked by cessation of high-frequency aperiodic noise in the spectrogram, to the onset of periodic voicing (Abramson and Whalen, 2017). Measurements were taken from stops and affricates with no palatalization or labialization. Examples of VOT measures are shown below in Figure 3, which are taken from the words [tá^{h}tě] (‘sir/father’), [kîǹǐ] (‘pig’), [t͡si^{h}kà] (‘grasshopper’), and [t͡ʃi^{n}t͡ʃi] (‘cricket’), respectively.

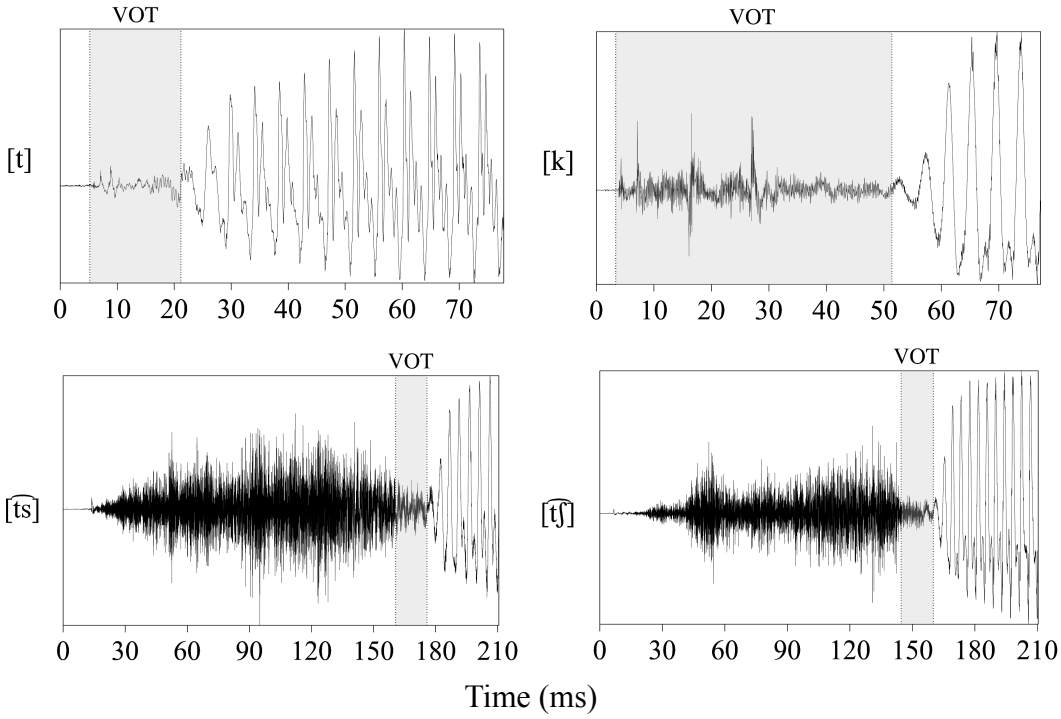


Figure 3: Representative examples of plain stops and affricates. Shaded portion shows VOT.

While VOT is an informative measure of voicing for the plain stops and affricates, it does not capture the internally-complex structure of prenasalized stops and affricates. These consonants are characterized by a sequence of periodically-voiced prenasalization, followed by a period of voicelessness during the oral closure and/or frication, and there is also almost always a positive VOT measured from the stop burst or cessation of frication. In many cases, weak voicing persisted from prenasalization into a portion of the stop closure, though this voicing was never strong and almost always ceased prior to the stop burst (c.f. Cortés et al. 2022). Figure 4 below illustrates this sequencing of prenasalization (PN), weak voicing (WV), no voicing (NV), and a positive VOT. The examples are taken from the words [^{n}tiβi] (‘beautiful’), [lí^{ŋ}ko] (‘fruit of the

maguey cactus'), [ⁿt͡sîʰkà] ('wide'), and [ᶮt͡ʃíʰʃí] ('corn cob'), respectively. The example illustrating [ᵑk] is root-medial, since this segment never occurs root-initially.

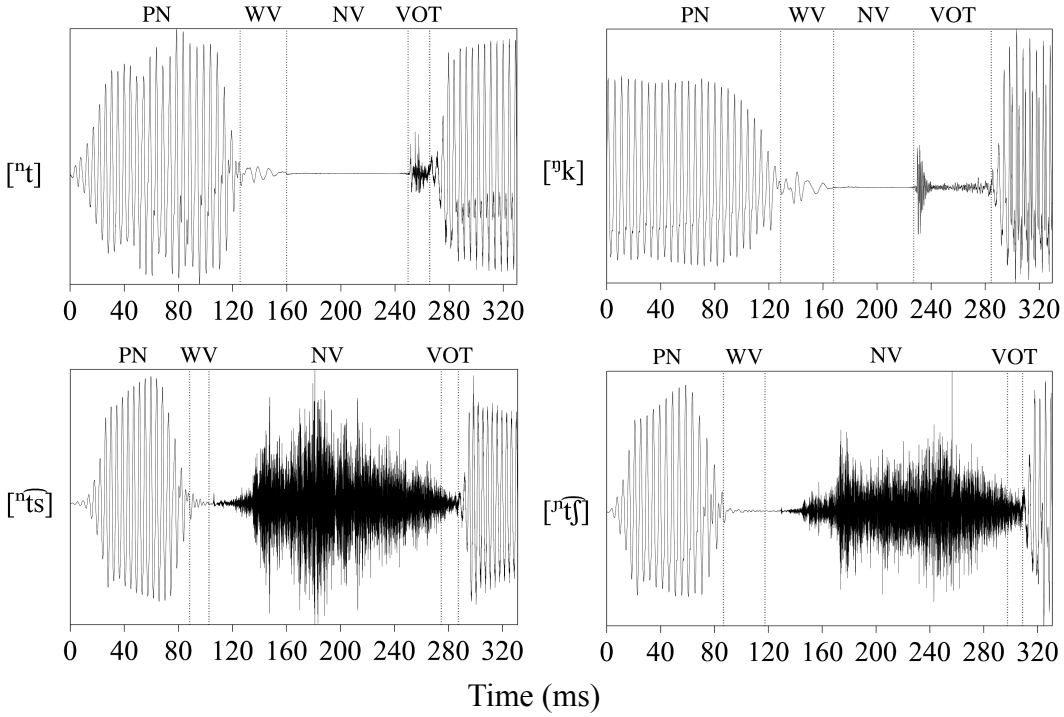


Figure 4: Representative examples of prenasalized stops and affricates. PN = prenasalized, WV = weak voicing, NV = no voicing, VOT = VOT from stop burst or cessation of oral frication

One point of interest is that both plain and prenasalized stops and affricates have a positive VOT, measured either from the stop burst or from the cessation of oral frication. As shown in Figure 5, VOT values roughly line up between the plain and prenasalized versions of a consonant. Note, though, that the VOT values for [ᵑk] are taken from root-medial tokens, since there are no root-initial tokens of [ᵑk], and that [ᶮt͡ʃ] was excluded because of a low number of tokens for analysis (3 per consultant). The data were measured in Praat and illustrated using the ggplot package (Wickham, 2016) in R (R Core Team, 2013).

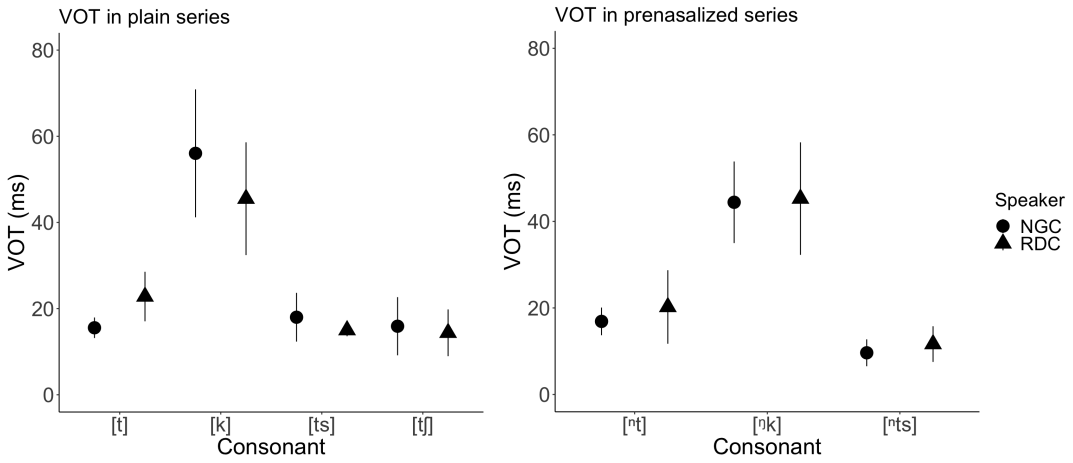


Figure 5: VOT by consonant type and speaker in the series of plain stops and affricates (left) and prenasalized stops and affricates (right). Error bars represent one standard deviation.

NGC [t͡ʃ] = 35, [k] = 51, [t] = 18, [t͡s] = 6, [ⁿt] = 18, [ᵑk] = 12, [ⁿt͡s] = 11.

RDC [t͡ʃ] = 18, [k] = 51, [t] = 15, [t͡s] = 6, [ⁿt] = 18, [ᵑk] = 15, [ⁿt͡s] = 17.

The VOT of [t], [tʃ], and [ts] are all short lag, with averages around 20ms. As is usually the case with backer places of articulation (Lisker and Abramson, 1967), [k] has higher VOT than [t]. However, the difference between [t] and [k] here appears larger than the difference in VOT between [t] and [k] and English (Lisker and Abramson, 1967:6), as well as in other Mixtec languages (DiCanio et al., 2020; Cortés et al., 2022). Notably, [k] also displays more variance than [t]. The higher VOT variance for [k], and potentially the higher average value, may stem from the distinction between speech rates in the production task.

As discussed above, prenasalized stops and affricates are made up of a sequence of prenasalization (PN), weak voicing (WV), and voicelessness (no voicing, NV). The voiceless portion also includes a positive VOT measured from the offset of oral constriction to the beginning of periodic voicing. Figure 6 below shows the duration of each of these subparts of a prenasalized consonant as a proportion of the entire duration of the consonant. Given that the period of voicelessness and positive VOT are two subparts of the voiceless period of the consonant, they are combined under the category “voiceless” here.

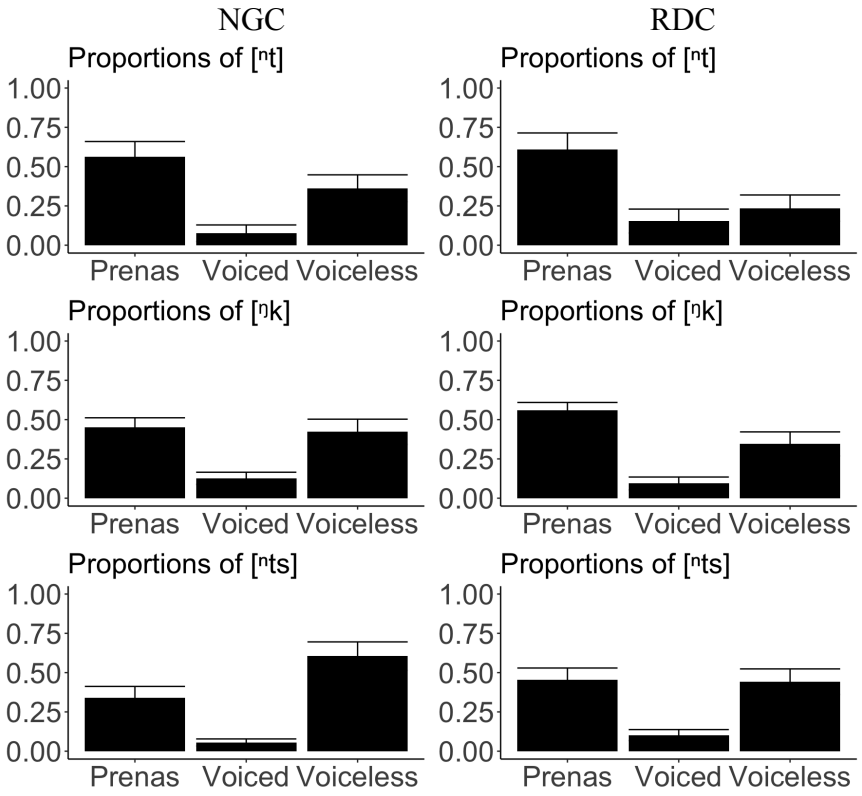


Figure 6: Proportional duration of prenasalization, weak voicing, and voicelessness in prenasalized consonants for NGC (left) and RDC (right). Error bars show one standard deviation.

Before moving on, it is worth noting that a prenasalized stops in a number of Mixtec languages alternate between fully voiced forms and forms with a voiceless interval (see, e.g., Rueda Chaves, 2019:139 and sources therein). In these varieties, the devoiced versions of prenasalized consonants appear in particular phonological environments, such as root-initially, motivating an analysis in which their voicelessness is derived by either a phonological or phonetic process of strengthening. In San Martín Peras Mixtec, however, prenasalized stops and affricates almost always have a voiceless interval regardless of position in the root. Because of this, we do not analyze the voicelessness in prenasalized consonants as derived by a phonological process. It is possible that across-the-board voicelessness in prenasalized consonants is an innovation in the phonological system of SMP Mixtec, though we leave this question to future research.

Sonorants

In addition to its obstruent consonants, San Martín Peras Mixtec has a set of seven sonorant consonants. There are three contrastive nasal consonants with bilabial, alveolar, and palatal

places of articulation, e.g. *málì* [málì] ‘godmother of one’s child, or mother of one’s godchild’, *nánà* [nánà] ‘mother’ or ‘madam’, and *ñani* [ɲānī] ‘brother (of a male)’. In addition, there are three non-nasal approximants in the language: /l/, /β/, and /j/. Finally, there is one voiced alveolar tap /ɾ/. Of these consonants, /ɾ/ is the most clearly restricted. To our knowledge, it only occurs almost exclusively in the clitic pronoun series and function words. [ɾà] is used for human males, [ɾí] is used for animals and round objects, [ɾá] is used for liquids, and [ɾa] is the conjunction *and.* There are very few other native lexical items with /ɾ/ in the language. Of the approximants, /j/ appears to vary most widely in its phonetic realizations, even between productions of the same word in the same context by the same speaker. The examples below, which were all produced by the same consultant in the same carrier phrase, show /j/ realized as an approximant (Figure 7, left), a transition from an approximant into a nearly voiceless fricative (Figure 7, middle), and a voiceless palatal fricative (Figure 7, right). Voiced fricative realizations of /j/ can also be seen in (4a-c).

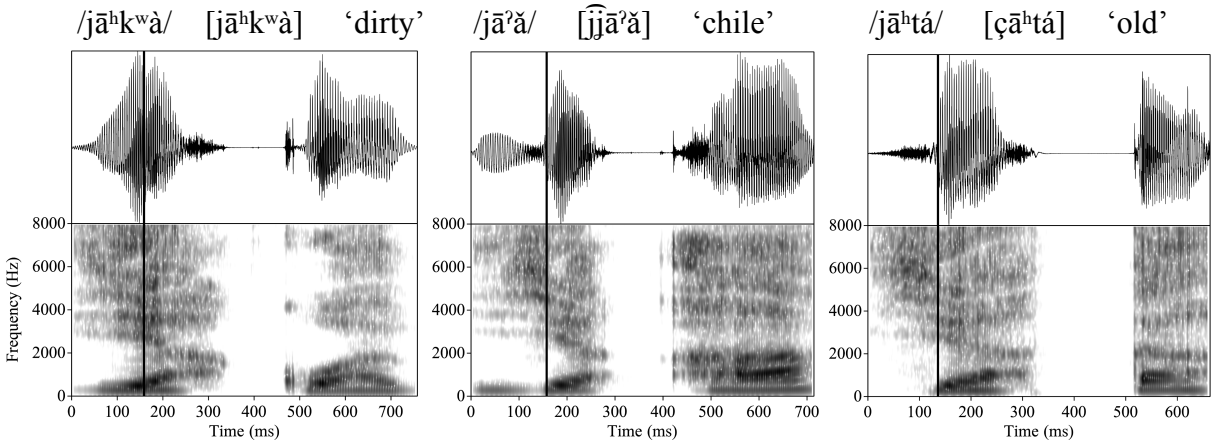


Figure 7: Waveforms and spectrograms showing four distinct realizations of /j/. Vertical black bar represents the offset of the sound.

This variation is not limited to couplet-initial position, unlike cases of fortition in other Mixtec languages (see, e.g., Rueda Chaves, 2019). In couplet-medial position, /j/ can be realized as a glide (Figure 8, left), or as partially voiced and fricated (Figure 8, right). In these examples, the black vertical lines mark the left and right boundary of /j/.

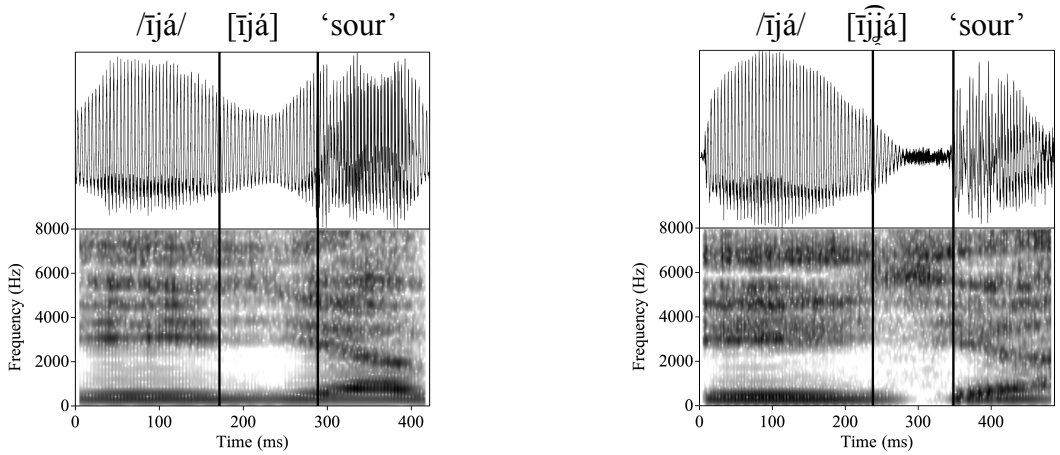
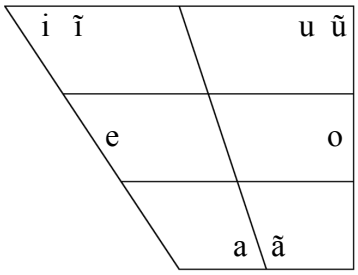


Figure 8: Waveforms and spectrograms showing two distinct realizations of /j/. Vertical black bars represent the onset and end of the sound.

Despite the fricative realizations of /j/, we classify it as an approximant and not a voiced fricative, unlike in other Mixtec varieties (e.g., Cortés et al, 2022). This is because, if it were a fricative, it would be the only phonemically voiced consonant in SMPM's obstruent series. Additionally, like the other approximants (/l/ and /β/), it is only ever followed by oral vowels in non-morphologically-complex contexts.

2 Vowels

San Martín Peras Mixtec has five oral vowels and three contrastively nasal vowels.[6]



[kátaʔă]	kátaʔă	‘fights’
[kéβàʔă]	kévàʔă	‘wins’
[kítaʔă]	kítaʔă	‘accompanies’

[6] Following common phonological practice, we use the symbol [a] to represent the low vowel in SMPM, despite the fact that it is a perceptually central vowel.

[kò-kɑ́ʔà]	kò-ká'àn	'does not talk'
[kúkáʔnù īnī]	kúká'nù ini	'forgives' (lit. 'be big inside')
[kʷáà]	kuáà	'blind'
[kʷɑ́à̰]	kuáàn	'yellow'
[kʷíì]	kuíì	'clear'
[kʷɪ́̀ɪ̀]	kuíìn	'striped'
[kūū]	kuu	'will die'
[kṵ̄ṵ̄]	kuun	'will fall' (used for rain)

While there are only three contrastive nasal vowels, we note that many speakers pronounce nasal [ũ] lower in the vowel space than oral [u], leading it to sound like [õ]. However, we know of no examples of [ũ] contrasting with [õ]. Nasal vowels only contrast after plain stops and affricates. Vowels are predictably nasal when following nasal consonants, and oral when following prenasalized consonants and approximants.

Below are two plots showing the average formant values for SMPM's five oral vowels and 3 nasal vowels for NGC and RDC. The graphs represent root-final vowels, where there is no phonation type contrast.

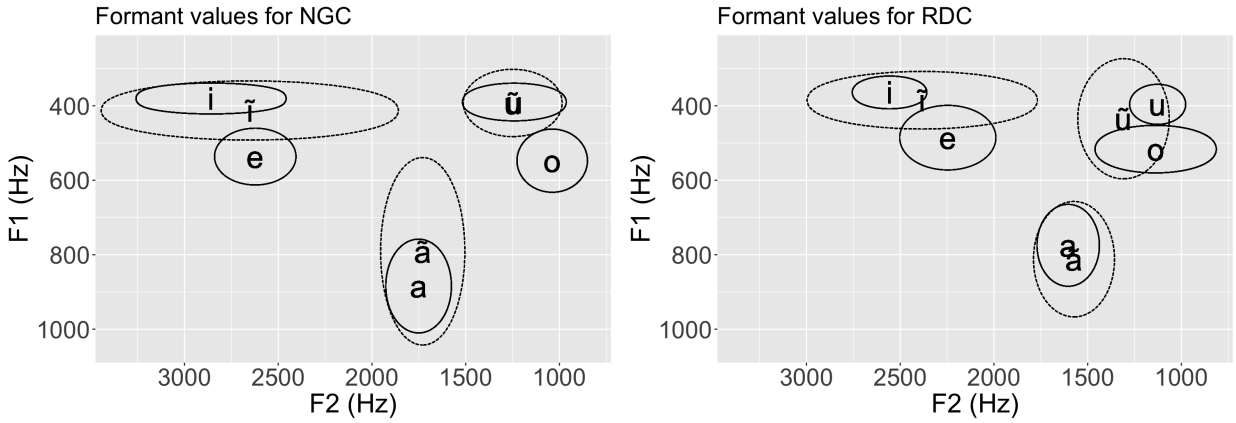


Figure 9: Plots of average formant values (Hz) for NGC and RDC. Ellipses show one standard deviation around the mean. Solid lines represent oral vowels; dotted lines represent nasal vowels.

Number of tokens for NGC: 69 [a], 36 [ã], 21 [e], 77 [ĩ], 88 [i], 84 [o], 32 [ũ], 30 [u].

Number of tokens for RDC: 75 [a], 39 [ã], 18 [e], 78 [ĩ], 84 [i], 87 [o], 33 [ũ], 36 [u].

3 Prosodic Features

Syllable structure and the couplet

In arguably all Mixtec languages, words are organized around a bimoraic unit known in the Mixtecanist literature as the 'couplet' (Pike, 1948; see Penner, 2019 for a comprehensive overview). This is the case in SMPM, where lexical roots are minimally, and usually maximally, bimoraic (Peters, 2018; Eischens, 2022). They are made up of two monomoraic short vowels, or one bimoraic long vowel, and there is a ban on coda consonants, which gives rise to the canonical root shapes of (C)VCV and (C)VV. The couplet in SMPM is the locus of phonation contrasts and tonal melodies, which are the subjects of the following sections. Given that the bimoraic lexical root and the couplet are usually interchangeable, we use the more generic term 'root' throughout.

Phonation Type

Across Mixtec languages, the glottal stop [ʔ] patterns differently from other consonants, as outlined in Macaulay and Salmons (1995). For example, it is usually the only licit coda consonant (e.g., *ka'nu* [káʔnṵ̀] 'big'), and it never occurs phonemically in root-initial or root-final position in most Mixtec languages (though see Pankratz and Pike, 1967 and Herrera Zendejas, 2014 for Ayutla Mixtec; and Towne, 2011 for Zacatepec Mixtec). Additionally, there may only be one glottal stop per root, and when it occurs between two vowels in a mono-morphemic context, the vowels always match in quality and nasalization. In addition, in some Mixtec languages, CVʔV and CVV roots act as a natural class regarding tone sandhi processes, to the exclusion of roots with a medial oral consonant (Macaulay and Salmons, 1995:58). Because of these characteristics, many researchers have adopted the hypothesis that the glottal stop in Mixtec languages is not a consonant proper, but rather a supra-segmental feature of the vowel, root, or word (Macaulay and Salmons, 1995; Gerfen, 2013; McKendry, 2013; Becerra Roldán, 2015; León Vázquez, 2017; Penner, 2019; Rueda Chaves, 2019, Cortés et al., 2023, a.o., but see Herrera Zendejas, 2014 for treatment of glottal stop as a consonant). Because most of these characteristics also hold of the glottal stop in SMPM, we follow the trend in the Mixtec literature and represent glottal stop as a contrastive phonation type, referred to throughout as glottalization. It may surface root-medially between two homorganic vowels or before root-medial sonorants and prenasalized consonants. In addition, SMPM makes use of a

contrastive [h] with the same phonotactic distribution as [ʔ], occurring root-medially between two homorganic vowels or before root-medial sonorants and prenasalized consonants.[7] [h] is uncommon in other Mixtec languages, and thus is likely a relatively recent innovation in SMPM (Peters, 2018). Because of its phonotactic similarity to [ʔ], we also analyze [h] as a contrastive phonation type, referred to herein as breathy phonation. Following Ve'e Tu'un Savi, we represent glottalization orthographically as an apostrophe. Due to the rarity of breathy phonation in Mixtec languages, there is no orthographic convention to represent it proposed by Ve'e Tu'un Savi. In what follows, we choose to represent breathiness orthographically as *j*.[8] All five phonemic oral vowels can be contrastively breathy (b examples) and glottalized (c examples) (4-8).

(4)	a. [jaa]	*yaa*	'white'
	b. [jâ^{h}ă]	*yájă*	'tongue'
	c. [ja^{ʔ}ă]	*ya'ă*	'chile'
(5)	a. [kú^{n}t͡s^{j}ēē]	*kúntsiee*	'puts up with'
	b. [^{n}t͡s^{j}è^{h}ě]	*ntsièjě*	'strength'
	c. [nt͡s^{j}é^{ʔ}é]	*ntsié'é*	'ground bean soup'
(6)	a. [k^{w}íì]	*kuíì*	'clear (Adj)'
	b. [k^{w}ǐ^{h}ǐ]	*kuíjǐ*	'green'
	c. [k^{w}ì^{ʔ}ī]	*kuì'i*	'fruit'
(7)	a. [kōō]	*koo*	'welcome'
	b. [kò^{h}ǒ]	*kòjǒ*	'snake'
	c. [kò^{ʔ}ǒ]	*kò'ǒ*	'plate' or 'bowl'
(8)	a. [^{n}túù]	*ntúù*	'a small black flying insect with a white face'
	b. [^{n}tù^{h}ǔ]	*ntùjǔ*	'large, hard fruit seeds'
	c. [^{n}tū^{ʔ}ú]	*ntu'ú*	'fat (Adj)'

In addition, all three nasalized vowels can be contrastively breathy and glottalized (9-11).

[7] We know of one exception to this, which is that the post-nominal demonstrative [háan] (`that') begins with [h].
[8] The choice of *j* is meant to mirror the Spanish grapheme *j*, which is realized as [x] and [h] in dialects of Mexican Spanish (Canfield, 1981).

(9)	a. [kʷáǎ̀]	*kuáàn*	‘yellow’
	b. [kʷâʰǎ]	*kuâjǎn*	‘unmarried’
	c. [kwã̄ʔǎ]	*kua’ǎn*	‘leaves’ (V.)
(10)	a. [ĩ́ĩ́]	*íín*	‘hail’
	b. [ĩ̂ʰĩ̌]	*íjǐn*	‘skin’ or ‘leather’
	c. [ĩ́ʔĩ̀]	*í’ìn*	‘temazcal (a type of sweat lodge)’
(11)	a. [ɲū̄ù]	*ñuù*	‘town’
	b. [ɲùʰŭ]	*ñùjǔ*	‘palm plant’
	c. [ɲū̄ʔù]	*ñu’ù*	‘fire’

[h] is allophonically nasalized between nasal vowels, and it surfaces as the palatal fricative [ç] after the high front vowel [i] (Eischens, 2022). In addition to contrastive breathiness, vowels are predictably aspirated when they precede a plain, root-medial consonant.[9] Examples of this allophonic variation can be found throughout this Illustration, including, for example, *tátà* [táʰtà] ‘father/sir’ and *tyìkí* [t͡ʃìʰkí] ‘prickly pear fruit’. However, plain consonants that are non-root-medial (that is, word-medial consonants that surface after a prefix) are not pre-aspirated. For example, *kò-ká’àn* [kò-kǎʔà] ‘does not talk’ and *kúká’nù ini* [kúkáʔnù īnī] ‘to forgive (lit. to be big inside)’. At present, it is unclear whether allophonic [h] is best understood as preaspiration of plain consonants, as is found in Alcozauca (Mendoza Ruiz, 2016) and Ayutla Mixtec (Pankratz and Pike, 1967), or as allophonic breathy phonation similar to the allophonic glottalization found in Coatzospan Mixtec (Gerfen, 2013). On the one hand, both allophonic [h] and contrastive [h] are restricted to root-medial position, lending support to the view that they both constitute breathy phonation. On the other hand, vowels preceding allophonic [h] may host any tone, while vowels preceding contrastive [h] may only host a subset of the possible tones, as discussed in the section on lexical tone. This latter point lends support to a view of allophonic and contrastive [h] as phonologically distinct. Given the contradictory evidence, we leave a definitive answer to this question for future research.

In what follows of this section, we briefly discuss the phonetic realization of glottalization, breathiness, and preaspiration in SMPM. As is the case in many Mixtec languages (Pike and Small, 1974:122-124; Macaulay, 1996:42; Gerfen and Baker, 2005; Herrera Zendejas,

[9] Note that we do not represent preaspiration orthographically because it is predictable allophonic variation.

2014:72-74; Becerra Roldán, 2019:112-116; Penner, 2019:254; Cortés et al., 2022:11-14), the articulation of glottalization varies greatly both within and between speakers. Though the most common realization of glottalization in the examples in Figure 10 involves full glottal closure, glottalization is often produced with creaky voice or periodic voicing accompanied by amplitude and/or pitch modulations (see Eischens, 2022 for more details on the variable realization of glottalization). Breathiness is most commonly realized as a short period of breathy voicing followed by voiceless aspiration, as shown in Figure 11. The difference between breathiness and glottalization can be seen by comparing the examples in Figure 11 to their corresponding (near-)minimal pairs in Figure 10.

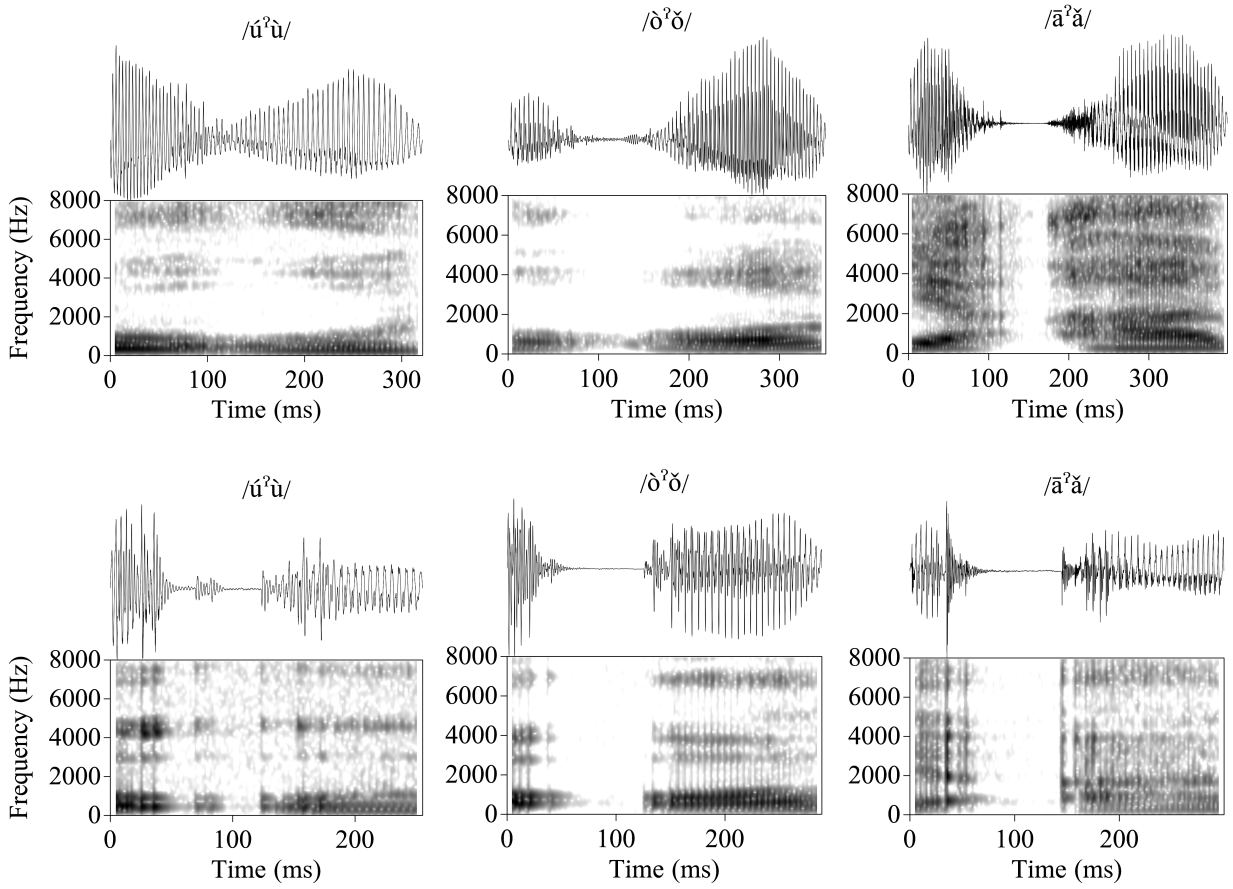


Figure 10: Individual tokens of glottalized vowels from the words /kúʔù/ (‘sick’), /kòʔǒ/ (‘plate’), and /jāʔǎ/ (‘chile pepper’) from NGC (top) and RDC (bottom).

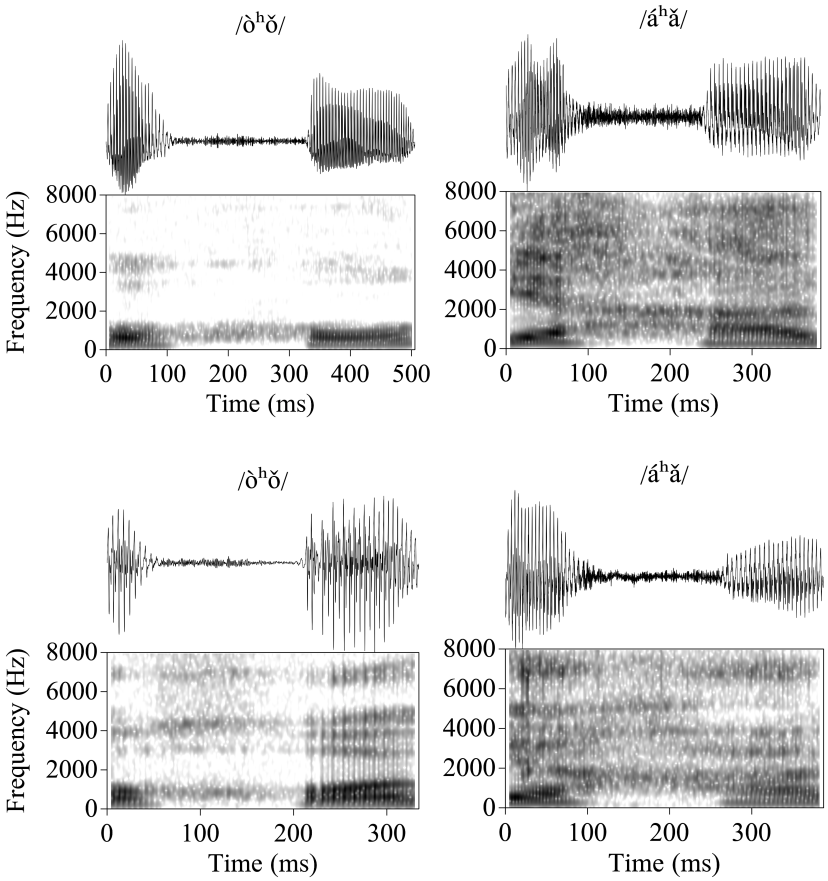


Figure 11: Individual tokens of breathy vowels from the words /kòʰǒ/ (‘snake’) and /jáʰǎ/ (‘tongue’) from NGC (top) and RDC (bottom).

Preaspiration is also most commonly realized as an interval of breathy voicing followed by voiceless frication, though the frication has a much shorter duration than in breathy vowels. This can be seen below in Figure 12, which shows a vowel followed by a preaspirated [t].

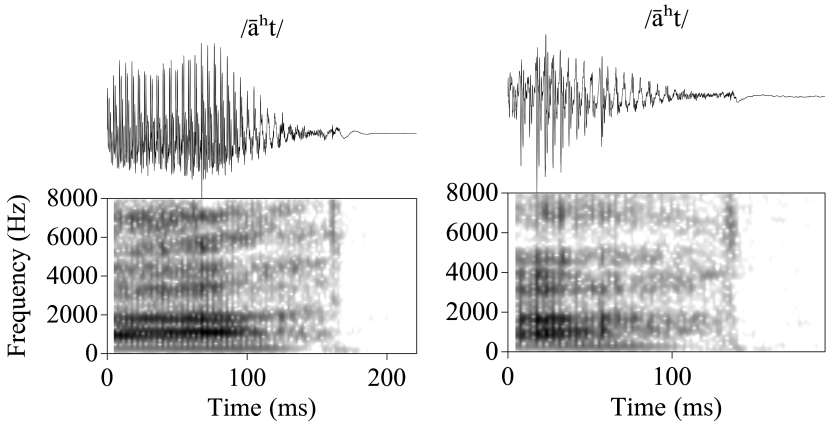


Figure 12: Individual tokens of vowels followed by preaspiration from the word /āʰtū/ (‘bitter’) from NGC (left) and RDC (right).

To quantitatively examine the phonetic realization of glottalization, breathiness, and preaspiration, we calculated Strength of Excitation (SoE) across phonation types, using the same data set used to calculate VOT, formant frequencies, and fricative spectra. SoE is a measure of the relative amplitude of voicing in the speech signal (see, e.g., Murty & Yegnanarayana, 2008; Garellek et al., 2023), and as such is a useful tool for examining the strength of periodic voicing throughout the implementation of non-modal phonation types (see Cortés et al, 2022 for a recent example on another Mixtec language). Following similar methods in Cortés et al (2022) and Garellek et al. (2023), SoE was calculated using VoiceSauce (Shue et al. 2011) at 1 ms intervals over a 10 ms window, then averaged over 20 equally-spaced intervals for each token. SoE measurements were log-transformed and then normalized by subtracting a speaker’s minimum SoE value every measurement, and dividing the result by the difference between the speaker’s maximum and minimum SoE. The results ranged between 0 and 1, with 1 representing the speaker’s highest SoE, and 0 representing the speaker’s lowest SoE.

The plots in Figure 13 show aggregated SoE contours over the course of the vocalic portion of laryngealized and breathy vowels (e.g., [VʔV] and [VhV]). The steep dip in SoE during the middle of the timecourse is consistent with creaky voice and glottal closure during the realization of laryngealization, and aperiodic frication during the realization of breathiness. In general, RDC’s productions show a steeper dip in SoE than NGC’s productions, suggesting that the realization of laryngealization and breathiness is likely subject to interspeaker variation. The plots in Figure 14 show the SoE contours for vowels and following preaspiration (e.g., the

underlined portion of [t͡ʃi̲ʰkí] (‘prickly pear fruit’)) alongside that of vowels and following prenasalization (e.g., the underlined portion of [lí̲ⁿko] (‘fruit of the maguey cactus’)). Since preaspiration involves aperiodic noise and prenasalization involves periodic voicing, SoE stays relatively high for a sequence of a vowel and following prenasalization, but dips for a sequence of a vowel and following preaspiration. This is consistent with the presence of breathy voicing and eventual aspiration, which lowers the relative strength of voicing in the acoustic signal.

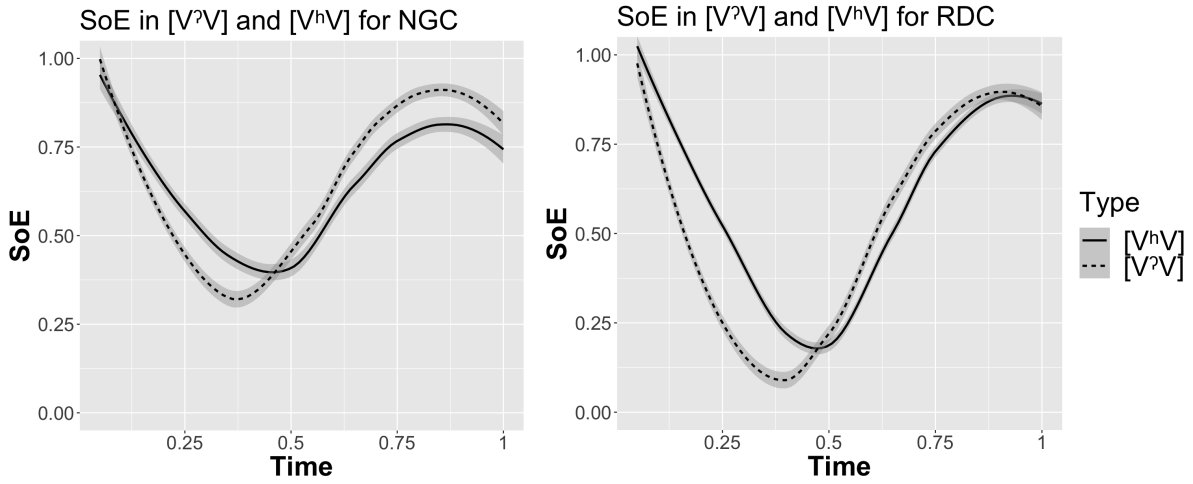


Figure 13: SoE for [VhV] (solid) and [VʔV] (dashed) sequences. Lines are smoothed loess regression lines, and grey bars represent a 95% confidence interval around the regression line. NGC [VʰV] = 57, [VʔV] = 58. RDC [VʰV] = 66, [VʔV] = 63.

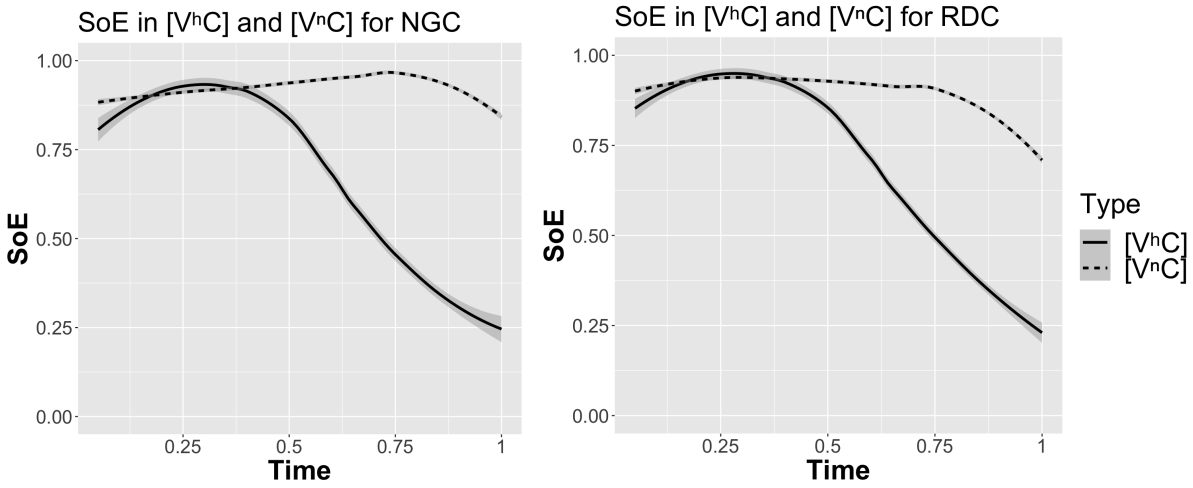


Figure 14: SoE for vowels followed by preaspiration (solid) and prenasalization (dashed). Lines are smoothed loess regression lines, and gray bars represent a 95% confidence interval around the regression line. NGC [VʰC] = 84, [VⁿC] = 63. RDC [VʰC] = 66, [VⁿC] = 66.

Lexical Tone

There are at least five phonemic tones in SMPM, with three level tones and at least two contour tones. The three level tones are High tone (marked with an acute accent *á*), Mid tone (no diacritic), and Low tone (marked with a grave accent *à*). There is one phonemic rising tone, which rises from Low to High (marked with a hacek *ǎ*), and at least one falling tone (marked with a circumflex accent *â*). Phonological tone sandhi evidence suggests that HM, ML, or HL falls may all occur. HL and H certainly contrast, as evidenced by the difference between the H-L root [tá^{h}tà] (‘señor’) and the HL-L root [ntsî^{h}kà] (`wide’). However, it is not at present clear whether HM and H contrast, or whether ML and M contrast. The mora is the tone-bearing unit, and any one of SMPM’s five tones may appear on a mono-moraic vowel (Peters, 2018). Additionally, SMPM is a laryngeally-complex language (Silverman, 1997), meaning that contrastive tone and contrastive phonation type are cross-classified: any one of SMPM’s five tones may appear on modal and non-modal vowels alike, with the exception of contrastively breathy vowels, which almost exclusively host falling or low tones. The initial vowel in roots with only modal vowels (both CVCV (12) and CVV (13) roots), in roots with glottalized vowels (14), and in roots with initial vowels followed by non-contrastive [h] (15) may all host any one of SMPM’s five contrastive tones.

(12)	a. [nánà]	*nánà*	‘mother’ or ‘madam’
	b. [kōnì]	*konì*	‘know.POT’
	c. [t͡sìnà]	*tsìnà*	‘dog’
	d. [ʃǐyò]	*xǐyò*	‘dress’
	e. [kînì]	*kînì*	‘pig’
(13)	a. [k^{w}ǻa̋]	*kuáàn*	‘yellow’
	b. [sāà]	*saà*	‘bird’
	c. [ʃàà]	*xàà*	‘chin’
	d. [ɲŭű]	*ñŭù*	‘night’
	e. [(jiβa) t^{j}ôŏ]	*(yiva) tiôŏ*	‘a type of edible mallow plant’
(14)	a. [mǻ^{ʔ}à]	*má’à*	‘raccoon’
	b. [tā^{ʔ}mǻ]	*ta’mà*	‘hillside’
	c. [ɲǻ^{ʔ}mǻ]	*ñà’mă*	‘smooth’

	d. [mǎˀnǎ]	*mă'nà*	'tired'
	e. [sʲâˀă]	*siâ'ă*	'Tecomaxtlahuaca (a town)'
(15)	a. [táʰtà]	*tátà*	'father' or 'sir'
	b. [ūʰsù]	*usù*	'deer'
	c. [t͡sìkàʰtsʲà]	*tsìkàtsià*	'round (Adj)'
	d. [jùtăʰtá]	*yùtătá*	'mirror'
	e. [ⁿt͡sîʰkà]	*ntsîkà*	'wide (Adj)'

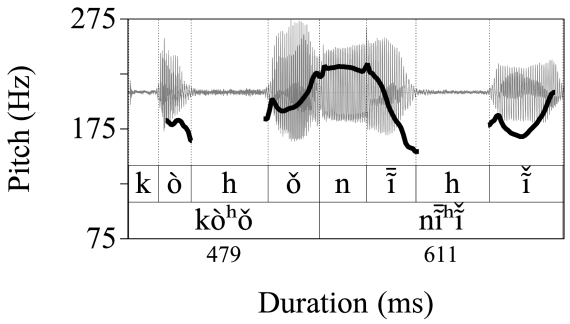
Unlike other phonation types, contrastively-breathy vowels almost always host Low or Falling tones on their first mora. These falling tones contrast, and they differ in their starting pitch. For example, the word for 'skinny' in (16a) begins with a fall whose pitch begins roughly at the level of a high tone, while the word for 'ear of corn' in (16b) begins with a fall whose pitch starts roughly at the level of a mid tone. We analyze these distinct falling tones as derived from underlying low tones, as represented in the difference between the phonemic transcription in slashes and the allophonic transcription in square brackets.

(16)	a. /nǐʰĭ̌/ [nîʰĭ̌]	*nîjĭ*	'skinny'
	b. /nī̀ʰĭ̌/ [nîʰĭ̌]	*nîjĭ*	'ear of corn'
	c. /nì̀ʰĭ̌/ [nì̀ʰĭ̌]	*nìjĭ*	'blood'

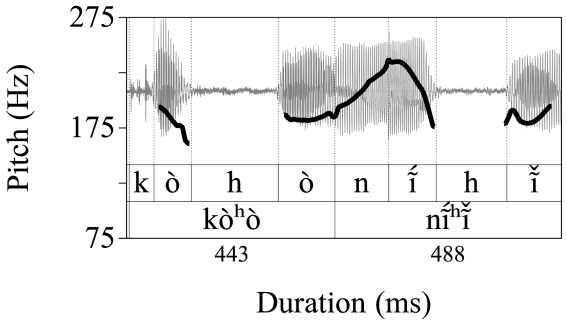
The motivation to analyze the falling tone in (16a) and the falling tone in (16b) as derived from underlying high and mid tones, respectively, is due to their asymmetric behavior with respect to the phonological tone sandhi process of Rise Flattening. In Rise flattening, discussed in the Tone Sandhi section below, word-final rising tones flatten to level low tones when followed by high tones across a word boundary. Example (17) below shows that the underlying final rise of [kòʰǒ] ('snake') surfaces faithfully before /nī̀hĭ̌/ ('ear of corn'), suggesting that the fall in 'ear of corn' does not begin with an underlying high tone. However, the final rise of [kòʰǒ] ('snake') surfaces as a low tone before /nǐʰĭ̌/ ('skinny') in (18), suggesting that the fall in 'skinny' does begin with an underlying high tone.[10] Given the sandhi patterns and that HL and ML falls on breathy vowels are in complementary distribution with level tones, we analyze the falling tones on breathy vowels as derived from underlying level tones.

[10] The fact that (17) and (18) involve distinct syntactic constructions is irrelevant here, since Rise Flattening may apply across the boundary between subject and object NPs (Eischens, 2022:79).

(17) ʃàʰʃī kòʰǒ nī̂ʰǐ̌
ate snake ear.of.corn
'The snake ate an ear of corn.'



(18) ʃàʰʃī kòʰò nī̂ʰǐ̌ nī̄hǐ̌
ate snake skinny ear.of.corn
'The skinny snake ate an ear of corn'



We know of only one instance of a non-falling high tone on the first mora of a contrastively breathy vowel, shown in (22). This word is likely derived from the root [tà̰ʰā̰] ('for there to be an earthquake'). Importantly, the following tone is mid, not rising. There appear to be no surface level high or mid tones on contrastively-breathy vowels followed by a rising tone.

(19) [tā́ʰā̰] *tájan* 'earthquake/tremor'

On the second vowel of a bi-moraic root, only four tones are found in non-derived contexts; we have found no evidence of contrastive falling tones in this context (c.f., Peters and Mendoza, 2020). With this restriction in mind, any one of the four remaining phonemic tones—high, mid, low, or rising—may occur on the second mora of the bi-moraic root. This is the case whether the preceding vowel is modal (whether in a CVCV (20) or CVV (21) root), is glottalized vowel (22), or is followed by non-contrastive [h] (23).

(20) a. [kānḯ] *kaní* 'slippery'
b. [kōlō] *kolo* 'turkey'
c. [kōnì] *konì* 'know.POT'
d. [(ñà̰) jīβǐ] *(ñà) yivǐ* 'person'

(21) a. [t͡sī̋ī̋] *tsìín* 'rat'
b. [ī̄ī̄] *iin* 'one'
c. [ʃàà] *xàà* 'chin'
d. [(jīβā) tʲôǒ̰] *(yiva) tiôǒ* 'a type of edible mallow plant'

(22) a. [ⁿt͡sʲéʔé] *ntsié'é* 'ground bean soup'
b. [βēʔē] *ve'e* 'house'
c. [ʃèʔè] *xè'è* 'garbage'
d. [ʃēʔě] *xe'ě* 'ring'

(23) a. [t͡ʃúʰtú] *tyútú* 'full (Adj)'
b. [lēʰsō] *leso* 'rabbit'
c. [ūʰsù] *usù* 'deer'
d. [jûʰkǔ] *yûkǔ* 'mountain/wilderness'

Vowels following contrastively-breathy vowels always host mid tones (24a) or rising tones (24b). We know of no examples where a high, low, or falling tone occurs on the second mora of a contrastively breathy root.

(24) a. [t͡ʃèʰē] *tyèje* 'male (Adj)'
b. [nîʰǐ] *nîjǐn* 'skinny'

Below are plots showing pitch contours for tonal categories, with values aggregated across many productions for two speakers. The plots on the left show pitch for high, mid, low, and falling tones on the first short vowel of bi-syllabic word, since this is the environment in which most falling tones occur. Falling tones are divided between those that start at the level of high tones (coded as HL) and those that start at the level of mid tones (coded as ML). The vast majority of these falls occur on contrastively breathy vowels, which almost always host falling or low tones. Rising tones are excluded from the V1 plots because they are very rare in this position, and roots with medial [ʔ] were excluded from the V1 plots because pitch readings for the vowel preceding a [ʔ] are often unreliable and sometimes even absent in this language. The plots on the right show pitch for high, mid, low, and rising tones on the second short vowel of a root. Rises were included because most rising tones in the language occur on the second vowel of the root. Falling tones were excluded because we know of no underlying Falling tones in this environment. Roots with medial [ʔ] were included in the V2 plots because pitch on the second mora is not significantly perturbed by the preceding [ʔ].

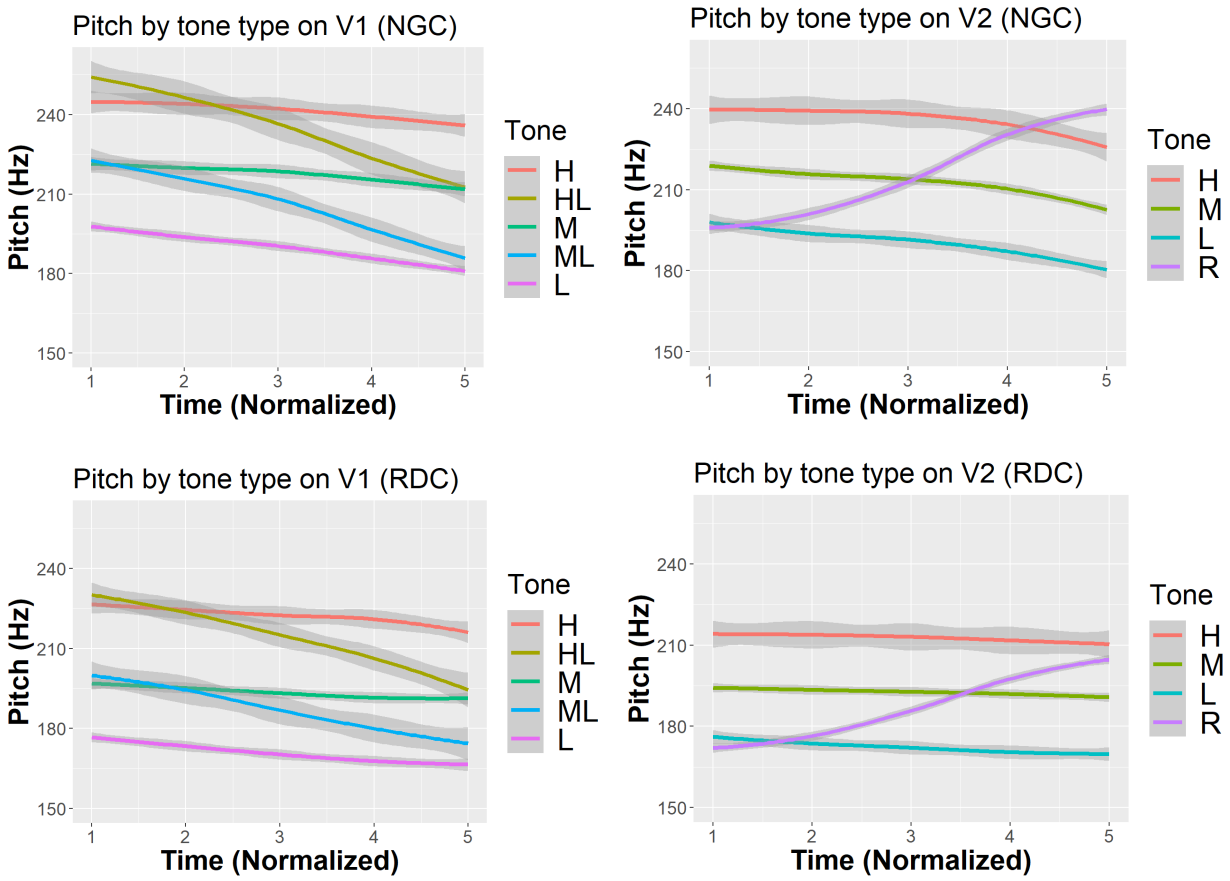


Figure 15: Pitch (Hz) by tone type on V1 (left) and V2 (right) for consultants NGC (top) and RDC (bottom). Colored lines are loess regression curves, and gray bars around the lines represent 95% confidence intervals for the curve.

Number of tokens for NGC V1: H = 51, HL = 33, M = 93, ML = 24, L = 113.

Number of tokens for RDC V1: H = 36, HL = 33, M = 93, ML = 24, L = 129.

Number of tokens for NGC V2: H = 42, M = 107, L = 80, R = 157.

Number of tokens for RDC V2: H = 51, M = 119, L = 102, R = 180.

Tone Sandhi

San Martín Peras Mixtec has relatively few tone sandhi processes when compared to some other varieties of Mixtec, like the Yucuquimi de Ocampo and Nochixtlan varieties (León Vázquez, 2017; Mckendry, 2013). In this regard, it is similar to Alcozauca Mixtec (Mendoza Ruiz, 2016; Uchihara and Mendoza Ruiz, 2021), though Alcozauca Mixtec has more tone levels than SMPM—four as opposed to three. We know of two tone sandhi processes in the language, which were first described in Hedding (2019). The first, which we call Rise Flattening, changes

an underlying rising tone to a low tone when it is immediately followed by a high tone. This process can be seen in that the underlying word-final rise on [nṵ̀ʰnḭ̌] ‘corn’, seen in (25a), surfaces as a low tone when the following word begins with a high tone (25b). Additionally, rising tones often surface with level low pitch at the end of an utterance, suggesting that an identical or similar process applies at utterance edges.

(25) a. nṵ̀ʰnḭ̌ jāā
corn white
‘white corn’

b. nṵ̀ʰnḭ̀ kʷā́à̰
corn yellow
‘yellow corn’

The second tone sandhi process, which we refer to as Low Tone Spread, changes an underlying high tone to a rising tone when it is immediately preceded by a low tone. Additionally, as demonstrated by Eischens (2022), this process is only triggered when the high tone is docked on a glottalized low vowel [aˀ], and the only words known to undergo it are adjectives. This process can be seen below, where the word [káˀnṵ̀] `big’ surfaces with its underlying initial high tone in (26a), but with an initial rise when it follows a low tone in (26b).

(26) a. ī̄ lāⁿt͡ʃī káˀnṵ̀
one lamb big
‘A big lamb’

b. ī̄ ūʰtù kǎˀnṵ̀
one corn.field big
‘A big corn field’

As noted in Hedding (2019), these two sandhi processes interact opaquely with each other. While Rise Flattening may create the conditioning environment for Low Tone Spread to apply, the second process does not apply; that is, low tones derived by Rise Flattening do not trigger Low Tone Spread. For example, the word /jûʰkŭ/ (`mountain/wilderness’) has an underlying, final rising tone (23d). This rising tone becomes a low tone when followed by the initial high tone of [káˀnṵ̀] (`big’). However, the initial high tone of ‘big’ does not undergo Low Tone Spread (27), even though it is immediately preceded by a derived low tone.

(27) ī̄ī̄ jûʰkù káʔnǔ̀
one mountain big
'A big mountain.'

4 Illustrative Passage

The story of the North Wind and the Sun is not a native Mixtec story, so consultant NGC read the story in Spanish and translated it into San Martín Peras Mixtec. To familiarize the consultant with the story, she first translated it sentence-by-sentence into Mixtec. Once she was familiar enough with the story, she told it several times from beginning to end without referencing her sentence-by-sentence translation, and she chose the telling that seemed most natural and accurate to her. This telling differed substantially from the sentence-by-sentence translation and is the one written below. We have included a transcription of the story in the working orthography described earlier in the article, as well as a narrow transcription including a three-line gloss.

Orthographic version

Tàtyĭ nórtè xí'ìn tsìkàntsìjĭ kìxă nà kâ'àn nà yòó nà ntakù tyáákà, tá nìyà'ă iin rà xĭka ìníìn ñuù yivĭ xí'ìn tsiàà kă'nù ítiví rà. Nìkà'àn nà yòó nà kevà'ă kasa ntúxa nà xí'ìn rà tàvă nà míí tsiàà ítiví rà, rà kúu nà nà ntakù nùjŭ ntsìkúu ñà'ă ñuù yivĭ. Míí tàtyĭ ñà nórtè ntakù và'ă tsìvyà xí'ìn ntsìkúu ntsièjâ, só nú kuà'ă và'ă tsívyà, kuà'ă tyáákà ná tìjvyâ tsyàà kă'nù míí rà xĭka ìníìn ñuù yivĭ. Nŭ ntsí'i kùntaà ini míí tàtyĭ nórtè kòníkùù tyiñà. Sáá nàye'è tsìkàntsìjĭ xí'ìn ntsìkúu ñà i'níà. Kamà và'ă tàvă míí rà xíka ìníìn ñuù yivĭ tsyàà kă'nù ítiví rà. Nŭ kùndaà ini tàtyĭ tsìkàntsìjĭ yă ntakù tyáákà nùjŭ ntsìkúu ñà'ă.

Transcription

(1) tàʰt͡ʃĭ nǒrtè ʃĭʔ̀ì t͡sìkàⁿt͡sìʰĭ kìʰʃă nǎ kǎʔ̀ă nǎ jòó
wind north with sun begin.COMP 3PL talk.CONT 3PL who

nǎ ⁿtāʰkù t͡ʃáá=kà
3PL strong more=ADD
"The North Wind and the Sun were disputing which was the stronger,"

(2) tá nì̀-jàʔ̀ă ī̄ī̄ ɾà ʃĭʰkā ì̀nʔ́ì̀ ɲū̄ǔ jīβĭ ʃĭʔ́ì̀ t͡sʲàà
when COMP-walk one 3SG.M walk.CONT around town person with clothes

kǎˀnǔ ítīβí ɾà
big wrap.CONT 3SG.M
“when a traveler came along wrapped in a warm cloak.”

(3) nɨ̀-kɑ̀ˀɑ̀ nɑ̀ jòó nɑ̀ kēβàˀă kāʰsā ⁿtúʰʃā nɑ̀ ʃɪ́ˀɪ̀
COMP-talk 3PL who 3PL win.POT make strength 3PL with

ɾà tàβă nɑ̀ mɪ́ɪ́ t͡sʲàà ítβí ɾà
3SG.M take.off.POT 3PL FAM clothes wrap 3SG.M
“They agreed that the one who first succeeded in making the traveler take his cloak off”

(4) ɾà kūū nɑ̀ nɑ̀ ⁿtāʰkù nǔ̀ʰǔ̀ ⁿt͡sìkúū ɲɑ̀ˀɑ̌ ɲūǔ̀ jīβĭ
3SG.M be.POT 3PL 3PL strong over all thing town person
“should be considered stronger than the other” (lit. would be stronger than anything in the world)

(5) mɪ́ɪ́ tàʰt͡ʃī ɲɑ̀ nɔ́ɾtè ⁿtāʰkù βàˀā t͡sìβī=à ʃɪ́ˀɪ̀ ⁿt͡sìkúū
FAM wind 3SG.N north strong well blow.COMP=3SG.N with all

t͡sʲèʰĕ=à
strength=3.SG.N
“Then the North Wind blew as hard as he could,”

(6) só nǔ́ kʷàˀă βàˀā t͡síβī=à
but when very well blow.CONT=3SG.N
“but the more he blew”

(7) kʷàˀă t͡ʃáákà nɑ́ tìʰβĭ=à t͡sʲàà kǎˀnǔ̀ mɪ́ɪ́ ɾà
very more HORT wear.COMP=3SG.N clothes big FAM 3SG.M

ʃĭʰkā ɨ̀nɪ́ɪ̀ ɲūǔ̀ jīβĭ
walk.CONT around town person
“the more closely did the traveler fold his cloak around him;”

(8) nǔ̌ ⁿtsíˀī kùⁿtāà īnɪ̄ mɪ́ɪ́ tàʰt͡ʃĭ nɔ́ɾtè kò-nɪ́-kùù
when finish.CONT realize.COMP inside FAM wind north NEG-COMP-able

t͡ʃɪ̄ñū̄=à
work=3SG.N
“and at last the North Wind gave up the attempt.”

(9) sáá nàjēˀè t͡sìkàⁿt͡sìʰĭ ʃɪ́ˀɪ̀ ⁿt͡sìkúū ñɑ̀ iˀnɪ́=à
so shine.COMP sun with all 3SG.N heat=3SG.N
“Then the Sun shone out warmly,”

(10) kāmǎ̀ βàʔā tàβă mḯḯ ɾà ʃǐʰkā ɲ̀ɲ́ḭ̀ ɲūǔ jīβĭ
fast good take.off.COMP FAM 3SG.M walk.CONT around town person

t͡sʲàà kǎʔnǔ́ ítīβí ɾà
clothes big wrap.CONT 3SG.M
“and immediately the traveler took off his cloak.”

(11) nǔ̌ kùⁿdāà īnī tàʰt͡ʃǐ
when realize.COMP inside wind
“And so the North Wind was obliged to confess,”

(12) t͡sìkàⁿt͡sìʰĭ já ⁿtāʰkù t͡ʃáá=kà nǔ̀ʰǔ̌ ⁿt͡sìkúū ɲà̰ʔǎ̰
sun 3SG.N strong more=ADD over all thing
“that the Sun was the stronger of the two.”

Acknowledgments

Placeholder for acknowledgments

References

Abramson, Arthur S., and Whalen, D.H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, 75-86.

Auderset, Sandra, Greenhill, Simon J., DiCanio, Christian T., and Campbell, Eric W. (2023). Subgrouping in a 'dialect continuum': A Bayesian phylogenetic analysis of the Mixtecan language family. *Journal of Language Evolution*, 8, 33-63.

Beatham, Bruce and Beatham, Candice (2019). Propuesta de convenciones para escribir el mixteco del oeste de Juxtlahuaca. Instituto Lingüístico de Verano.

Becerra Roldán, Braulio. "Un estudio fonológico del mixteco de Santo Domingo Huendio, Oaxaca." (2015). Licentiate Thesis, Escuela Nacional de Antropología e Historia.

Becerra Roldán, Braulio (2019). *Análisis sincrónico y consideraciones diacrónicas sobre la fonología del mixteco de San Pedro Tuliztlahuaca*. Master's Thesis, La Universidad Autónoma de México.

Boersma, Paul and Weenink, David (2020). Praat: Doing phonetics by computer (version 6.1.27).

Bravo Vargas, Gerardo (2014). *Vocabulario Náhuatl del maguey y el pulque*. Asociación de amigos del museo del maguey y el pulque.

Caballero, Gabriela, Claudia Juárez Chávez, and Michelle Yuan (to appear). The representation of tone in San Juan Piñas Mixtec (Tò'ōn Ndā'ví): Phonological and orthographic implications. In: *Proceedings of WCCFL 39*.

Campbell, Eric W. (2017). Otomanguean historical linguistics: Exploring the subgroups. *Language and Linguistics Compass*, 11(e12244).

Canfield, D. Lincoln (1981). *Spanish Pronunciation in the Americas*. University of Chicago Press.

Cortés, Félix, Mantenuto, Iara, and Steffman, Jeremy (2023). San Sebastián del Monte Mixtec. *Journal of the International Phonetic Association,* 1-22.

Daly, John P., and Larry M. Hyman (2007). "On the representation of tone in Peñoles Mixtec." *International Journal of American Linguistics,* 73 (2): 165-207.

DiCanio, Christian, Jonathan Amith, and Rey Castillo García (2014). "The phonetics of moraic alignment in Yoloxóchitl Mixtec."

DiCanio, Christian, Joshua Benn, and Rey Castillo García (2018). "The phonetics of information structure in Yoloxóchitl Mixtec." *Journal of Phonetics* 68: 50-68.

DiCanio, Christian, Joshua Benn, and Rey Castillo García (2021). "Disentangling the effects of position and utterance-level declination on the production of complex tones in Yoloxóchitl Mixtec." *Language and Speech,* 64 (3): 515-557.

DiCanio, Christian T., Zhang, Caicai, Whalen, Douglas H., and Castillo García, Rey (2020). Phonetic structure in Yoloxóchitl Mixtec consonants. *Journal of the International Phonetic Association*, 50(3), 333-365.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig, editors (2022). *Ethnologue: Languages of the World.* Twenty-fifth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Eischens, Ben (2022). *Tone, Phonation, and the Phonology-Phonetics Interface in San Martín Peras Mixtec.* PhD Thesis, University of California, Santa Cruz.

Eischens, Ben (2023). Rate-sensitive differences between modal and non-modal vowels in San Martín Peras Mixtec. In: Radek Skarnitzl & Jan Volín (Eds.), Proceedings of the 20th International Congress of Phonetic Sciences (pp. 3212–3416). Guarant International.

Embriz Osorio, Arnulfo and Óscar Zamora Alarcón, editors (2012). *México: lenguas indígenas nacionales en riesgo de desaparición*. Instituto Nacional de Lenguas Indígenas.

Garellek, Marc, Chai, Yuan, Huang, Yaqian, and Van Doren, Maxine (2023). Voicing of glottal consonants and non-modal vowels. *Journal of the International Phonetic Association*, 53(2), 305-332.

Gerfen, Chip (2013). *Phonology and phonetics in Coatzopan Mixtec*. Springer.

Gerfen, Chip, and Kirk Baker (2005). "The production and perception of laryngealized vowels in Coatzospan Mixtec." *Journal of Phonetics,* 33 (3): 311-334.

Hedding, Andrew A. (2019). *Two Tone Sandhi Processes in San Martín Peras Mixtec.* Manuscript, University of California, Santa Cruz.

Herrera Zendejas, Esther (2014). *Mapa fónico de las lenguas mexicanas.* Ciudad de México, México: El Colegio de México.

Instituto Nacional de Estadística y Geografía (2020). Censo de población y vivienda.

Instituto Nacional de Lenguas Indígenas (2008). Catálogo de las lenguas indígenas nacionales: Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas.

Iverson, Gregory K., and Salmons, Joseph C. (1996). Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics*, 62(2): 165-175.

Josserand, Judy K. (1983). *Mixtec Dialect History*. PhD Thesis, Tulane University.

Keating, Patricia, Wymark, Daniel, and Sharif, Ryan. 2019. Proposal for superscript diacritics for prenasalization, preglottalization and preaspiration. *Journal of the International Phonetic Association*, 51(1), 75-90.

Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry* 10(3), 421-441.

León Vázquez, Octavio (2017). *Sandi tonal en el mixteco de Yucuquimi de Ocampo*. Master's Thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social.

Lisker, Leigh, and Abramson, Arthur S. (1967). Some effects of context on Voice Onset Time in English stops. *Language and Speech*, 10(1): 1-28.

Macaulay, Monica (1996). *A grammar of Chalcatongo Mixtec*. Vol. 127. University of California Press.

Macaulay, Monica and Joseph C. Salmons (1995). The phonology of glottalization in Mixtec. *International Journal of American Linguistics*, 61(1): 38-61

Mantenuto, Iara (2020). 'Other' and 'More' in San Sebastián del Monte Mixtec and Beyond. PhD thesis, University of California, Los Angeles.

Marlett, Stephen A. (1992). "Nasalization in Mixtec languages." *International Journal of American Linguistics,* 58 (4): 425-435.

McKendry, Inga (2013). Tonal association, prominence and prosodic structure in South-Eastern Nochixtlán Mixtec. PhD Thesis, University of Edinburgh.

Mendoza, Gabriel (2020). *Syntactic sketch of San Martín Peras Tu'un Savi*. BA Thesis, University of California, Santa Barbara.

Mendoza Ruiz, Juana (2016). *Fonología segmental y patrones tonales del Tu'un Savi de Alcozauca de Guerrero*. Master's Thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social.

Murty, K. S. R. and Yegnanarayana, B. (2008). Epoch Extraction From Speech Signals. In *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1602-1613.

Norma de Escritura del Tu'un Savi (Idioma Mixteco) (2022). Instituto Nacional de Lenguas Indígenas.

North, Joanne, and Jäna Shields (1977). "Silacayoapan mixtec phonology." *Studies in Otomanguean Phonology* 54: 21-33.

Pankratz, Leo and Eunice V. Pike (1967). Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics*, 33 (4): 287-299.

Penner, Kevin L. (2019). *Prosodic structure in Ixtayutla Mixtec: Evidence for the foot*. PhD Thesis, The University of Alberta.

Peters, Simon L. (2018). *The inventory and distribution of tone in Tù'un Ndá'vi, the Mixtec of Piedra Azul (San Martín Peras), Oaxaca*. Master's thesis, University of California, Santa Barbara.

Peters, Simon L. and Mendoza, Gabriel (2020). Morphophonological processes in Piedra Azul Tù'un Ntá'vi. Presentation at the Winter 2020 meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA).

Piedra-Malagón, Eva M., Sosa, Victoria, Angulo, Diego F., and Díaz-Toribio, Milton H. (2022). Edible native plants of the Gulf of Mexico province. *Biodiversity Data Journal*, 10.

Pike, Kenneth L. (1948). *Tone Languages.* University of Michigan Press.

Pike, Eunice V., and Priscilla Small (1974). "Downstepping terrace tone in Coatzospan Mixtec." *Advances in tagmemics*: 105-134.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rueda Chaves, John E. (2019). *La interacción entre el tono y el acento en el mixteco de San Jerónimo de Xayacatlán*. PhD Thesis, El Colegio de México.

Shue, Yen-Liang, Keating, Patricia, Vicenik, Chad, and Yu, Kristine (2011). Voicesauce: A program for voice analysis. In Proceedings of the International Congress of Phonetic Sciences XVII. 1846-1849.

Silverman, Daniel (1997). Laryngeal complexity in Otomanguean vowels. *Phonology*, 14(2): 235-261.

Stremel, Sophia (2022). /i/-deletion in San Martín Peras Mixtec. Manuscript, University of California, Santa Cruz.

Towne, Douglas (2011). *Gramática popular del tacuate (mixteco) de Santa María Zacatepec, Oaxaca.* Instituto Lingüístico de Verano, Mexico.

Uchihara, Hiroto and Juana Mendoza Ruiz (2022). Minimality, maximality and perfect prosodic word in Alcozuaca Mixtec. *Natural Language & Linguistic Theory*, 40: 599-649

Wickham, Hadley (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Appendix A:

The vocabulary items produced by JGO in this manuscript were elicited in carrier phrases in order to control for the effect of tone sandhi and utterance-level intonation on the realization of tones. However, the carrier phrase that was used was not the same in every case. The carrier phrases were adjusted as needed to create naturalistic and plausible utterances according to the preferences of our language consultant. Consequently, we used multiple carrier phrases, with the constraint that, whenever possible, the tones immediately preceding and following the target word be mid tones. Because the carrier phrases are helpful in hearing tonal contrasts, and because the carrier phrases were not always identical, the following list contains the gloss of each carrier phrase for each target word given throughout this Illustration. The glosses are numbered in the order that they occur in the paper. For ease of cross-referencing, any example or figure number is also included following an underscore, if applicable.

(1) ī̄ī̄ pá^{h}tò lō^{ʔ}ō
one duck small
'A small duck.'

(2) ī̄ī̄ ^{m}páà lō^{ʔ}ō
one godfather small
'A small godfather.'

(3) ī̄ī̄ tá^{h}tà lō^{ʔ}ō ɾà
one sir small he
`He is a small man.'

(4) ī̄ī̄ ɾà t͡s^{j}à^{h}a ntáʔβī
one CLF man poor
`A poor man.'

(5) t^{j}à^{h}tá
oak.tree
'(A type of oak tree)' (no carrier phrase)

(6) ^{n}tʲaʔmǐ t͡sʲáʔji
radish rotten
`A rotten radish'

(7) kōnī̄ kàʰkū ī̃ī̃ ɾà lōʔō
yesterday was.born one he small
'A boy was born yesterday.'

(8) ɾà kìʃi ĩ̃ⁿkà ɲū̃ù̃
he came other town
'He came from another town.'

(9) lēʰsō kʷáʔà lōʔō
rabbit red small
'A small, red rabbit.'

(10) ī̃ī̃ kʲáʔmǐ̌ lōʔō
one squash small
'A small squash.'

(11) ī̃ī̃ mā́lì lōʔō
one godmother small
'A small godmother.'

(12) ī̃ī̃ nā́nà̀ lōʔō
one mother/woman small
'A small woman.'

(13) ī̃ī̃ ɲã̄nī̄ ēstéβà
one brother Estébà
'A brother of Esteban.'

(14) ī̃ī̃ ɾà lōʔō
one he small
'A boy.'

(15) ī̃ī̃ sâʔmǎ̌ lōʔō
one tortilla.cloth small
'A small tortilla cloth.'

(16) ñū̃ù̃ sʲâʔă βà
town Tecomaxtlahuaca particle
'The town of Tecomaxtlahuaca.'

(17) ĩĩ ʃàà lōʔō
one chin small
‘A small chin’

(18) ĩĩ t͡sʲǎʔà plástiko
one jug plastic
‘A plastic jug’

(19) lēʰsō ⁿt͡sʲàʔβă lōʔō
rabbit toothless small
‘A small toothless rabbit’

(20) kʷáʔà t͡ʃááɣa kìʰʃă sufríì ⁿtù
much more began suffer we.exclusive
‘We (excl.) began to suffer much more.’

(21) ĩĩ ⁿt͡ʃĩʰʃĭ lōʔō
one corncob small
‘A small corncob.’

(22) lēʰsō βálí jāʔá
rabbit small.Pl orange
‘Small, orange rabbits.’

(23) lēʰsō jāā lōʔō
rabbit white small
‘A small, white rabbit.’

(24) ĩĩ lāⁿt͡ʃĩ lōʔō
one lamb small
‘A small lamb.’

(27_1a) ĩĩ t͡ʃìʰkí lōʔō
one prickly.pear.fruit small
‘A small prickly pear fruit.’

(28_1b) ĩĩ ⁿt͡ʃĩʰʃĭ lōʔō
one corncob small
‘A small corncob.’

(29_2a) ĩĩ ⁿt͡sîʰkă lōʔō
one banana small
‘A small banana.’

(30_2b) íjō ^{n}t͡sʲêʰě ɾà
there.is strength his
'He is strong'

(31_2c) lēʰsō ^{n}t͡sʲàʔβă lōʔō
rabbit toothless small
'A small, toothless rabbit.'

(32_2d) ^{n}t͡sʲôʰǒ
moonlight
'Moonlight' (no carrier phrase)

(33_2e) t͡sʲı̀ı̄ t͡sʲùʔú̠ t͡ʃúʰtū
bit venomous.spider cat
'The venomous spider bit the cat.'

(34_2a) t͡ʃĭʔⁿkì
acorn
'Acorn' (no carrier phrase)

(35_2b) ı̄ı̄ līⁿkō íʰt͡sʲă lōʔō
one maguey.bud soft/young little
'A small, soft maguey flower bud.'

(49) nɑ̀ kátáʔă nɑ̀
they fight they
'They are fighting.'

(50) nɑ̀ kéβāʔă nɑ̀
they win they
'They are winning.'

(51) nɑ̀ kítáʔă nɑ̀
they accompany they
'They are accompanying.'

(52) nɑ̀ káā kò-kɑ́ʔɑ̀ nɑ̀
they DEM NEG-speak they
'They do not speak.'

(53) nɑ̀ káā kúkáʔnū̄ īnī̄ nɑ̀
they DEM forgive inside they
'They forgive.'

(54) lēʰsō kʷáà lōʔō
rabbit blind small
‘A small, blind rabbit.’

(55) lēʰsō kʷáǎ lōʔō
rabbit yellow small
‘A small, yellow rabbit.’

(56) ī̄ī̄ tsi̯kʷǐì kʷíì lōʔō
one water clear little
‘Some water that is a little bit clear.’

(57) lēʰsō kʷíˆ̀ī̀ lōʔō
rabbit striped little
‘A small, striped rabbit.’

(58) ɾà kūū ɾà
he will.die he
‘He will die.’

(59) ɾá kū̃ū̃ ɾá
it.liquid will.fall it.liquid
‘It (rain) will fall.’

(60_4a) lēʰsō jāā lōʔō
rabbit white small
‘A small, white rabbit.’

(61_4b) ī̄ī̄ jâʰă lōʔō
one tongue small
‘A small tongue.’

(62_4c) ī̄ī̄ jāʔă lōʔō
one chile.pepper small
‘A small chile pepper.’

(63_5a) kúⁿt͡sʲēē ɾà
puts.up.with he
‘He puts up with it.’

(64_5b) íjō ⁿt͡sʲêʰě ɾà
there.is strength his
‘He is strong’

(65_5c) ʃìʃi ɾà lōʔō ⁿt͡sʲéʔé kōnī̄
ate he small ground.bean.soup yesterday
‘The boy ate ground bean soup yesterday.’

(66_6a) ĩ̄ĩ̄ tsi̯kʷǐì kʷíì lōʔō
one water clear little
‘Some water that is a little bit clear.’

(67_6b) lēʰsō kʷî̌ʰǐ lōʔō
rabbit green small
‘A small, green rabbit.’

(68_6c) ĩ̄ĩ̄ kʷìʔī lōʔō
one fruit small
‘A small fruit.’

(69_7a) kōō βēʔē
welcome house
‘Welcome to the house.’

(70_7b) ĩ̄ĩ̄ kòʰǒ lōʔō
one snake small
‘A small snake.’

(71_7c) ĩ̄ĩ̄ kòʔǒ lōʔō
one plate small
‘A small plate.’

(72_8a) ĩ̄ĩ̄ ⁿtúù lōʔō
one nduu small
‘A small *nduu.*’ (type of insect)

(73_8b) ĩ̄ĩ̄ ⁿtùʰǔ lōʔō
one fruit.seed small
‘A small fruit seed.’

(74_8c) ɾà ⁿtūʔú ɾà
he fat he
‘He is fat.’

(75_9a) lēʰsō kʷǻǎ lōʔō
rabbit yellow small
‘A small, yellow rabbit.’

(76_9b) ɾà kʷâ̂ʰǎ ɾà
he unmarried he
‘He is unmarried.’

(77_9c) ɾà kʷáʔǎ ɾà
he leaves he
'He leaves.'

(78_10a) ī̄ ī́ lōʔō
one hail small
'Some small hail.'

(79_10b) ī̄ ī̂ʰǐ lōʔō
one skin small
'A little bit of skin/leather.'

(80_10c) ī̄ íʔī lōʔō
one temazcal small
'A small temezcal.' (a type of sweat lodge)

(81_11a) ī̄ ɲū̃ù̀ lōʔō
one town small
'A small town.'

(82_11b) ī̄ ɲù̀ʰǔ̌ lōʔō
one palm.plant small
'A small palm plant.'

(83_11c) ī̄ ɲū̃ʔù̀ lōʔō
one fire small
'A small fire.'

(96_12a) ī̄ nǎ́nà̀ lōʔō
one mother/woman small
'A small woman.'

(97_12b) ʃǐ́nī̃.ɲú̃ʔù̀ kōnǐ̀ ɾà ñá́
has.to know he she
'He has to know her.'

(98_12c) ī̄ t͡sìnà̀ lōʔō
one dog small
'A small dog.'

(99_12d) ī̄ ʃĭjò lōʔō
one dress small
'A small dress.'

(100_12e) ī̄ī̄ kînǐ̀ lōʔō
one pig small
'A small pig.'

(101_13a) lēʰsō kʷáà̌ lōʔō
rabbit yellow small
'A small, yellow rabbit.'

(102_13b) ī̄ī̄ sāà lōʔō
one bird small
'A small bird.'

(103_13c) ī̄ī̄ ʃàà lōʔō
one chin small
'A small chin.'

(104_13d) nìjàʔā ɲǔǔ̀ βīʰt͡sī̄
arrived night now
'It is now nighttime.'

(105_13e) ʃiʰʃî jīβā tʲôǒ kōnī̄
I.ate plant mallow yesterday
'I ate *yiva tioo* yesterday.' (a type of mallow plant)

(106_14a) ī̄ī̄ mā́ʔà̌ lōʔō
one raccoon small
'A small raccoon.'

(107_14b) ī̄ī̄ tāʔmà̌ lōʔō
one hillside small
'A small hillside.'

(108_14c) ī̄ī̄ sāà lōʔō mī́ī́ ɲà̌ʔmǎ̌ ɾí
one bird small EMPH smooth it.Animal
'A small bird that is smooth.'

(109_14d) ī̄ī̄ mǎ̌ʔnà̌ lōʔō
one tired small
'Someone who is a little tired.'

(110_14e) ɲū̃ǔ̀ sʲâʔă βà
town Tecomaxtlahuaca particle
'The town of Tecomaxtlahuaca.'

(111_15a) ĩ̄ĩ̄ tá^{h}tà lō^{ʔ}ō ɾà
one father small he
'He is a small man.'

(112_15b) ĩ̄ĩ̄ u^{h}sù lō^{ʔ}ō
one deer small
'A small deer.'

(113_15c) βē^{ʔ}ē ts̩ìkà^{h}ts^{j}à lō^{ʔ}ō
house round small
'A small, round house.'

(114_15d) ĩ̄ĩ̄ jùtă^{h}tá lō^{ʔ}ō
one mirror small
'A small mirror.'

(115_15e) ĩ̄ĩ̄ βē^{ʔ}ē ^{n}tsî^{h}kà lō^{ʔ}ō
one house wide small
'A small, wide house.'

(116_16a) lē^{h}sō nî^{h}ǐ̌ lō^{ʔ}ō
rabbit skinny small
'A small, skinny rabbit.'

(117_16b) ĩ̄ĩ̄ nî^{h}ǐ̌ lō^{ʔ}ō
one ear.of.corn small
'A small ear of corn.'

(118_16c) íjō lō^{ʔ}ō nì^{h}ǐ̌ βī^{h}tsī̄
there.is small blood now
'There is now a little bit of blood.'

(121_19) ĩ̄ĩ̄ tā́^{h}ã̄ lō^{ʔ}ō
one earthquake small
'A small earthquake.'

(122_20a) lē^{h}sō kānǐ́ lō^{ʔ}ō
rabbit slippery small
'A small, slippery rabbit.'

(123_20b) ĩ̄ĩ̄ kōlō lō^{ʔ}ō
one turkey small
'A small turkey.'

(124_20c) ʃǐnī̀.ɲúʔù̀ kōnì̀ ɾà ñá́
has.to know he she
‘He has to know her.’

(125_20d) ī̄ī̄ ñà̀ jīβĭ lōʔō
one they person small
‘A small person.’

(126_21a) ī̄ī̄ tsì̋ī̋ lōʔō
one rat small
‘A small rat.’

(127_21b) íjō ī̄ī̄ lēʰsō
there.is one rabbit
‘There is a rabbit.’

(128_21c) ī̄ī̄ ʃàà lōʔō
one chin small
‘A small chin.’

(129_21d) ʃiʰʃî jīβā tʲôǒ kōnī̄
I.ate plant mallow yesterday
‘I ate *yiva tioo* yesterday.’ (a type of mallow plant)

(130_22a) ʃìʃī ɾà lōʔō ⁿt͡sʲéʔé kōnī̄
ate he small ground.bean.soup yesterday
‘The boy ate ground bean soup yesterday.’

(131_22b) ī̄ī̄ βēʔē lōʔō
one house small
‘A small house.’

(132_22c) ī̄ī̄ ʃèʔè lōʔō
one trash small
‘A little bit of trash.’

(133_22d) ī̄ī̄ ʃèʔě lōʔō
one ring small
‘A small ring.’

(134_23a) ʃǐnì̀ ī̄ī̄ βēʔē t͡ʃúʰtú kōnī̄
I.saw one house full yesterday
‘I saw a full house yesterday.’

(135_23b) ī̄ī̄ lēʰsō lōʔō
one rabbit small
‘A small rabbit.’

(136_23c) ī̄ī̄ ūʰsù lōʔō
one deer small
‘A small deer.’

(137_23d) ī̄ī̄ jûʰkǔ lōʔō
one wilderness small
‘A small mountain/wilderness.’

(138_24a) lēʰsō t͡ʃèʰē lōʔō
rabbit male small
‘A small, male rabbit.’

(139_24b) lēʰsō nî̌ʰǐ̌ lōʔō
rabbit skinny small
‘A small, skinny rabbit.’