ILLUSTRATION OF THE IPA

San Martín Peras Mixtec

Ben Eischens[1],* ⓘ and Andrew A. Hedding[2] ⓘ

[1]University of California, Los Angeles and [2]University of Washington
*Corresponding author. Email: beischens@ucla.edu

San Martín Peras Mixtec

San Martín Peras Mixtec (autonym: *Tu'un Nta'vi* [tṹ$^{\text{ʔ}}$ū̄ ⁿtá$^{\text{ʔ}}$βī] or *Tu'un Savi* [tṹ$^{\text{ʔ}}$ū̄ sàʰβĭ]) 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), as shown in Figure 1. 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 12 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, 2022). A recent Bayesian phylogenetic analysis of Mixtecan languages identified 23 distinct subgroups and classified 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 in San Martín Peras has contrastive breathy phonation (see section 3), while the Mixtec spoken in neighboring Coicoyán de las Flores does not (Beatham & Beatham, 2019). Moreover, as migration and increased connectedness

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

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 to say that the term "San Martín 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."

This work adds to a substantial list of phonological and phonetic studies of the sound systems of Mixtec languages, going back to the mid-twentieth century (Pankratz & Pike, 1967; Pike & Small, 1974; North & Shields, 1977; Josserand, 1983; Marlett, 1992; Macaulay & Salmons, 1995; Iverson & Salmons, 1996; Gerfen & Baker, 2005; Daly & 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 & Mendoza, 2020; DiCanio et al., 2021; Cortés et al., 2023; Eischens, 2022; Stremel, 2022; Uchihara & Mendoza, 2021; Eischens, 2023; Caballero et al., 2024; among others).

The data in this article were collected from three 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 1,200 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 fifties (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., [t͡sìkū ʰʃí] vs. [t͡ʃìkū ʰʃí] 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]

	Bilabial	Alveolar	Palatalized Alveolar	Postalveolar	Palatal	Velar	Labialized velar	Palatalized Velar
Plosive	p/^{m}p	t/^{n}t	tʲ/^{n}tʲ			k/^{ŋ}k	kʷ	kʲ
Nasal	m	n			ɲ			
Tap		ɾ						
Fricative		s	sʲ	ʃ				
Affricate			t͡sʲ/^{n}t͡sʲ	t͡ʃ/^{ɲ}t͡ʃ				
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*)

---

[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, 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 "obstruents" in section 1.1.

[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 (v̀), rising tones with a caron (v̌), and falling tones with a circumflex (v̂). 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 on 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.

| [táʰtà] | *tátà* | 'father' or 'sir' |
[ⁿtáˀβī]	*ntá'vi*	'poor'
[tʲàʰtá]	*tiàtá*	'a type of oak tree'
[ⁿtʲāˀmǐ]	*ntia'mǐ*	'radish'
[kàʰkū]	*kàku*	'was born'
[ĩᵐkà]	*iinkà*	'other'
[kʷáˀà]	*kuá'à*	'red'
[kʲáˀmĩ]	*kiá'mǐ*	'(a type of) squash'
[mãlì]	*málì*	'godmother of one's child, or mother of one's godchild'
		(cf. Spanish *comadre*)
[nãnā]	*nánà*	'mother' or 'madam'
[ɲãnĩ]	*ñani*	'brother (of a male)'[3]
[rà lōˀō]	*rà lo'o*	'boy' (literally "small male")
[sāˀmǎ]	*sa'mǎ*	'tortilla cloth'
[sʲâˀǎ]	*siâ'ǎ*	'Tecomaxtlahuaca' (a town)
[ʃàà]	*xàà*	'chin'
[t͡sʲǎˀà]	*tsiǎ'à*	'jug/pitcher'
[ⁿt͡sʲàˀβǎ]	*ntsià'vǎ*	'toothless'
[t͡ʃáákà]	*tyáákà*	'more/most'
[ɲ͡t͡ʃĩʰʃĩ]	*ntyixǐ*	'corncob'
[βálí]	*válí*	'small (pl)'
[jāā]	*yaa*	'white'
[lāⁿt͡ʃĩ]	*lantyi*	'lamb'

1.1 Obstruents

San Martín Peras Mixtec has three 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 /t͡ʃáá=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'mǐ*

---

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

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

	RDC	NGC
CoG mean (SD)	7669 (596)	8258 (792)
[s]	n = 34	n = 34
CoG mean (SD)	4610 (550)	5394 (527)
[ʃ]	n = 21	n = 18
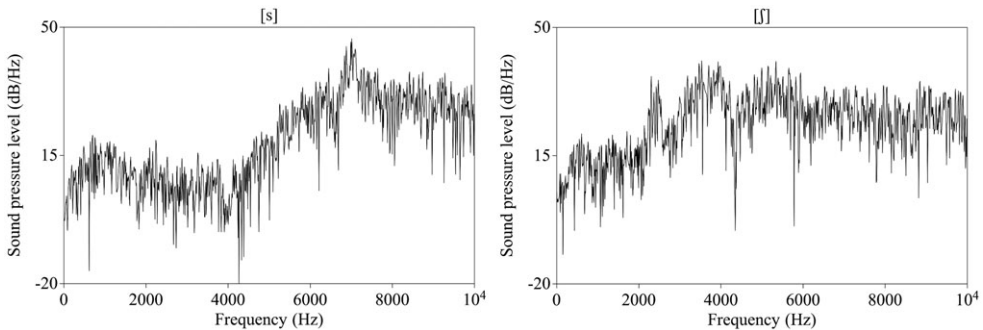
Figure 2. FFTs for [s] (left) and [ʃ] (right) produced by NGC. Examples taken from the words [sânã̌] ('crazy') and [ʃânǔ] ('cigarette').

[kʲáˀmǐ] '(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 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 *ñà yivǐ* [ɲà jĩβǐ] 'person' (20d), but palatalized consonants never precede the vowel [i]. That is, a hypothetical word like [kʲĩβǐ] 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 two voiceless fricatives: /s/ and /ʃ/. In a small number of words, /s/ is contrastively palatalized, e.g., *siâ'ǎ* [sʲâˀǎ] 'Tecomaxtlahuaca (a town)'. Fast Fourier transforms (FFTs) of individual tokens of [s] and [ʃ] from NGC are shown in Figure 2, along with average center of gravity (CoG) measurements for both fricatives for NGC and RDC in Table 1. 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 & Weenink, 2023).

In addition to fricatives, SMPM has two 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 /ⁿt͡ʃ/ can be seen in (1) below.

(1) a. [t͡ʃĩʰkí]   *tyìkí*   'prickly pear fruit'

      b. [ʲⁿt͡ʃĩʰʃǐ]   *ntyixǐ*   'corncob'

Unlike the plosive series, these affricates are not contrastively palatalized. /t͡ʃ/ and /ⁿt͡ʃ/ 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. [ⁿt͡sîʰkǎ]   *ntsîkǎ*   'banana'

     b. [ⁿt͡sʲê̂ʰě]   *ntsiêjě*   'strength'

     c. [ⁿt͡sʲà̂ʔβǎ]   *ntsià'vǎ*   'toothless'

     d   [ⁿt͡sʲô̂ʰǒ]   *ntsiôjǒ*   'moonlight'

     e. [t͡sʲṹ̂ʔű̂]   *tsiu'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 article, 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* [lā̄ⁿt͡ʃĩ] 'lamb'. On the other hand, Eischens (2022) analyzes them as complex segments, since onset consonant clusters are banned in the language, but prenasalized 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 & 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. [t͡ʃĩ̂ʔᵑkì]   *tyi'nkì*   'acorn'

     b. [lĩᵑkō]   *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

---

[4] A third word with a prenasalized word-medial [ᵑk] (ĩ̄ᵑkà 'other') is likely derived from the numeral one (ĩ̄) and an "additive" clitic (=ka) (though see Mantenuto, 2020 for a more nuanced discussion of its meaning).
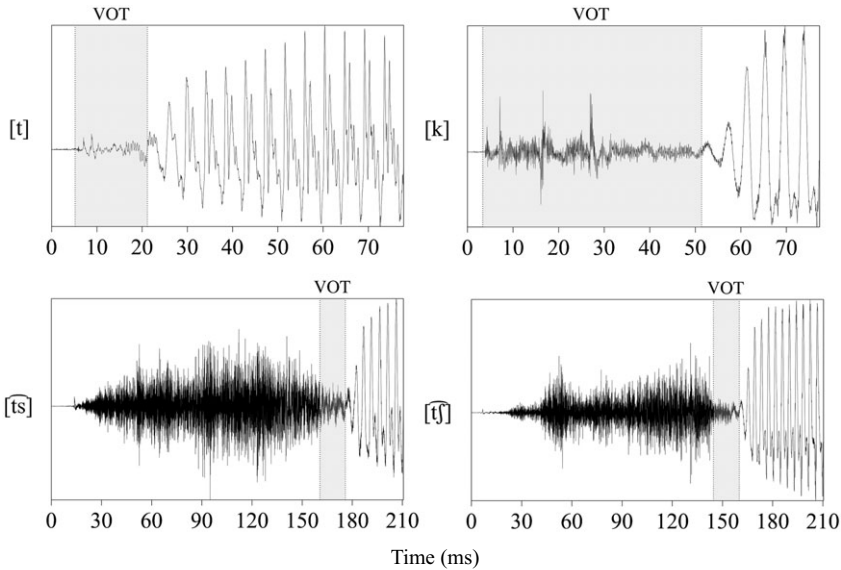
Figure 3. Representative examples of plain stops and affricates. Shaded portion shows VOT.

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 & 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 VOT and the internal structure of prenasalized stops and affricates. For plain stops, VOT was 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 & Whalen, 2017). Measurements were taken from stops and affricates with no palatalization or labialization. Examples of VOT measures are shown in Figure 3, which are taken from the words [táʰtě] ('sir/father'), [kînǐ] ('pig'), [t͡siʰkà] ('grasshopper'), and [t͡ʃîʳt͡ʃi] ('cricket'), respectively.

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 (cf. Cortés et al., 2023). Figure 4 below illustrates this sequencing of prenasalization (PN), weak voicing (WV),

---

[5] *Gualumbo* is itself a borrowing from the Otomí word *uadombo* (Bravo Vargas, 2014).
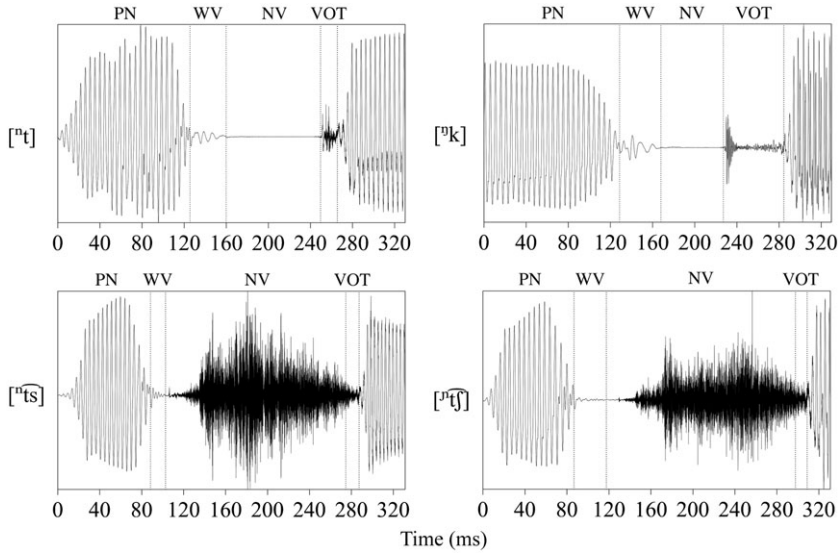
Figure 4. Representative examples of prenasalized stops and affricates. NV = no voicing; PN = prenasalized; VOT = VOT from stop burst or cessation of oral frication; WV = weak voicing.
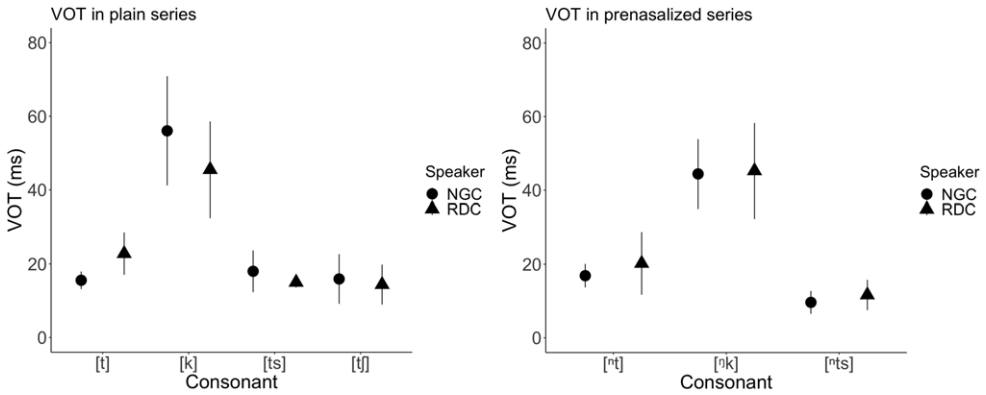
Figure 5. VOT by consonant type and speaker in the series of plain stops and affricates (left) and prenasalized stops and affricates (right). Error bars represent one standard deviation.
NGC [t͡ʃ] = 35, [k] = 51, [t] = 18, [t͡s] = 6, [ⁿt] = 18, [ᵑk] = 12, [ⁿt͡s] = 11.
RDC [t͡ʃ] = 18, [k] = 51, [t] = 15, [t͡s] = 6, [ⁿt] = 18, [ᵑk] = 15, [ⁿt͡s] = 17.

no voicing (NV), and a positive VOT. The examples are taken from the words [ⁿtiβi] ('beautiful'), [lí ᵑko] ('bud of the flower of the maguey cactus'), [ⁿt͡sî ʰkà] ('wide'), and [ᵖt͡ʃi ʰʃi] ('corn cob'), respectively. The example illustrating [ᵑk] is root-medial, since this segment never occurs root-initially.

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 [ᵖt͡ʃ] was excluded because of a

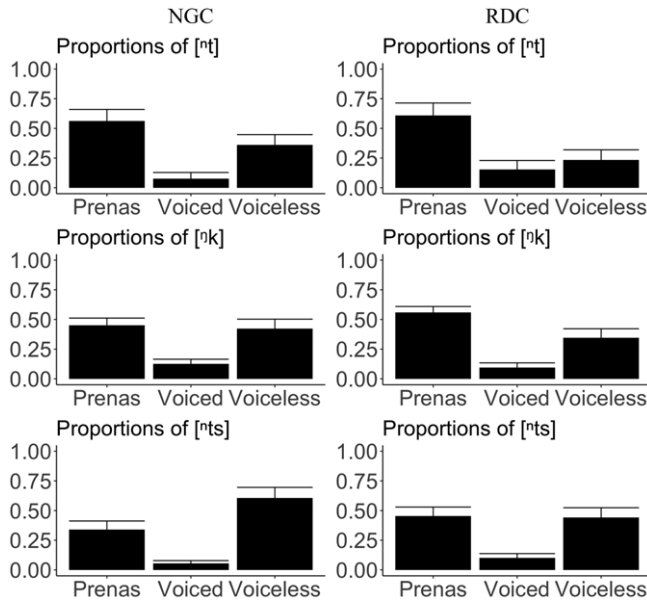**Figure 6.** Proportional duration of prenasalization, weak voicing, and voicelessness in prenasalized consonants for NGC (left) and RDC (right). Error bars show one standard deviation.

low number of tokens for analysis (three per consultant). The data were measured in Praat and illustrated using the ggplot package (Wickham, 2016) in R (R Core Team, 2013).

The VOT of [t], [t͡ʃ], and [t͡s] are all short lag, with averages around 20ms. As is usually the case with backer places of articulation (Lisker & 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] in English (Lisker & Abramson, 1967:6), as well as in other Mixtec languages (DiCanio et al., 2020; Cortés et al., 2023). 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 previously, prenasalized stops and affricates are made up of a sequence of PN, WV, and NV. The voiceless portion also includes a positive VOT measured from the offset of oral constriction to the beginning of periodic voicing. Figure 6 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.

Before moving on, it is worth noting that 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.

1.2 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* [mālì] 'godmother of one's child, or mother of one's godchild', *nánà* [nānà] '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).
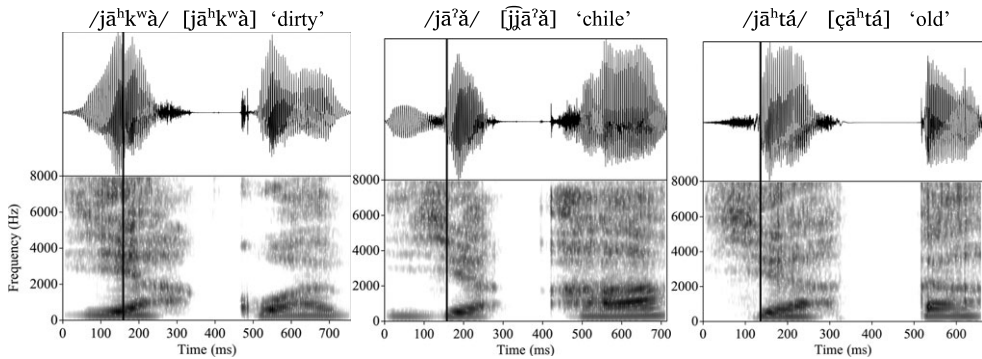

Figure 7. Waveforms and spectrograms showing four distinct realizations of /j/. Vertical black bar represents the offset of the sound.
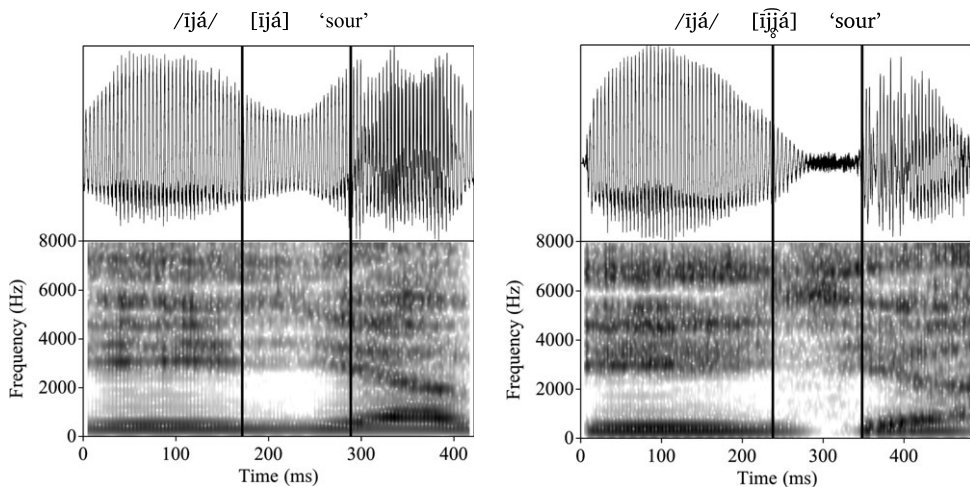

Figure 8. Waveforms and spectrograms showing two distinct realizations of /j/. Vertical black bars represent the onset and end of the sound.

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/.

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., 2023). 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]
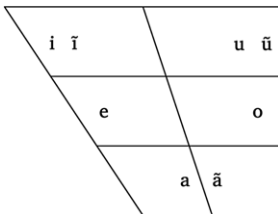
[kátá^{ʔ}ǎ]	*kátá'ǎ*	'fights'
[kéβà^{ʔ}ǎ]	*kévà'ǎ*	'wins'
[kítá^{ʔ}ǎ]	*kítá'ǎ*	'accompanies'
[kò-ká́^{ʔ}ã̀]	*kò-ká'àn*	'does not talk'
[kúká^{ʔ}nù īnī]	*kúká'nù ini*	'forgives' (lit. 'be big inside')
[kʷáà]	*kuáà*	'blind'
[kʷá́ã̀]	*kuáàn*	'yellow'
[kʷíì]	*kuîì*	'clear'
[kʷí̃ì̃]	*kuîn*	'striped'
[kūū]	*kuu*	'will die'
[kũ̄ũ̄]	*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.

---

[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.
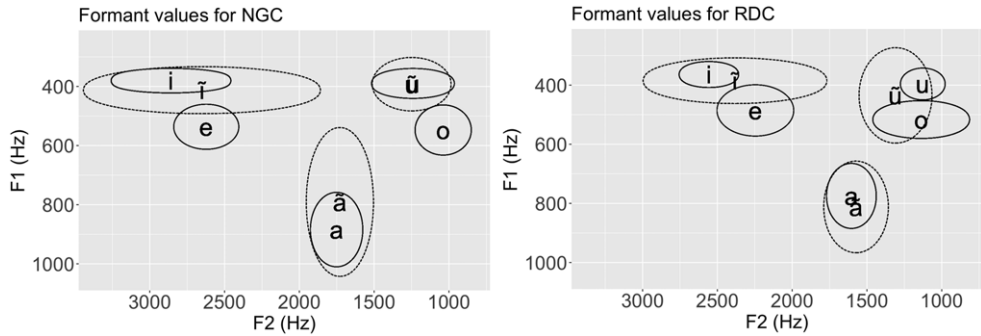
Figure 9. Plots of average formant values (Hz) for NGC and RDC. Ellipses show one standard deviation around the mean. Solid lines represent oral vowels; dotted lines represent nasal vowels.
Number of tokens for NGC: 69 [a], 36 [ã], 21 [e], 88 [i], 77 [ĩ], 84 [o], 30 [u], 32 [ũ].
Number of tokens for RDC: 75 [a], 39 [ã], 18 [e], 84 [i], 78 [ĩ], 87 [o], 36 [u], 33 [ũ].

Figure 9 displays two plots showing the average formant values for SMPM's five oral vowels and three nasal vowels for NGC and RDC. The graphs represent root-final vowels, where there is no phonation type contrast.

3. Prosodic features

3.1 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.

3.2 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., *kiá'mǐ* [kʲá²mǐ] 'squash'), and it never occurs phonemically in root-initial or root-final position in most Mixtec languages (though see Pankratz & 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 & 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 & 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. [jāā]   *yaa*   'white'
     b. [jâʰǎ]   *yájǎ*   'tongue'
     c. [jāˀǎ]   *ya'ǎ*   'chile'

(5) a. [kún͡tsʲēē]   *kúntsiee*   'puts up with'
     b. [n͡tsʲêʰě]   *ntsiêjě*   'strength'
     c. [n͡tsʲéˀé]   *ntsî'é*   'ground bean soup'

(6) a. [kʷíì]   *kuîì*   'clear (Adj)'
     b. [kʷîʰǐ]   *kuíjǐ*   'green'
     c. [kʷìˀī]   *kuì'i*   'fruit'

(7) a. [kōō]   *koo*   'welcome'
     b. [kòʰǒ]   *kòjǒ*   'snake'
     c. [kòˀǒ]   *kò'ǒ*   'plate' or 'bowl'

(8) a. [ⁿtúù]   *ntúù*   'a small black flying insect with a white face'
     b. [ⁿtùʰǔ]   *ntùjǔ*   'large, hard fruit seeds'
     c. [ⁿtūˀú]   *ntu'ú*   'fat (Adj)'

In addition, all three nasalized vowels can be contrastively breathy and glottalized (9–11). For clarity, we transcribe the underlying level tones as opposed to the surface falling tones in examples (17) and (18).

---

[7] We know of one exception to this, which is that the post-nominal demonstrative [hã̀ã́] ('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). The majority of residents of San Martín Peras are literate in Spanish (Instituto Nacional de Estadística y Geografía, 2020), so using *j* to represent breathiness seems to us to be a natural choice.

(9)   a.   [kʷáá̀]   *kuáàn*   'yellow'

    b.   [kʷâ̂ʰǎ]   *kuâjǎn*   'unmarried'

    c.   [kʷā̄ʔǎ]   *kua'ăn*   'leaves (V)'

(10) a.   [íí́]   *íín*   'hail'

    b.   [ì̀ʰǐ]   *íjǐn*   'skin' or 'leather'

    c.   [í́ʔì̄]   *í'ìn*   'temazcal (a type of sweat lodge)'

(11) a.   [ɲū̄ù̃]   *ñuù*   'town'

    b.   [ɲù̀ʰǔ]   *ñùjǔ*   'palm plant'

    c.   [ɲū̄ʔù̃]   *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* [t͡ʃ ì̀ʰkí] '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-ká́ʔà̀] 'does not talk' and *kúká'nù ini* [kúká́ʔnù 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 & 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 & Small, 1974:122–124; Macaulay, 1996:42; Gerfen & Baker, 2005; Herrera Zendejas, 2014:72–74; Becerra Roldán, 2019:112–116; Penner, 2019:254; Cortés et al., 2023:11–14), 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.

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

---

[9] Note that we do not represent preaspiration orthographically because it is predictable allophonic variation.
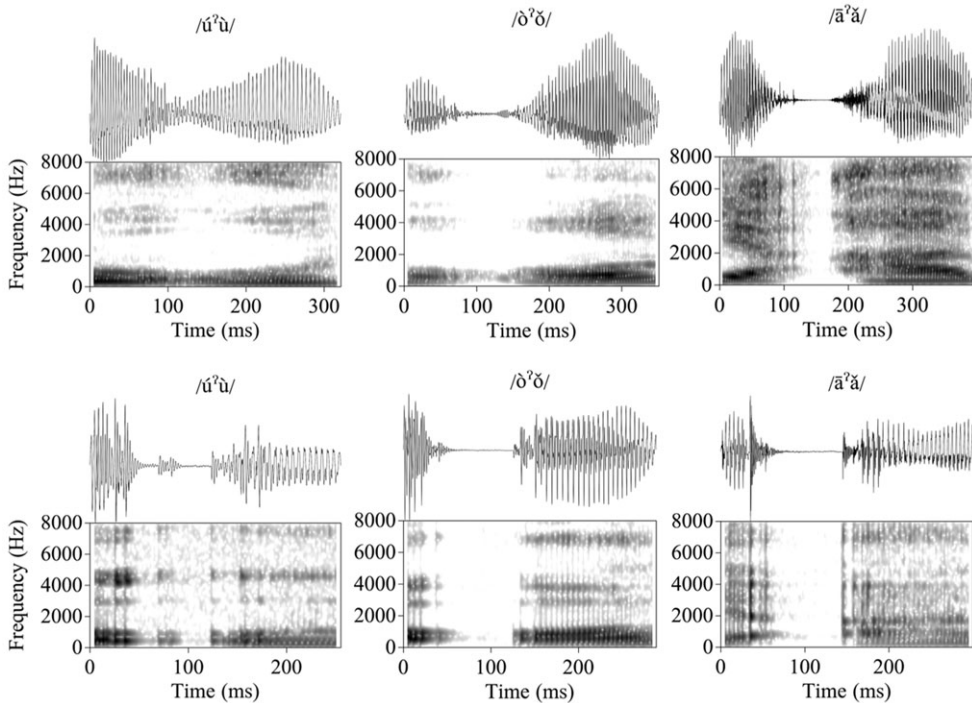
Figure 10. Individual tokens of glottalized vowels from the words /kúˀù/ ('sick'), /kòˀð̌/ ('plate'), and /jā̄ˀǎ/ ('chile pepper') from NGC (top) and RDC (bottom).

in breathy vowels. This can be seen in Figure 12, which shows a vowel followed by a preaspirated [t].

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., 2023 for a recent example on another Mixtec language). Following similar methods in Cortés et al. (2023) 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 from 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., [VˀV] 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
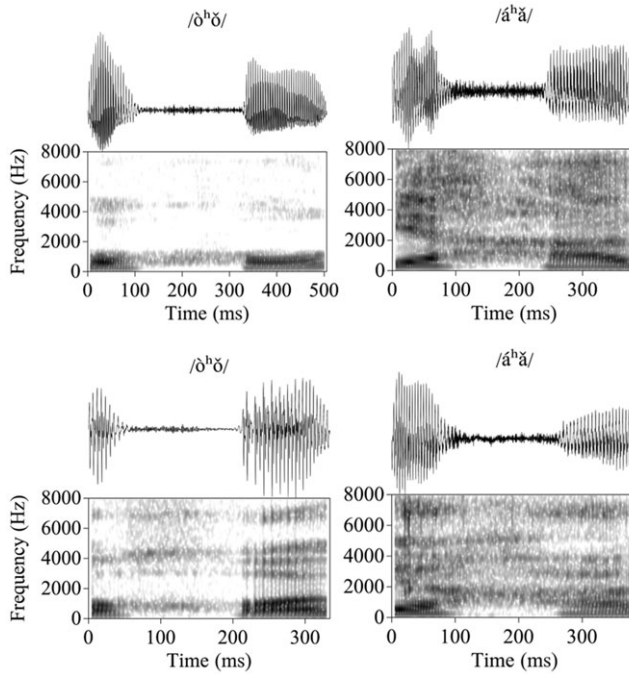
Figure 11. Individual tokens of breathy vowels from the words /kòʰǒ/ ('snake') and /jáʰǎ/ ('tongue') from NGC (top) and RDC (bottom).
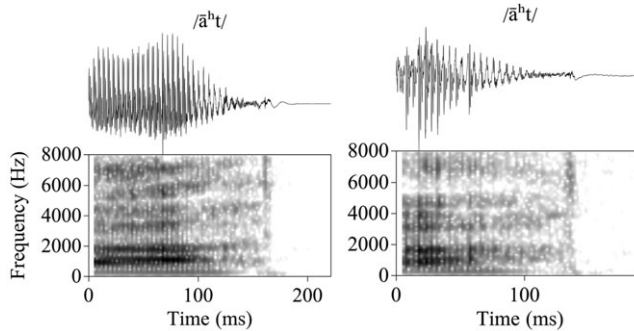
Figure 12. Individual tokens of vowels followed by preaspiration from the word /āʰtū/ ('bitter') from NGC (left) and RDC (right).

to interspeaker variation. The plots in Figure 14 show the SoE contours for vowels and following preaspiration (e.g., the underlined portion of [t͡ʃìʰkí] ('prickly pear fruit')) alongside that of vowels and following prenasalization (e.g., the underlined portion of [lí̃ⁿko] ('bud of the flower 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.
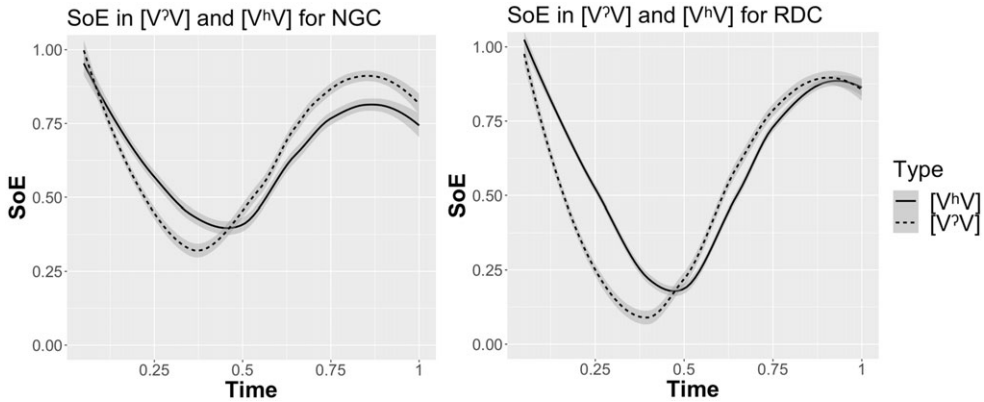
Figure 13. SoE for [VʰV] (solid) and [VˀV] (dashed) sequences. Lines are smoothed LOESS regression lines, and gray bars represent a 95% confidence interval around the regression line. NGC [VʰV] = 57, [VˀV] = 58. RDC [VʰV] = 66, [VˀV] = 63.
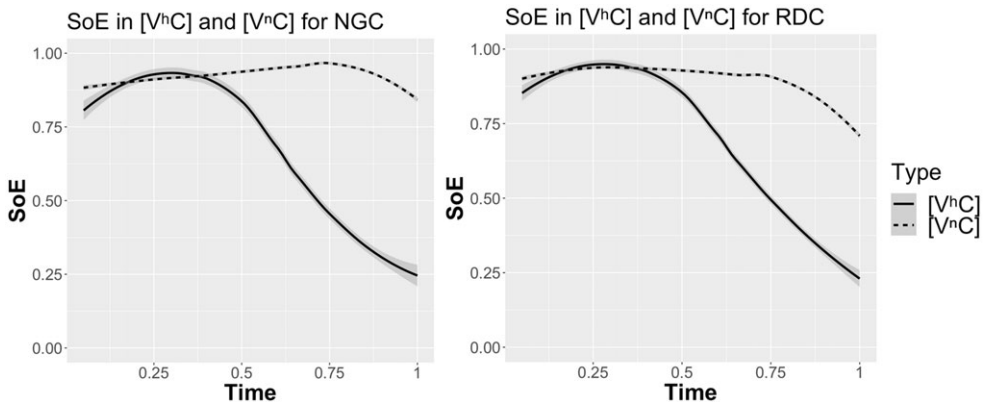
Figure 14. SoE for vowels followed by preaspiration (solid) and prenasalization (dashed). Lines are smoothed LOESS regression lines, and gray bars represent a 95% confidence interval around the regression line. NGC [VʰC] = 84, [VⁿC] = 63. RDC [VʰC] = 66, [VⁿC] = 66.

3.3 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 [táʰtà] ('señor') and the HL-L root [ntsîʰkà] ('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. [ʃǐjò]   *xǐyò*   'dress'

e. [kînì]   *kînì*   'pig'

(13) a. [kʷáâ̌ǎ]   *kuáàn*   'yellow'

b. [sāà]   *saà*   'bird'

c. [ʃàà]   *xàà*   'chin'

d. [ɲǔ̌ù]   *ñǔù*   'night'

e. [(jīβā) t�/ jôǒ]   *(yiva) tiôǒ*   'a type of edible plant' (in Spanish, alache)

(14) a. [máˀ̂ǎ̃]   *máʼà*   'raccoon'

b. [tāˀmã̀]   *taʼmà*   'hillside'

c. [ɲà̃ˀmǎ]   *ñàʼmǎ*   'smooth'

d. [mǎ̃ˀnã̀]   *mǎʼnà*   'tired'

e. [sʲâˀǎ]   *siâʼǎ*   'Tecomaxtlahuaca (a town)'

(15) a. [táʰtà]   *tátà*   'father' or 'sir'

b. [ūʰsù]   *usù*   'deer'

c. [t͡sìkàʰtsʲà]   *tsìkàtsià*   'round (Adj)'

d. [jùtǎʰtá]   *yùtǎtá*   'mirror'

e. [ⁿt͡sî̂ʰkà]   *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 level tones, as represented in the difference between the phonemic transcription in slashes and the allophonic transcription in square brackets.

(16) a. /níˊʰǐ/   [nî̂ʰǐ]   *nîjǐ*   'skinny'

b. /nīˉʰǐ/   [nî̂ʰǐ]   *nîjǐ*   'ear of corn'

c. /nìˋʰǐ/   [nìˋʰǐ]   *nìjǐ*   '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 (section 3.4), word-final rising tones flatten to level low tones when followed by high tones across a word boundary. Example (17) shows that the underlying final rise of [kòʰǒ] ('snake') surfaces faithfully before /nῑ̄ʰǐ/ ('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 /nίʰǐ/ ('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.

(17) ʃàʰʃῑ   kòʰǒ   nῑ̄ʰǐ   (18) ʃàʰʃῑ   kòʰò   nίʰǐ   nῑ̄ʰǐ

    ate   snake   ear.of.corn   ate   snake   skinny   ear.of.corn

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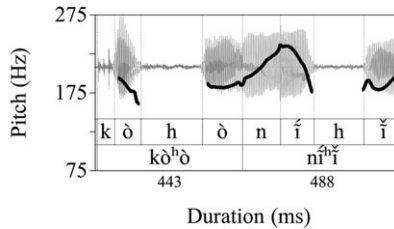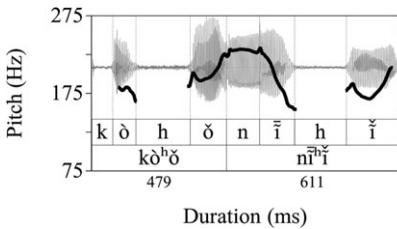
We know of only one instance of a non-falling high tone on the first mora of a contrastively breathy vowel, shown in (22). This word is likely derived from the root [tà̀ʰā́] ('for there to be an earthquake'). Importantly, the following tone is mid, not rising. There appear to be no surface level high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

(19)   [tá́ʰā̄]   *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 (though see Peters (2018) and Peters & Mendoza (2020) for examples). 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 a glottalized vowel (22), or is followed by non-contrastive [h] (23).

(20)   a.   [kānί]   *kaní*   'slippery'

     b.   [kōlō]   *kolo*   'turkey'

     c.   [kōnῑ̀]   *konì*   'know (POT)'

     d.   [(ñà̀) jɪβɪ̌]   *(ñà) yivǐ*   'person'

---

[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).

(21) a. [t͡sɪ̀ɪ́]   *tsìín*   'rat'

      b. [ɪ̄ɪ́]   *iin*   'one'

      c. [ʃàà]   *xàà*   'chin'

      d. [(jīβā) t͡ʲôŏ]   *(yiva) tiôŏ*   'a type of edible plant' (in Spanish, alache)

(22) a. [ⁿt͡sʲéʔé]   *ntsié'é*   'ground bean soup'

      b. [βē̄ʔē]   *ve'e*   'house'

      c. [ʃè̀ʔè]   *xè'è*   'garbage'

      d. [ʃē̄ʔě]   *xe'ě*   'ring'

(23) a. [t͡ʃúʰtú]   *tyútú*   'full (Adj)'

      b. [lē̄ʰsō]   *leso*   'rabbit'

      c. [ū̄ʰsù]   *usù*   'deer'

      d. [jû̂ʰkǔ]   *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. [t͡ʃè̀ʰē]   *tyèje*   'male (Adj)'

      b. [nɪ̂ʰɪ̌]   *nîjǐn*   'skinny'

Figure 15 shows 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 words, 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].

3.4 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 & 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
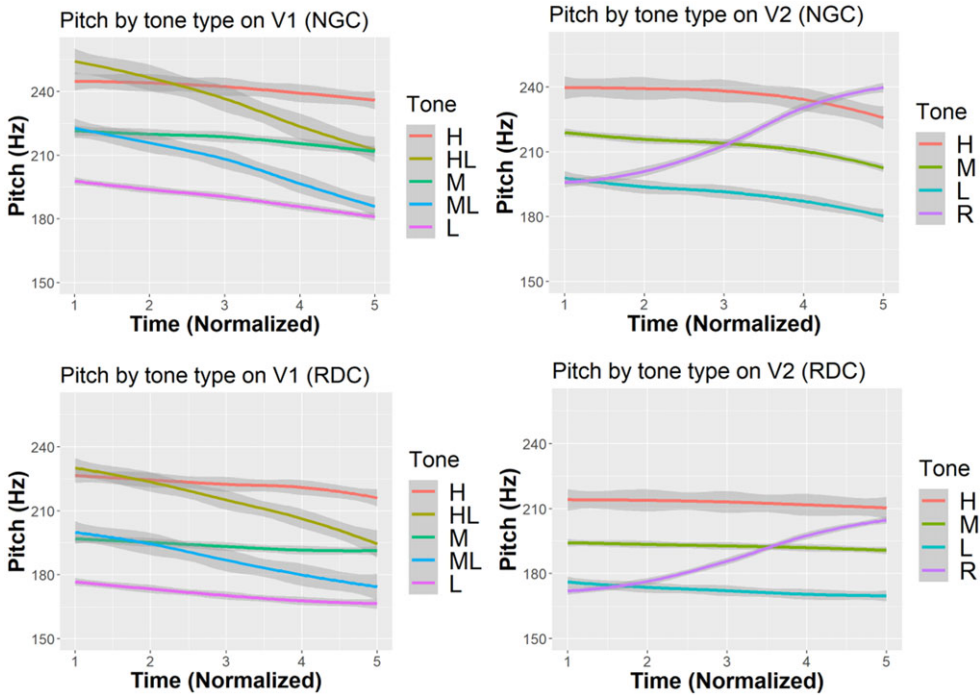
Figure 15. (Colour online) Pitch (Hz) by tone type on V1 (left) and V2 (right) for consultants NGC (top) and RDC (bottom). Colored lines are LOESS regression curves, and gray bars around the lines represent 95% confidence intervals for the curve.
Number of tokens for NGC V1: H = 51, HL = 33, M = 93, 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.

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 [nǜʰnǐ] '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. nǜʰnǐ jāā
          corn   white
          'white corn'

     b. nǜʰnì kʷáà̀
          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 [káˀnǜ] '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.   ī̄ī̄   lā̃ⁿt͡ʃī̄   káˀnǜ

          one   lamb   big

          'A big lamb'

      b.   ī̄ī̄   ū̃ʰtù   kǎˀnǜ

          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 /jũ̃ʰkǔ/ ('mountain/wilderness') has an underlying, final rising tone (23d). This rising tone becomes a low tone when followed by the initial high tone of [káˀnǜ] ('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   jũ̃ʰkù   káˀnǜ

      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.

4.1 Orthographic version

Tàtyĭ nórtè xí'ìn tsìkàntsìjĭ kìxaà nà kâ'àn nà yòó nà ntakù tyáákà, tá nìyà'ǎ iin rà xíka ìníì ñuù yivĭ xí'ìn tsiàà kǎ'nù ítiví rà. Nìka'àn nà yòó nà kevà'a kasa ntúxa nà xí'ìn rà tàvǎ nà míí tsiàà ítiví rà, rà kúu nà nà ntakù nùjǔ ntsìkúu ñà'a ñuù yivĭ. Míí tàtyĭ ña nórtè ntakù và'a tsìvià xí'ìn ntsìkúu ntsiêjâ, só nú kuà'á và'ǎ tsívià, kuà'á tyáákà ná tìjviâ tsiàà kǎ'nù míí rà xíka ìníì ñuù yivĭ. Nǔ ntsí'i kùntaà ini míí tàtyĭ nórtè kòníkùù tyiñà. Sáá nàye'è tsìkàntsìjĭ xí'ìn ntsìkúu ñà i'níà. Kamà và'a tàvǎ míí rà xíka ìníì ñuù yivĭ tsiàà kǎ'nù ítiví rà. Nǔ kùndaà ini tàtyĭ tsìkàntsìjĭ yá ntakù tyáákà nùjǔ ntsìkúu ñà'a.

4.2 Transcription

(1) tàʰt͡ʃǐ nőɾtè ʃíʔĩ t͡sìkàⁿt͡sìʰĩ kìʃāà   nǎ̰ kâ̰ʔǎ̰   nǎ̰ jòó
     wind north with sun   begin.COMPL 3PL talk.CONT 3PL who

     nǎ̰   ⁿtāʰkù   t͡ʃáá=kà
     3PL   strong   more=ADD

     "The North Wind and the Sun were disputing which was the stronger,"

(2) tá   nǐ-jà̰ʔǎ̰   ĩ̄ĩ̄   rà   ʃíʰkā   ĩnĩ̂ĩ   ɲṵ̃ṵ̃ jīβǐ   ʃíʔĩ   t͡sʲàà
     when COMPL-walk one 3SG.M walk.CONT around town person with clothes

     kǎ̰ʔnṵ̌   ítīβí   rà
     big   wrap.CONT   3SG.M

     "when a traveler came along wrapped in a warm cloak."

(3) nǐ-kã̰ʔā̰   nǎ̰   jòó   nǎ̰   kēβà̰ʔa   kāʰsā   ⁿtúʰʃā   nǎ̰   ʃíʔĩ
     COMPL-talk 3PL who 3PL win.POT make strength 3PL with

     rà   tàβǎ   nǎ̰ mǐ́ǐ t͡sʲàà   ítīβí rà
     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) rà   kūū   nǎ̰   nǎ̰   ⁿtāʰkù   nṵ̀ʰṵ̌   t͡sìkúū   ɲà̰ʔā̰   ɲṵ̃ṵ̃ jīβǐ
     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) mǐ́ǐ   tàʰt͡ʃǐ   ɲà̰   nőɾtè   ⁿtāʰkù   βà̰ʔā   t͡sìβī=à   ʃíʔĩ
     FAM   wind 3SG.N north strong well blow.COMPL=3SG.N with

     ⁿt͡sìkúū   t͡sʲêʰě=à
     all   strength=3.SG.N

     "Then the North Wind blew as hard as he could,"

(6) só   nṹ   kʷàˀá   βàˀā   t͡síβī̄=à
     but   when   much   well   blow.CONT=3SG.N
     "but the more he blew"

(7) kʷàˀá   t͡ʃáákà   nã́   tìʰβĭ=à   t͡sʲàà   kǎˀnṳ̀   mĩ́í   ɾà
     much   more   MOOD   wear.COMPL=3SG.N   clothes   big   FAM   3SG.M

     ʃíʰkā   ìnĩ̂ĭ   ɲṵ̄ṳ̀   jīβĭ
     walk.CONT   around   town   person
     "the more closely did the traveler fold his cloak around him;"

(8) nǔ   ⁿtsíˀī   kùⁿtāà   īnī̄   mĩ́í   tàʰt͡ʃĭ nő́ɾtè
     when   finish.CONT   realize.COMPL   inside   FAM   windnorth

     kò-nĩ́-kùù   t͡ʃĩ̄ñṵ̄=à
     NEG-COMPL-able   work=3SG.N
     "and at last the North Wind gave up the attempt."

(9) sáá   nàjē̄ˀè   t͡sìkàⁿt͡sìʰĭ   ʃíˀī̆   ⁿt͡sìkúū   ñæ̀   iˀnĩ́=à
     so   shine.COMPL   sun   with   all   3SG.N   heat=3SG.N
     "Then the Sun shone out warmly,"

(10) kāmæ̀   βàˀā   tàβǎ   mĩ́í   ɾà   ʃíʰkā   ìnĩ̂ĭ   ɲṵ̄ṳ̀   jīβĭ
     fast   well   take.off.COMPL   FAM   3SG.M   walk.CONT   around   town   person

     t͡sʲàà   kǎˀnṳ̀   ítīβí   ɾà
     clothes   big   wrap.CONT   3SG.M
     "and immediately the traveler took off his cloak."

(11) nǔ   kùⁿdāà   īnī̄   tàʰt͡ʃĭ
     when   realize.COMPL   inside   wind
     "And so the North Wind was obliged to confess,"

(12) t͡sìkàⁿt͡sì^hǐ já ⁿtā^hkù t͡ʃáá=kà nǔ^hǔ ⁿt͡sìkúū ɲà̰ˀā̰

sun 3SG.N strong more=ADD over all thing

"that the Sun was the stronger of the two."

Abbreviations We use the following abbreviations (adopting, where possible, the Leipzig Glossing Rules): 1 = first person, 3 = third person, ADD = Additive, COMPL = Completive aspect, CONT = Continuative aspect, DEM = Demonstrative, EXCL = Exclusive, F = Feminine, FAM = Familiar referent, FOC = Focus, LIQ = Liquid pronoun, M = Masculine, MOOD = Mood marker, N = Neuter, NEG = Negative, PL = Plural, POT = Potential aspect, SG = Singular.

Appendix

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 organized by folder section name and numbered in the order that they occur in their respective folders. For ease of cross-referencing, any example or figure number is also included following an underscore, if applicable. Any examples for which carrier phrases were not available are excluded from the list, and their number in each folder is skipped.

1. Consonants

(01) ī̄ī̄ pá^htò lō^ˀō

one duck small

'A small duck.'

(02) ī̄ī̄ ^mpáà lō^ˀō

one godfather small

'A small godfather.'

(03)   ī̄ī   tá^htà   lō^ʔō   rà
         one   sir   small   3SG.M
         'He is a small man.'

(04)   ī̄ī   rà   t͡sʲà^hā   ntá^ʔβī
         one   3SG.M   man   poor
         'A poor man.'

(06)   ⁿt͡ʲā^ʔmǐ   t͡sʲá^ʔjī
         radish   rotten
         'A rotten radish'

(07)   kōnī̄   kà^hkū   ī̄ī   rà   lō^ʔō
         yesterday   born.COMPL   one   3SG.M   small
         'A boy was born yesterday.'

(08)   rà   kì^hʃī   ī̄īⁿkà ɲū̄ù̃
         3SG.M   come.COMPL   other town
         'He came from another town.'

(09)   lē̄^hsō   k^wá^ʔà   lō^ʔō
         rabbit   red   small
         'A small, red rabbit.'

(10)   ī̄ī   k^ʲá^ʔmǐ   lō^ʔō
         one   squash   small
         'A small squash.'

(11)   ī̄ī   mɑ́lì   lō^ʔō
         one   godmother   small
         'A small godmother.'

(12)   ī̄ī   nɑ́nɑ̀   lō^ʔō
         one   mother/woman   small
         'A small woman.'

(13)   ī̄ī̃   ɲã̄nī̃   ēstéβà
        one   brother   Estéban
        'A brother of Esteban.'

(14)   ī̄ī̃   ɾà   lō̄ˀō̄
         one   3SG.M   small
         'A boy.'

(15)   ī̄ī̃   sā̄ˀmǎ   lō̄ˀō̄
         one   tortilla.cloth   small
         'A small tortilla cloth.'

(16)   ñ̃ū̃ū̃   sʲâ̄ˀǎ   βà
        town   Tecomaxtlahuaca   FOC
        'The town of Tecomaxtlahuaca.'

(17)   ī̄ī̃   ʃàà   lō̄ˀō̄
         one   chin   small
         'A small chin'

(18)   ī̄ī̃   t͡sʲǎ̄ˀà   plástīkō̄
         one   jug   plastic
         'A plastic jug'

(19)   lē̄ʰsō̄   ⁿt͡sʲà̄ˀβǎ   lō̄ˀō̄
        rabbit   toothless   small
        'A small toothless rabbit'

(20)   kʷá̄ˀà   t͡ʃááɣà   kì̄ʰʃā̄   sū̄fɾí̄ì   ⁿtù
        much   more   do.COMPL   suffer   1PL.EXCL
        'We (excl.) began to suffer much more.'

(21)   ī̄ī̃   ⁿt͡ʃī̄ʰʃǐ   lō̄ˀō̄
         one   corncob   small
         'A small corncob.'

(22)   lēʰsō   βálí   jāˀá
        rabbit   small.PL   orange
        'Small, orange rabbits.'

(23)   lēʰsō   jāā   lōˀō
        rabbit   white   small
        'A small, white rabbit.'

(24)   ĩ̄ĩ̄   lā͡ⁿtʃĩ   lōˀō
        one   lamb   small
        'A small lamb.'

1.1 Obstruents

(01)   kʷáˀà   t͡ʃááɣà kìʰʃā   sūfríî ⁿtù
        much   more   do.COMPL   suffer 1PL.EXCL
        'We (excl.) began to suffer much more.'

(03)   ĩ̄ĩ̄   kʲáˀmǐ   lōˀō
        one   squash   small
        'A small squash.'

(04)   lēʰsō   kʷáˀà   lōˀō
        rabbit   red   small
        'A small, red rabbit.'

(05)   ĩ̄ĩ̄   ñã̀   jɪβɪ̌   lōˀō
        one   3PL   person   small
        'A small person.'

(06)   ɲũ̄ũ̀   sʲâˀǎ   βà
        town   Tecomaxtlahuaca   FOC
        'The town of Tecomaxtlahuaca.'

(09_1a)   ĩ̄ĩ̄   t͡ʃìʰkí   lōˀō
           one   prickly.pear.fruit   small
           'A small prickly pear fruit.'

(10_1b)   ī̄ī̀   ⁿt͡ʃî̀ʰʃĭ   lō̄ˀō̄
            one   corncob   small
            'A small corncob.'

(11_2a)   ī̄ī̀   ⁿts͡î̀ʰkă   lō̄ˀō̄
            one   banana   small
            'A small banana.'

(12_2b)   íjō̄   ⁿts͡ʲê̂ʰě   ɾà
            exist.CONT   strength   3SG.M
            'He is strong'

(13_2c)   lē̄ʰsō̄   ⁿts͡ʲà̀ˀβă   lō̄ˀō̄
            rabbit   toothless   small
            'A small, toothless rabbit.'

(14_2d)   ⁿts͡ʲô̂ʰǒ
            moonlight
            'Moonlight' (no carrier phrase)

(15_2e)   ts͡î̀ī̀   ts͡ʲū̂ˀṹ   t͡ʃúʰtū
            bite.COMPL   venomous.spider   cat
            'The venomous spider bit the cat.'

(16)   ī̄ī̀   lā̄ⁿt͡ʃĭ   lō̄ˀō̄
            one   lamb   small
            'A small lamb.'

(17_3a)   t͡ʃî̀ˀⁿkì
            acorn
            'Acorn' (no carrier phrase)

(18_3b)   ī̄ī̀   lìⁿkō̄   íʰts͡ʲă   lō̄ˀō̄
            one   maguey.bud   soft/young   little
            'A small, soft/young bud of the flower of the maguey cactus.'

(19_fn4)   ɾà   kìʰʃī   īⁿkà ɲū̃ū̃
             3SG.M   come.COMPL   other town
             'He came from another town.'

1.2 Sonorants

(01)   ī̃ī̃   málì   lō̰ʔō
         one   godmother   small
         'A small godmother.'

(02)   ī̃ī̃   nã́nã̀   lō̰ʔō
         one   mother/woman   small
         'A small woman.'

(03)   ī̃ī̃   ɲã̄nī̃   ēstéβà
         one   brother   Estéban
         'A brother of Esteban.'

2. Vowels

(01)   nã̀   kátáʔǎ   nã̀
         3PL   fight.CONT   3PL
         'They are fighting.'

(02)   nã̀   kéβā̰ʔǎ   nã̀
         3PL   win.CONT   3PL
         'They are winning.'

(03)   nã̀   kítáʔǎ   nã̀
         3PL   accompany.CONT 3PL
         'They are accompanying.'

(04)   nã̀   káā   kò-ká́ʔã̀   nã̀
         3PL   DEM   NEG-speak.CONT   3PL
         'They do not speak.'

(05)   nǎ̰   káā   kúkáˀnṵ̄   īnḭ̄   nǎ̰
3PL   DEM   forgive.CONT   inside   3PL
'They forgive.'

(06)   lḛ̄ʰsō   kʷáà   lōˀō
rabbit   blind   small
'A small, blind rabbit.'

(07)   lḛ̄ʰsō   kʷá̰ằ   lōˀō
rabbit   yellow   small
'A small, yellow rabbit.'

(08)   ḭ̄ḭ̄   tsḭ̀kʷḭ̂   kʷíì   lōˀō
one   water   clear   little
'Some water that is a little bit clear.'

(09)   lḛ̄ʰsō   kʷḭ̂ḭ̂   lōˀ ō
rabbit   striped   little
'A small, striped rabbit.'

(10)   ɾà   kūū   ɾà
3SG.M   die.POT   3SG.M
'He will die.'

(11)   ɾá   kṵ̄ṵ̄   ɾá
3SG.LIQ   fall.POT   3SG.LIQ
'It (rain) will fall.'

3.2 Phonation type

(01)   ḭ̄ḭ̄   kʲáˀmǐ   lōˀō
one   squash   small
'A small squash.'

(02)   tsìnǎ̰   hằá̰   ɾí   tsḭ̀ḭ̄   lḛ̄ʰsō
dog   DEM   3SG.AML   bite.COMPL   rabbit
'That dog bit the rabbit'

(03_4a)   lē^hsō   jāā   lō^ʔō

             rabbit   white   small

             'A small, white rabbit.'

(04_4b)   ī̄ī̄   jâ^hǎ   lō^ʔō

             one   tongue   small

             'A small tongue.'

(05_4c)   ī̄ī̄   jā^ʔǎ   lō^ʔō

             one   chile.pepper   small

             'A small chile pepper.'

(06_5a)   kúⁿt͡sʲēē   rà

             put.up.with.CONT   3SG.M

             'He puts up with it.'

(07_5b)   íjō   ⁿt͡sʲê^hě   rà

             exist.CONT   strength   3SG.M

             'He is strong'

(08_5c)   ʃì^hʃī   rà   lō^ʔō   ⁿt͡sʲé^ʔé   kōnī̄

             eat.COMPL 3SG.M small ground.bean.soup yesterday

             'The boy ate ground bean soup yesterday.'

(09_6a)   ī̄ī̄   tsì̀k^wĩ̀ĩ   k^wíì   lō^ʔō

             one   water   clear   little

             'Some water that is a little bit clear.'

(10_6b)   lē^hsō   k^w î̂^hĭ   lō^ʔō

             rabbit   green   small

             'A small, green rabbit.'

(11_6c)   ī̄ī̄   k^wì^ʔĩ   lō^ʔō

             one   fruit   small

             'A small fruit.'

(12_7a)   kōō   βē²ē
           welcome   house
           'Welcome to the house.'

(13_7b)   ĩ̄ĩ̄   kò^hǒ   lō²ō
           one   snake   small
           'A small snake.'

(14_7c)   ĩ̄ĩ̄   kò²ǒ   lō²ō
           one   plate   small
           'A small plate.'

(15_8a)   ĩ̄ĩ̄   ⁿtúù   lō²ō
           one   ntuu   small
           'A small *ntuu.*' (type of insect)

(16_8b)   ĩ̄ĩ̄   ⁿtù^hǔ   lō²ō
           one   fruit.seed   small
           'A small fruit seed.'

(17_8c)   rà   ⁿtū²ú   rà
           3SG.M   fat   3SG.M
           'He is fat.'

(18_9a)   lē^hsō   kʷá̰ā̰   lō²ō
           rabbit   yellow   small
           'A small, yellow rabbit.'

(19_9b)   rà   kʷâ̰^hǎ   rà
           3SG.M   unmarried   3SG.M
           'He is unmarried.'

(20_9c)   rà   kʷá̰²ǎ   rà
           3SG.M   leave.CONT   3SG.M
           'He leaves.'

(21_10a)   ī̄ī̄   í̋í̋   lō̄ʔō̄
            one   hail   small
            'Some small hail.'

(22_10b)   ī̄ī̄   î̂ʰǐ   lō̄ʔō̄
            one   skin   small
            'A little bit of skin/leather.'

(23_10c)   ī̄ī̄   í̃ʔĩ   lō̄ʔō̄
            one   temazcal   small
            'A small temezcal.' (a type of sweat lodge)

(24_11a)   ī̄ī̄   ɲṹ̃ũ̀   lō̄ʔō̄
            one   town   small
            'A small town.'

(25_11b)   ī̄ī̄   ɲũ̀ʰǔ   lō̄ʔō̄
            one   palm.plant   small
            'A small palm plant.'

(26_11c)   ī̄ī̄   ɲū̃ʔũ̀   lō̄ʔō̄
            one   fire   small
            'A small fire.'

(27)   ī̄ī̄   táʰtà   lō̄ʔō̄   ɾà
        one   father   small   3SG.M
        'He is a small man.'

(28)   ī̄ī̄   t͡ʃĩʰkí   lō̄ʔō̄
        one   prickly.pear.fruit   small
        'A small prickly pear fruit.'

(29)   nã̀   káā   kò-ká̋ʔã̀   nã̀
        3PL   DEM   NEG-speak.CONT   3PL
        'They do not speak.'

(30)   nằ   káā   kúká ˀnū̄   īnī̃   nằ
         3PL DEM   forgive.CONT   inside 3PL
         'They forgive.'

(43)   ī̃ī̃   t͡ʃì ʰkí   lōˀō
         one   prickly.pear.fruit   small
         'A small prickly pear fruit.'

(44)   ī̃ī̃   lìⁿkō   í ʰt͡sʲ ̌ă   lōˀō
         one   maguey.bud   soft/young   little
         'A small, soft/young bud of the flower of the maguey cactus.'

3.3 Lexical tone

(01)   ī̃ī̃   táʰtà   lōˀō   rà
            one   sir   small   3SG.M
            'He is a small man.'

(02)   ī̃ī̃   βēˀē   ⁿtsî ʰkà   lōˀō
            one   house   wide   small
            'A small, wide house.'

(03_12a)   ī̃ī̃   nã́nằ   lōˀō
            one   mother/woman   small
            'A small woman.'

(04_12b)   ʃínĩ̄ ɲű̃ˀũ̀   kōnì̃   rà   ñá̃
            must   know.POT   3SG.M   3SG.F
            'He has to know her.'

(05_12c)   ī̃ī̃   t͡sìnằ   lōˀō
            one   dog   small
            'A small dog.'

(06_12d)   ī̃ī̃   ʃíjò   lōˀō
            one   dress   small
            'A small dress.'

(07_12e)   ī̄ī̆   kînǐ   lōˀ̄ō
one   pig   small
'A small pig.'

(08_13a)   lē̄ʰsō   kʷá̆̀ằ   lōˀ̄ō
rabbit   yellow   small
'A small, yellow rabbit.'

(09_13b)   ī̄ī̆   sāà   lōˀ̄ō
one   bird   small
'A small bird.'

(10_13c)   ī̄ī̆   ʃàà   lōˀ̄ō
one   chin   small
'A small chin.'

(11_13d)   nì-jàˀ̄ā   ɲŭ̆ŭ̆   βī͡ʰts͡ī
COMPL-arrive   night   now
'It is now nighttime.'

(12_13e)   ʃî̄ʰʃ=î̂   jīβā   tʲôǒ   kōnī̄
eat.COMPL=1SG   plant   *alache*   yesterday
'I ate *yiva tioo* yesterday' (a type of edible plant, in Spanish, alache).

(13_14a)   ī̄ī̆   má̆ˀ̄ằ   lōˀ̄ō
one   raccoon   small
'A small raccoon.'

(14_14b)   ī̄ī̆   tāˀ̄mằ   lōˀ̄ō
one   hillside   small
'A small hillside.'

(15_14c)   ī̄ī̆   sāà   lōˀ̄ō   mî́í̄   ɲằˀ̄mǎ   ɾí
one   bird   small   FAM   smooth   3SG.AML
'A small bird that is smooth.'

(16_14d)   ī̄ī̄   mǎˀnằ   lō̄ˀō̄
one   tired   small
'Someone who is a little tired.'

(17_14e)   ɲū̄ṳ̀   sʲâˀǎ   βà
town   Tecomaxtlahuaca   FOC
'The town of Tecomaxtlahuaca.'

(18_15a)   ī̄ī̄   táʰtà   lō̄ˀō̄   ɾà
one   father   small   3SG.M
'He is a small man.'

(19_15b)   ī̄ī̄   ū̄ʰsù   lō̄ˀō̄
one   deer   small
'A small deer.'

(20_15c)   βē̄ˀē̄   tsɪ̀kàʰtsʲà   lō̄ˀō̄
house   round   small
'A small, round house.'

(21_15d)   ī̄ī̄   jùtǎʰtá   lō̄ˀō̄
one   mirror   small
'A small mirror.'

(22_15e)   ī̄ī̄   βē̄ˀē̄   ⁿtsî̄ʰkà   lō̄ˀō̄
one   house   wide   small
'A small, wide house.'

(23_16a)   lē̄ʰsō̄   nî̂ʰǐ   lō̄ˀō̄
rabbit   skinny   small
'A small, skinny rabbit.'

(24_16b)   ī̄ī̄   nî̂ʰǐ   lō̄ˀō̄
one   ear.of.corn   small
'A small ear of corn.'

(25_16c)   íjō   lōˀ̄ō   nȉ̤ʰ̌ɪ̌   βɪ̄ʰtsɪ̄
             exist.CONT   small   blood   now
             'There is now a little bit of blood.'

(28_19)   ῑ̄ɪ̄   tá̤ʰā̄   lōˀ̄ō
             one   earthquake   small
             'A small earthquake.'

(29_20a)   lē̤ʰsō   kānɪ̄́   lōˀ̄ō
             rabbit   slippery   small
             'A small, slippery rabbit.'

(30_20b)   ῑ̄ɪ̄   kōlō   lōˀ̄ō
             one   turkey   small
             'A small turkey.'

(31_20c)   ʃɪ̄́nɪ̄ɲṳ̋ˀ̰ṵ̄ kōnȉ   rà   ɲá̤
             must   know.POT   3SG.M   3SG.F
             'He has to know her.'

(32_20d)   ῑ̄ɪ̄   nȁ̰   jɪ̄βɪ̌   lōˀ̄ō
             one   3PL   person   small
             'A small person.'

(33_21a)   ῑ̄ɪ̄   tsȉ̤ɪ̄́   lōˀ̄ō
             one   rat   small
             'A small rat.'

(34_21b)   íjō   ῑ̄ɪ̄   lē̤ʰsō
             exist.CONT   one   rabbit
             'There is a rabbit.'

(35_21c)   ῑ̄ɪ̄   ʃàà   lōˀ̄ō
             one   chin   small
             'A small chin.'

(36_21d)   ∫ìʰ∫=î   jīβā   tʲôǒ   kōnī̄
             eat.COMPL=1SG   plant   *alache*   yesterday
             'I ate *yiva tioo* yesterday.' (a type of edible plant, in Spanish, alache)

(37_22a)   ∫ì∫ī   ɾà   lō²ō   ⁿt͡sʲé²é   kōnī̄
             eat.COMPL   3SG.M   small   ground.bean.soup   yesterday
             'The boy ate ground bean soup yesterday.'

(38_22b)   ī̄ī̄   βē²ē   lō²ō
             one   house   small
             'A small house.'

(39_22c)   ī̄ī̄   ∫è²è   lō²ō
             one   trash   small
             'A little bit of trash.'

(40_22d)   ī̄ī̄   ∫è²ě   lō²ō
             one   ring   small
             'A small ring.'

(41_23a)   ∫ǐn=ì̄   ī̄ī̄   βē²ē   t͡ʃúʰtú   kōnī̄
             see.COMPL=1SG   one   house   full   yesterday
             'I saw a full house yesterday.'

(42_23b)   ī̄ī̄   lēʰsō   lō²ō
             one   rabbit   small
             'A small rabbit.'

(43_23c)   ī̄ī̄   ūʰsù   lō²ō
             one   deer   small
             'A small deer.'

(44_23d)   ī̄ī̄   jǔʰkǔ   lō²ō
             one   wilderness   small
             'A small mountain/wilderness.'

(45_24a)   lē^hsō   t͡ʃè^hē   lō^ʔō

rabbit   male   small

'A small, male rabbit.'

(46_24b)   lē^hsō   nĩ̄^hĭ́   lō^ʔō

rabbit   skinny   small

'A small, skinny rabbit.'

References

Abramson, Arthur S., & 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., & 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, & 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, & Weenink, David (2023). Praat: Doing phonetics by computer (version 6.3.10), retrieved 6 May 2023 from http://www.praat.org/.

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

Caballero, Gabriela, Juárez Chávez, Claudia, & Yuan, Michelle (2024). 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, & Steffman, Jeremy (2023). San Sebastián del Monte Mixtec. *Journal of the International Phonetic Association*, 53 (3), 1–22.

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

DiCanio, Christian, Amith, Jonathan, & Castillo García, Rey (2014). The phonetics of moraic alignment in Yoloxóchitl Mixtec. In Carlos Gussenhoven, Yiya Chen, & Dan Dediu (Eds.), Proceedings from the 4th Tonal Aspects of Languages Conference in Nijmegen, the Netherlands (pp. 203–210). https://www.mpi.nl/publications/item_2070356

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

DiCanio, Christian, Benn, Joshua, & Castillo García, Rey (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., & Castillo García, Rey (2020). Phonetic structure in Yoloxóchitl Mixtec consonants. *Journal of the International Phonetic Association*, 50 (3), 333–365.

Eberhard, David M., Simons, Gary F., & Fennig, Charles D. (Eds.) (2022). *Ethnologue: Languages of the world* (25th edn.). SIL International. Online version available at: 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, & Zamora Alarcón, Óscar (Eds.) (2012). *México: lenguas indígenas nacionales en riesgo de desaparición.* Instituto Nacional de Lenguas Indígenas.

Garellek, Marc, Chai, Yuan, Huang, Yaqian, & 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, & Baker, Kirk (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., & 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, & 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, & 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, & Salmons, Joseph C. (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, Kodukula Sri Rama, & Yegnanarayana, Bayya (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, & Shields, Jäna (1977). Silacayoapan mixtec phonology. *Studies in Otomanguean Phonology*, 54, 21–33.

Pankratz, Leo, & Pike, Eunice V. (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 Tɨ̀'un Ndá'vi, the Mixtec of Piedra Azul (San Martín Peras), Oaxaca.* Master's thesis, University of California, Santa Barbara.

Peters, Simon L., & Mendoza, Gabriel (2020). Morphophonological processes in Piedra Azul Tɨ̀'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., & Díaz-Toribio, Milton H. (2022). Edible native plants of the Gulf of Mexico province. *Biodiversity Data Journal*, 10: e80565. https://doi.org/10.3897/BDJ.10.e80565.

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

Pike, Eunice V., & Small, Priscilla (1974). Downstepping terrace tone in Coatzospan Mixtec. In Brend, R. M., (ed.), *Advances in Tagmemics*, pp. 105–134. SIL International.

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, & 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, & Mendoza Ruiz, Juana (2021). 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.