

ILLUSTRATION OF THE IPA

# San Sebastián del Monte Mixtec

**Félix Cortés**

Club Proyecto Alma Mixteca  
*ctf214@gmail.com*

**Iara Mantenuto** 

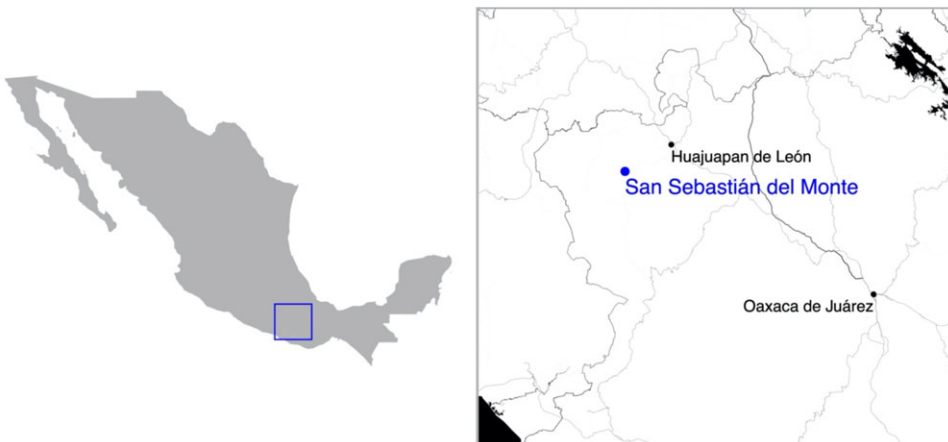
California State University Dominguez Hills  
*imantenuto@csudh.edu*

**Jeremy Steffman** 

Northwestern University  
*jeremysteffman@gmail.com*

San Sebastián del Monte Mixtec (SSM) (ISO:mks), also known as Tò'on Ndà'vi, is a language of the Mixtecan family, Otomanguan stock (Rensch 1976). The Mixtecan language family consists of Mixtec, Cuicatec and Trique, though Mixtec and Cuicatec are part of the same subgroup, also called Mixtecan (Josserand 1983: 99–101). SSM is part of the Mixteca Baja region of Oaxaca (Josserand 1983: 107). San Sebastián del Monte is a town in the Santo Domingo Tonalá municipality of Oaxaca State, Mexico, in the district of Huajuapán de León, 45 km southwest of Huajuapán de León (see Figure 1), with a population of approximately 2000 people (latitude: 17.677778, longitude: -98.021944).

Aside from very few elderly speakers who are monolingual, the majority of the people between the age of 19 years and 80 years are bilingual Mixtec and Spanish speakers. Recently, younger generations (below the age of 19 years) are mostly monolingual Spanish speakers. However, in the town, there is an ongoing effort to revitalize the language, teach



**Figure 1** (Color online) San Sebastián del Monte (right) as located within Mexico (left). Figure made with ggmaps (Kahle & Wickham 2013).

it to the younger members of the community, and to use it in activities like Taekwondo by the Ti Ndávi Ná Nisisa – Guerreros Mixteco. Many members of the community who live in California, Washington, Oregon and other US states are trilingual Mixtec/Spanish/English or bilingual Spanish/English or Spanish/Mixtec. In San Sebastián del Monte, Mixtec is the language used at home and in common interactions in the town, as well as in the city hall (*agencia*). However, the official educational language is Spanish. All the recordings come from community members living currently in Oaxaca.

The orthography has been in development in consultation with the members of the town of San Sebastián del Monte since 2019 and is still under development, mostly based on the official Mixtec orthography by the Academia de la Lengua Mixteca (Instituto Nacional de Lenguas Indígenas 2022).

There is extensive work done on the phonetics and phonology of Mixtec languages in general (Pike 1944, 1948; Bradley & Hollenbach 1988, 1990, 1991, 1992; Macaulay 1996; Paster & Beam de Azcona 2004; Daly & Hyman 2007; Hollenbach 2015; Campbell 2016; Palancar 2016; León Vázquez 2017; DiCanio, Benn & García 2018; Peters 2018; Mendoza 2020; DiCanio & Bennett 2021; Eischens 2021; Penner 2019 among others); however, each language is distinct from the others in multiple characteristics. Similarly to other Mixtec languages, SSM has VSO word order in a neutral context, a set of independent pronouns and a set of clitic pronouns, and compounds are very frequent.

We recorded our speech samples from seven speakers (four men and three women), ranging in age from 38 to 85 years old. Recordings were made using a Tascam DR-05 recorder, with recordings digitized at a sampling rate of 44.1 kHz and 32 bits.

## Consonants

The SSM consonantal inventory consists of 21 contrastive sounds. All transcriptions are broad phonemic transcriptions. Note that although we indicate prenasalization with superscript nasals, that is not official International Phonetic Alphabet (IPA) use, but we use it for simplicity following Keating, Wymark & Sharif (2019).

	Bilabial	Labiodental	Alveolar	Postalveolar	Palatal	Velar
Unaspirated plosive	p      b		t tʲ			k kʷ
Prenasalized plosive		<sup>m</sup> b	<sup>n</sup> d <sup>n</sup> dʲ			
Affricate				tʃ		
Nasal		m		n		ŋ
Trill				r		
Fricative			v      s	ʃ      ʒ		x      y
Lateral approximant				l		

CONSONANT	TRANSCRIPTION	GLOSS
/p/	pàpó	shawl
/b/	bàrrèta	barre
/t/	tata	father
/tʰ/	tʰàà	man
/k/	kanì	long
/kʷ/	kʷati	ask
/ᵐb/	ᵐbaɓí	godfather
/ᵐd/	ᵐdaʔà	hand
/ᵐdʲ/	ᵐdʲoʃi	God
/t͡ʃ/	t͡ʃáá	few
/m/	máʔá	raccoon
/n/	nàmà	soap
/ɲ/	ɲaʔá	woman
/r/	t͡ʃiriʔi	bat
/v/	vàʔa	well
/s/	satì	traditional pants
/ʃ/	ʃàʔá	lard
/ʒ/	ʒoò	pitcher
/x/	xàɲù	corncob
/ʎ/	ʎàròt͡ʃa	spear
/l/	laà	laà

The sound /p/ is used in the language infrequently, though not limited to only borrowed words as (1) shows.

- (1) pentyè /pent͡ʃè/ ‘orphan’

The phoneme /k/ at times surfaces intervocalically as [g], often limited to the morpheme *ka*. In some cases, such as in (2), it is produced as the allophone [k], but in others, as in (3), it is produced as the allophone [g].

- (2) Kásiì ínka ndìkà.<sup>1</sup>  
 /kásiì ínka <sup>n</sup>dìkà/  
 [kásiì ínka <sup>n</sup>dìkà]  
 Kási=ì ín=ka ndìkà.  
 eat.POT=1SG one=GA banana  
 ‘I am going to eat another banana.’
- (3) Nákòndisi va’akàrà jòò.  
 /nákò<sup>n</sup>disi va’a kà rà xòò/  
 [nákò<sup>n</sup>disi va’a gà rà xòò]  
 Nákòndisi va’a=kà=rà jòò.  
 cover.up.CONT well=GA=3HUM.M blanket  
 ‘The more tightly [the traveller] held his blanket.’

As Mantenuto (2020) has pointed out, the intervocalic voicing of /k/ in the morpheme *ka* is related to language change, thus the [k]–[g] alternation is a developing feature of the language, and indeed it is more common in younger speakers than in older speakers, though older speakers also present the variation, as (3) and all the occurrences in the recorded passage at the end of this paper show.

We consider /k<sup>w</sup>/, /tʃ/ and /<sup>n</sup>dʃ/, as in (4)–(6) to be contrastive sounds in the language because of the restricted occurrence of the glides only following specific sounds /k/ (for /w/), /t/ and /<sup>n</sup>d/ (for /j/).

- (4) kuati /kuati/ [k<sup>w</sup>ati] ‘ask’  
 (5) tiaà /tiaà/ [tʃaà] ‘man’  
 (6) Ndioxi /<sup>n</sup>diofi/ [<sup>n</sup>dʃofi] ‘God’

Assuming the glides to be underlyingly present in the language would result in non-valid syllable construction CCV.

The sounds [r] and [ɾ] are very infrequently used in the language, and they are not in contrast. [r] seems to be more common, as in (7)–(9), while [ɾ] seems to be less so, as it

<sup>1</sup> Sentences are reported on five lines: orthography, transcription, orthography with morphological segmentation, glosses and English translation.

Interlinear morphemic glosses adhere to the Leipzig Glossing Rules (Bickel, Comrie & Haspelmath 2015) except for the following morphemes: BASE = base morpheme, COMP = completive aspect, CONT = continuative aspect, GA = disanaphor morpheme, HUM = human, NEG = negation, POT = potential aspect, THING = classifier or pronoun referring to objects.

is restricted to only the clitic pronouns (10)–(11), even when in isolation and not occurring intervocalically.

(7)	mártèxè	/mártèʃè/	[mártèʃè]	‘Tuesday’
(8)	bàrrèta	/bàrrèta/	[bàrrèta]	‘barre’
(9)	tyirrìì	/tʃirìʔi/	[tʃirìʔi]	‘bat’
(10)	rá	/rá/	[rá]	‘clitic pronoun referring to liquids’
(11)	rà	/rà/	[rà]	‘clitic pronoun referring to men’

The sounds /b/, /<sup>m</sup>b/, /ɣ/ only marginally occur in the language, and they are often limited to borrowed words from Spanish, as in (12)–(14).

(12)	bàrrèta	/bàrrèta/	‘barre’
(13)	mpali	/ <sup>m</sup> balí/	‘godfather’
(14)	gàròcha	/ɣàròtʃa/	‘spear’

A spectrogram of the sound /<sup>m</sup>b/ is included in the section ‘Voicing in stops’.

We included both /b/ and /v/ in the language inventory because they are both present and in contrast, to the best of our knowledge, /b/ is used for Spanish borrowings, as in (15), and /v/ is used in non-Spanish borrowings, as in (16)–(17).

(15)	bàrrèta	/bàrrèta/	‘barre’
(16)	tìvulù	/tìvulù/	‘squash used to carry water’
(17)	vìkò	/vìkò/	‘cloud’

Phonetic realizations of [β] can be found as allophones of /v/ when occurring between two vowels, as the pair of words in (18) and (19) shows.

(18)	và’a	/và <sup>ʔ</sup> a/	[và <sup>ʔ</sup> a]	‘well’
(19)	ndáva	/ <sup>n</sup> dáva/	[ <sup>n</sup> dáβa]	‘fly.CONT’

/v/ is produced as [v] when at the beginning of a word, as in (18), while it surfaces as [β] when occurring intervocalically, as in (19).

/<sup>n</sup>d/ is at times produced as [<sup>n</sup>t] by some speakers; however, we assume /<sup>n</sup>d/ to be the underlying and most common form due to its frequency and the results of a more detailed phonetic analysis reported in the section ‘Voicing in stops’.

We are aware of two allophones for the sound /ʒ/, namely [ʒ] and [j]. The distribution of these allophones is beyond the purpose of this Illustration, but an example of [j] is in (20) and (21), which seems to indicate that /ʒ/ surfaces more commonly as [j] in intervocalic environments, counter to word initial /ʒ/ as in (22).

- (20) k<sup>w</sup>áyì /k<sup>w</sup>ázì/ [k<sup>w</sup>ájì] ‘horse’
- (21) niyà /nižà/ [nijà] ‘exceed.COMP’
- (22) yàà /žàà/ [žàà] ‘ash’

Two allophones are also available for /l/: at the beginning of the word, /l/ is produced as [l], as seen in (23), but intervocalically, /l/ surfaces as its fricative counterpart [ɮ], as in (24).

- (23) laà /laà/ [laà] ‘bird’
- (24) mpali /<sup>m</sup>balí/ [ᵐbaɮí] ‘godfather’

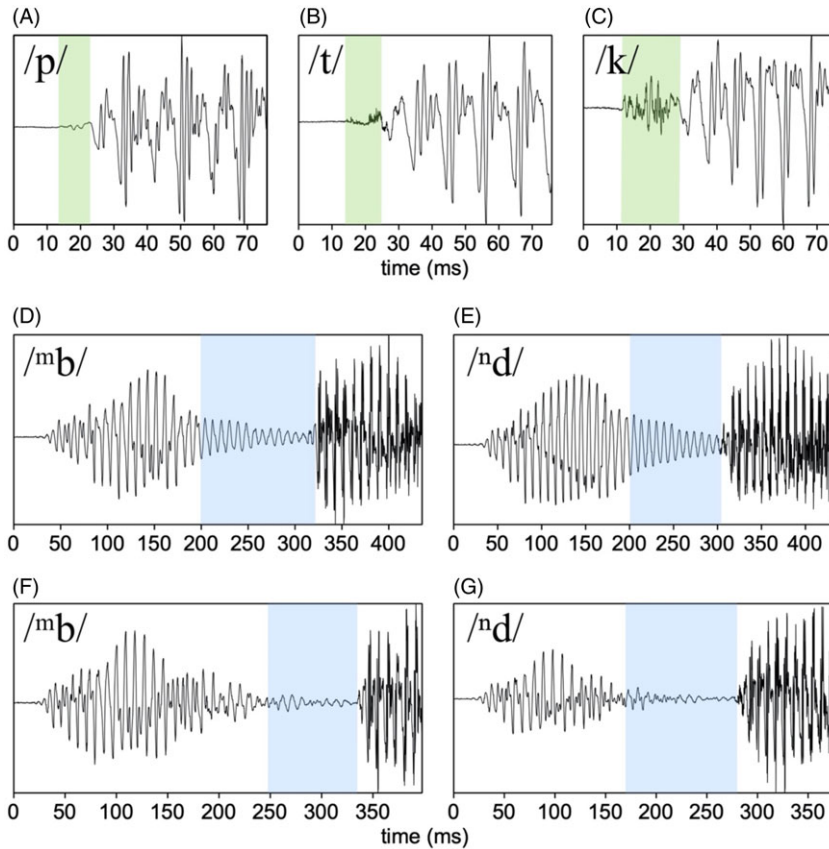
### Voicing in stops

We quantified stop voicing with two measures. For voiceless stops, we measured voice onset time (VOT). For pre-nasalized voiced stops, we measured the proportion of stop closure duration which was voiced. Voice onset time was measured from the release burst of a stop until periodic fluctuations were evident in the waveform, taking the zero crossing point of the first period of visible voicing in the waveform as the end of VOT. The proportion of voicing during stop voicing was calculated by looking at the interval which encompassed the oral stop closure and computing the proportion of the interval which was voiced. This interval was almost always voiced, though voicing could be weak and low in amplitude. In some rare cases the stop closure become voiceless shortly before the stop release. After presenting an overview of stop voicing, we present data from a larger corpus of data for three of the stops in question: /t k <sup>n</sup>d/.

Voiceless stops /p t k/ show VOT values which are typical of short-lag, or voiceless unaspirated stops. This is shown in Figures 1A, 1B and 1C. Each place of articulation exhibits a short burst interval, which also varies in burst amplitude by place. The release burst for /p/ is quieter than /t/, which is quieter than /k/ in the examples shown. Also evident in the examples, /k/ has the longest VOT, which is further confirmed in a larger corpus of data presented below. Panels D and E in Figure 2 show voiced stops with robust voicing during closure, which seems to be the canonical way in which these stops are produced. We note however, that in some productions voicing during closure was low amplitude and weaker, as exemplified in panels F and G.

We carried out a controlled elicitation to characterize consonants and vowels in SSM. This elicitation was originally designed to elicit a variety of vowels and tones, and to compare CVV and CV<sup>̃</sup>V shapes (each of which is described in detail in subsequent sections). This list also contained stops /t k <sup>n</sup>d/, allowing us to quantify voice onset time and closure voicing for these stops specifically with a larger corpus.

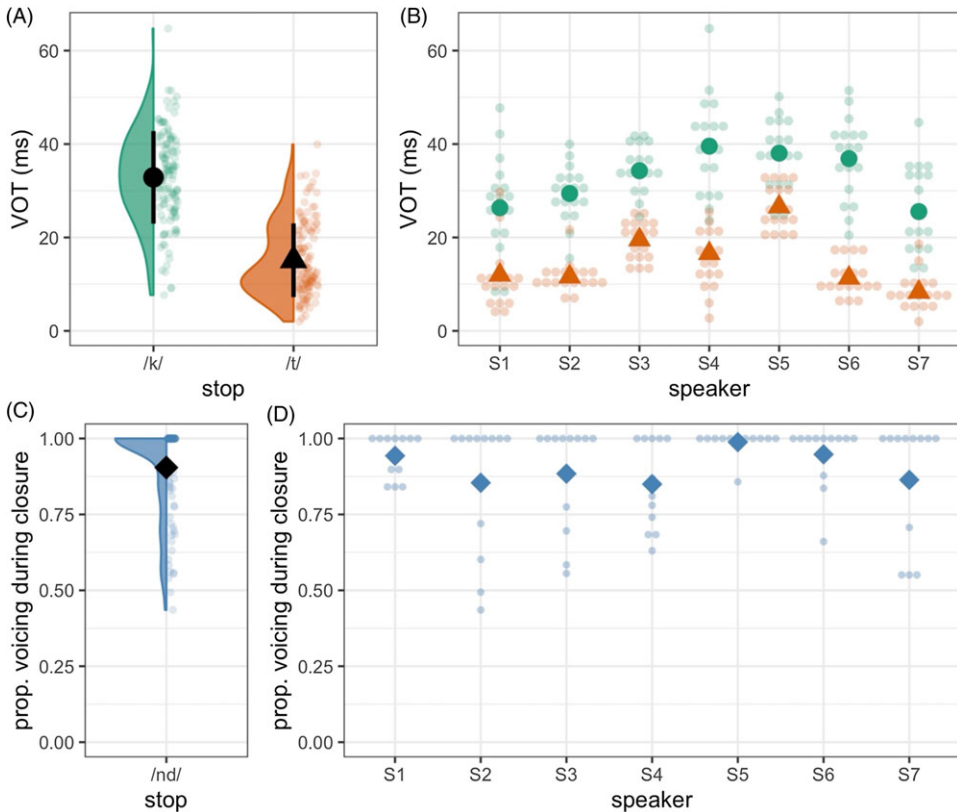
In the corpus (which we draw on throughout the paper) seven speakers of SSM (three female, four male) were recorded. Each speaker was recorded using a Tascan DR-05, in a quiet room in a house in the town of San Sebastián del Monte, Oaxaca, Mexico. The word list which was recorded consisted of words varying both in vowel and tone, some of which also contained rearticulated vowels (described below). All words were of CVV or CV<sup>̃</sup>V shape, where the latter represent rearticulation. The words were spoken in the following carrier sentence.



**Figure 2** (Colour online) Example waveforms of stops, showing VOT (Panels A–C) and closure voicing (panels D–G); with the stop in question in the top left of each panel. All tokens are produced by the same speaker (author FC). In panels A–C, the interval corresponding to measured VOT is highlighted. In Panels D–G, blue highlighting shows the interval of ORAL STOP CLOSURE, which is more strongly voiced in panel D and E, and less strongly voiced in panel F and G. All examples here are from the word list following the consonant table above, and are in the sound files included with the paper.

- (25) Tiáái \_\_\_ vîti.  
 /tʰáái \_\_\_ vîti/  
 Tiáái = ì \_\_\_ vîti.  
 write.POT = 1SG \_\_\_ now  
 ‘I am going to write \_\_\_ now.’

In total, we recorded 30 words, which were each presented six times, in a pseudo-randomized list to each speaker. Data analysis and visualization were implemented using the *ggplot2* package in R (Wickham 2016, R Core Team 2020). The words used to assess VOT and voicing in closure for /t k <sup>n</sup>d/ were selected to be matched in the following vowel across stops. We analyzed three words per voiceless stop, and two words for /<sup>n</sup>d/ (we had more /k/ initial words which we chose not to analyze here to have an equal sample of /t/ versus /k/).



**Figure 3** (Color online) VOT measures from the corpus for two stops (panel A), and the same measures split by individual speaker (panel B). Error bars in panel A represent one standard deviation. Panel C shows the proportion of voicing during closure for one stop  $^nd/$ , with data split by speaker in panel D. In all plots smaller points are individual measurements, and larger points are means.

We measured VOT for /t/ from /tíi/ ‘to be cold’, /too/ ‘frequent’ and /tùù/ ‘narrow’. We measured VOT for /k/ from /ki<sup>h</sup>i/ ‘to wear shoes’, /koo/ ‘to sit’ and /kuu/ ‘to be’. We measured the proportion of voicing in  $^nd/$  from  $^ndi/$  ‘deceased’, and  $^ndù<sup>h</sup>ù/$  ‘ours (EXCLUSIVE)’. The seven speakers each produced six repetitions of each word as described above, yielding 18 ( $6 \times 3$ ) tokens per speaker for both /t/ and /k/, and 12 per speaker for  $^nd/$ . The wordlist was originally designed for the examination of tone and vowel quality, and so lacked suitable items for /p/ and  $^mb/$ . As such, both labial stops are excluded from this analysis. The VOT measures are shown in Figure 3.

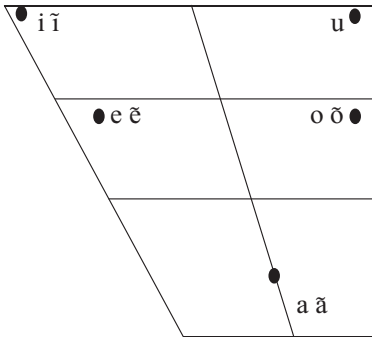
First, with respect to VOT measures for /t/ and /k/, we find confirmation of our observations based on Figure 2. Figure 3A shows each VOT measure for /t/ and /k/ as well as violin plots representing the distribution. In line with the examples in Figure 2, we see that /k/ (mean = 33 ms, SD = 10 ms) has VOT which is substantially longer than /t/ (mean = 15 ms, SD = 8 ms). Figure 3B additionally shows that, though each speaker evidences this pattern, speakers vary in terms of their distribution of VOT values for each stop, and the extent to which VOT values for /k/ and /t/ overlap.

As shown in panels C and D of Figure 3, closure voicing for  $^nd/$  is most often completely voiced. Some productions show voicing that ceases close to stop release, though in most cases unvoiced closure constitutes a small fraction of total closure duration. This pattern suggests voicing as a clear target for these stops.



## Vowels

San Sebastián del Monte Mixtec has five oral vowels and four nasal vowel phonemes, shown in the chart below. All transcriptions are broad phonemic transcriptions. It is important to note that we will use /a/ for typographical convenience to indicate a low central vowel.



SSM, like other Mixtec varieties, employs glottalization contrastively in vowels, where we use the term glottalization to refer to non-modal phonation produced with increased constriction of the vocal folds. Glottalization is evident in two types of syllable types (see Table 1). In  $\widehat{V}^?V$  and  $CV^?V$  shapes, so-called rearticulated vowels, a period of glottalization occurs between vowels which always share the same vowel quality. In  $\widehat{V}^?CV$  and  $CV^?CV$  shapes, glottalization occurs at the end of a vowel preceding a consonant. One analysis of the phonemic inventory of the language could be to posit /ʔ/ as a phoneme, though its distribution would be unusual in the sense that it would be the only allowable coda consonant in the language (in  $V^?.CV$  and  $CV^?.CV$ ), and would come with the stipulation that in  $V^?V$  and  $CV^?V$  forms the vowel quality must always be the same (see Gerfen 1999). For these reasons, we analyze glottalization as a feature of the vowel, in line with the general approach to analyzing glottalization in Mixtec at large (Macaulay & Salmons 1995, Gerfen 1999). Our transcriptions of glottalization in these vowels follow guidelines laid out by Garelle et al. (published online 19 July 2021), with the goal of capturing the fact that glottalization, is ‘a part’ of the vowel and enters into a particular phasing relationship with respect to the vowel articulation, being strongest in the middle of the vowel for  $\widehat{V}^?V$ , and at the end of the vowel for  $\widehat{V}^?$ . Each of the vowel qualities in the chart above may be produced as either  $\widehat{V}^?V$  or  $\widehat{V}^?$ , and glottalization is contrastive, as noted above (compare /ko<sup>ʔ</sup>o/ ‘to drink’ vs. /koo/ ‘to sit’). We provide a more in-depth and quantitative examination of  $\widehat{V}^?V$  rearticulated vowels

**Table 1** Syllable types. A period indicates a syllable boundary.

Syllable type	Transcription	Gloss
VV	ñ	one
$\widehat{V}^?V$	ã <sup>ʔ</sup> á	no
CVV	laà	bird
V.CV	i.tà	flower
$CV^?V$	lá <sup>ʔ</sup> á	weeping
CV.CV	ɲá.ɲá	cemetery
$CV^?.CV$	má <sup>ʔ</sup> .ɲó	middle
$\widehat{V}^?.CV$	í <sup>ʔ</sup> .ni	hot
CV.CV.CV	tʃí.ni.nò	below

below. Vowel-initial words also tend to be realized with initial glottalization on the vowel, in a narrow transcription this could be represented as [ʔV].

As in other varieties of Mixtec (Gerfen & Baker 2005), there are long and short vowels in SSM; the vowel length is predictable based on the shape of the morpheme, and is not contrastive. Morphemes that present a long vowel in the language are only CVV or VV. The vast majority of lexical items are bi- or tri-moraic, though some longer items do exist. No lexical word can be smaller than two moras (Pike 1948, Longacre 1957), but function words can be monomoraic.

VOWEL	TRANSCRIPTION	GLOSS
/a/	xaà	anger
/ã/	tãã	earthquake
/i/	isò	rabbit
/ĩ/	ĩĩ	one
/e/	tʃèe	ox
/ẽ/	sẽẽ	buy.POT
/u/	<sup>n</sup> duʔĩ	bean
/o/	xòò	bark
/õ/	ɲõ <sup>ʔ</sup>	land

All the oral vowels have a nasal counterpart, with the exception of /u/. More information on nasalization will be provided in the section ‘Nasalization’.

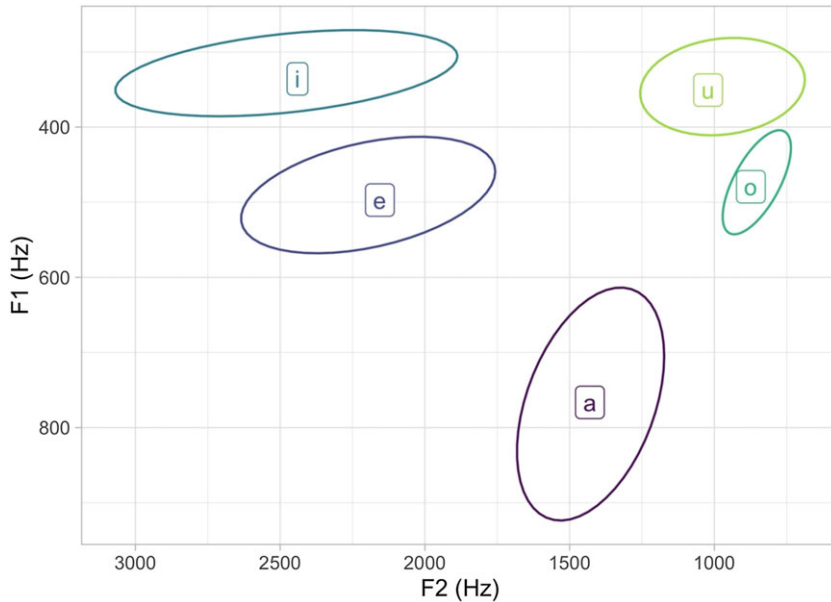
For the purpose of analyzing vowel formant structure, we compared a total of 10 words from our controlled elicitation, described above. Two words contained each of the five oral vowels in the language, as shown in Table 2.

**Table 2** Words used to calculate F1 and F2 for oral vowels in SSM.

Vowel	i	e	a	o	u
Word	tíí	kee	kaa	koo	kuu
Gloss	to be cold	will deposit	to abound	to sit	to be
	<sup>n</sup> dìi	leè	sàa	too	tùù
	deceased	baby	new	frequent	narrow

A total of 84 measurements were taken for each of the five oral vowels /i e a o u/ (2 words × 6 repetitions × 7 speakers). Vowels were segmented by hand using Praat (Boersma & Weenik 2020), and F1 and F2 measurements were extracted via a Praat script at the midpoint of each vowel. In Figure 4, we present F1 and F2 measurements in Hz.

As seen in Figure 4, /a/ shows F2 intermediate between the front and back vowels, with vowels positioned as expected in the vowel space.



**Figure 4** (Color online) F1 and F2 measurements for oral vowels in SSM. Each vowel label appears at the mean value, and ellipses show 1 standard deviation.

### Glottalization and rearticulated vowels

As described above, like other varieties of Mixtec (Macaulay & Salmons 1995, Gerfen & Baker 2005), SSM exhibits rearticulated vowels (also called ‘broken vowels’), which are characterized by a period of glottalized voice quality, or production of a glottal stop flanked by two vowels of the same quality. As described above, glottalization in Mixtec at large has generally been treated as a feature of rearticulated vowels (Macaulay & Salmons 1995). In this sense rearticulated vowels could be seen as instantiating a particular phasing relationship between a vowel articulation and glottal gesture, with a maximum of glottal constriction occurring roughly in the middle of the vowel (see Garellek et al., published online 19 July 2021). Following previous studies, we thus assume here that glottalization is a vowel feature and can be analyzed in term of dynamic changes over the whole of  $\widehat{V}^2V$ .

As described in Gerfen & Baker (2005), who report on glottalization in Coatzospan Mixtec, the realization of rearticulated vowels is highly variable. In their study, changes in pitch and spectral structure often served as cues to the contrast, also co-occurring with small changes in the amplitude envelope of the signal. The authors show in a perception experiment that these changes are sufficient to cue glottalization to speakers (see Hillenbrand & Houde 1996). More generally, the implementation of [ʔ] at large is known to be highly variable, with realizations that can range from a full and sustained stop closure, to subtler and continuous variations in voice quality (Pierrehumbert & Talkin 1992, Dilley, Shattuck-Hufnagel & Ostendorf 1996, Seid, Yegnanarayana & Rajendran 2012, Garellek 2013).

In this section, we offer a brief survey of the implementation of glottalization in SSM, focusing on rearticulated  $C\widehat{V}^2V$  forms. The words elicited in this section were elicited in the same controlled elicitation setting which was used to analyze vowel formant data. These were /s<sup>h</sup>e<sup>h</sup>e/ ‘trash’, /ki<sup>h</sup>i/ ‘to wear shoes’, /si<sup>h</sup>i/ ‘to die’, /ko<sup>h</sup>o/ ‘to drink’, /lo<sup>h</sup>o/ ‘small’, /<sup>m</sup>du<sup>h</sup>u/ ‘ours (exclusive)’, /ʒu<sup>h</sup>u/ ‘mouth’, /sa<sup>h</sup>a/ ‘foot’, and /ka<sup>h</sup>a/ ‘buttock’ (note there are two words per vowel, except for /e<sup>h</sup>e/, for which there is only one). Each of these nine words were elicited a total of six times from each of the seven speakers in the controlled elicitation (378 tokens

analyzed in total), giving us an opportunity to observe if/how implementation of glottalization may vary within and across speakers. We present two descriptions of this data. First, a qualitative coding of the frequency of production of a full and sustained stop closure. Second, a quantitative assessment of the amplitude of voicing across the rearticulated vowel, as measured by Strength of Excitation (SoE). SoE is a measure of voicing intensity, corresponding to the relative amplitude of excitation during glottal closure instants, i.e. ‘epochs’, during voicing. The measure allows for an assessment of the relative amplitude of voicing in the signal, independent of noise (Murty & Yegnanarayana 2008). Garellek et al. (published online 19 July 2021) recently used this measure to assess voicing during glottalization including in rearticulated vowels.<sup>2</sup> In assessing this quantitative measure of voicing during the rearticulated vowels we will focus descriptively on how this varies by speaker, correlates with the coding of the presence/absence of a sustained stop, and changes dynamically over the course of the  $/\widehat{V}^2V/$  interval.

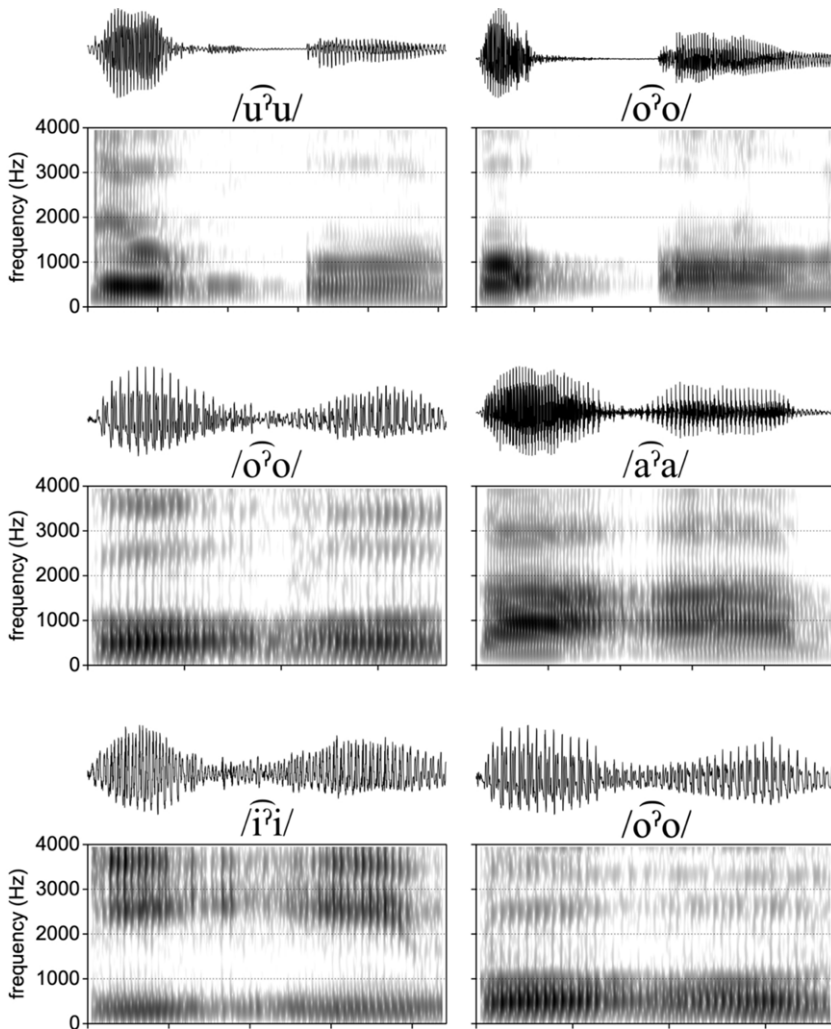
We observed that all seven speakers implemented glottalization in various ways, which formed a continuum from a full and sustained glottal stop, [ʔ], to continuous changes in pitch and voice quality. These latter cases were often accompanied by a dip in amplitude though in some cases voicing becomes highly aperiodic with low amplitude, low frequency, fluctuations in the waveform. Figure 5 shows examples of realizations exemplifying this variation, and ranging from a complete stop (top row) to a ‘lenited’ realization manifested by an amplitude dip (bottom row, in particular the rightmost example). This variation within speakers is occurring in the same elicitation setting, a controlled elicitation which might be expected to lead to more formal speech patterns. We thus conclude that, very generally speaking, speakers show substantial variability in their implementation of glottalization. To examine this variability in a more systematic fashion we coded how glottalization was realized across speakers and vowel qualities. This qualitative coding used the following criteria. A ‘full glottal stop’ was coded in the case of a sustained period of silence with unambiguous stop closure (e.g. topmost panels in Figure 5).

As shown in Figure 6, some speakers show a clear preference for a particular realization. For example, speakers 5 and 6 hardly ever produce a full glottal stop, while speakers 4 and 7 do quite frequently. Differences based on vowel quality do not show any particularly clear pattern across speakers. For example, /a/ shows the highest rate of sustained stop productions for speaker S4 and S7, while /i/ and /e/ show the highest rate for speaker S1 and S2.

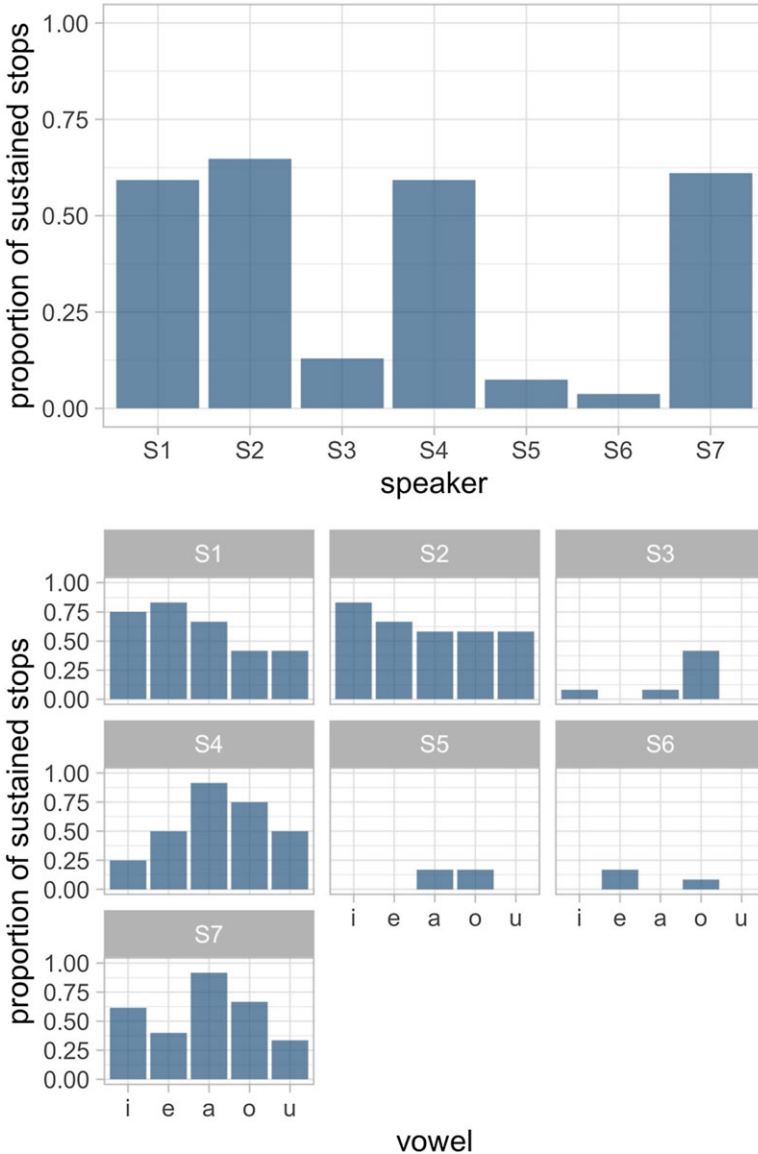
Figure 7 displays the SoE measures for each speaker over the course of a rearticulated vowel, with loess smooths fit to each individual trajectory, and for the speaker’s data as a whole. These visualizations allow us to examine the extent to which a speaker modulates the amplitude of voicing during glottalization. Speakers 5 and 6 show minimal changes in this regard, while speaker 2 shows the largest dip in amplitude. Other speakers show similar, though smaller-magnitude dips, phased roughly in the middle of the rearticulated vowel. We can also consider the consistency of the phasing and magnitude of SoE dips for each speaker. Speakers 3 and 4 seem to show the most variation in magnitude, suggesting a variability in the strength of glottalization across rearticulated vowels. The phasing of the SoE dip appears to be fairly consistent for a given speaker, though there are clearly some exceptions, for example

<sup>2</sup> We calculated SoE using VoiceSauce (Shue et al. 2011), with a temporal sampling rate of 10 ms, and then time-normalized these measures, binning them into 20 evenly-spaced samples across each  $/\widehat{V}^2V/$  interval. We subsequently took the log of the SoE at a given time, and range-normalized SoE values within a given speaker by taking the value at a given timepoint and subtracting that speaker’s minimum SoE value from it, and then dividing this by the difference between the speaker’s maximum and minimum SoE value following Garellek et al. (published online 19 July 2021). The resulting values ranged between 0 and 1, with the extremes representing the lowest amplitude of voicing (near voicelessness), and the highest amplitude of voicing for a speaker, respectively.

several productions from speakers 3 and 4 show notably later phasing of the SoE dip. The general variability in magnitude concurs with Garellek et al. (published online 19 July 2021), who suggest glottal gestures such as those in rearticulated vowels are not specified for their magnitude, in which case they vary based on other factors. In our case, the variation between speakers suggests clear speaker-specific differences in the magnitude of the glottal gestures, and additional differences in how consistent a given speaker is. It is worth reiterating here that all speakers read from the same materials in a rather formal controlled elicitation. We believe that these data show a fundamental variability in how glottalization is realized at least in rearticulated vowels. However, it appears that speakers can favor a particular implementation (as coded in our qualitative analysis), corresponding to a tendency for the magnitude of the glottal gesture (as indexed with SoE) to be realized in a particular way.



**Figure 5** Paired waveforms and spectrograms showing examples of variation in the realization of rearticulated vowels, ranging from a full and sustained glottal stop [ʔ], to continuous changes in amplitude and voice quality. Note the frequency range in the spectrograms is 0–4000 Hz, and each tick on the x axis below each spectrogram represents 100 ms, showing the approximate duration of each example. Examples are from two different speakers.

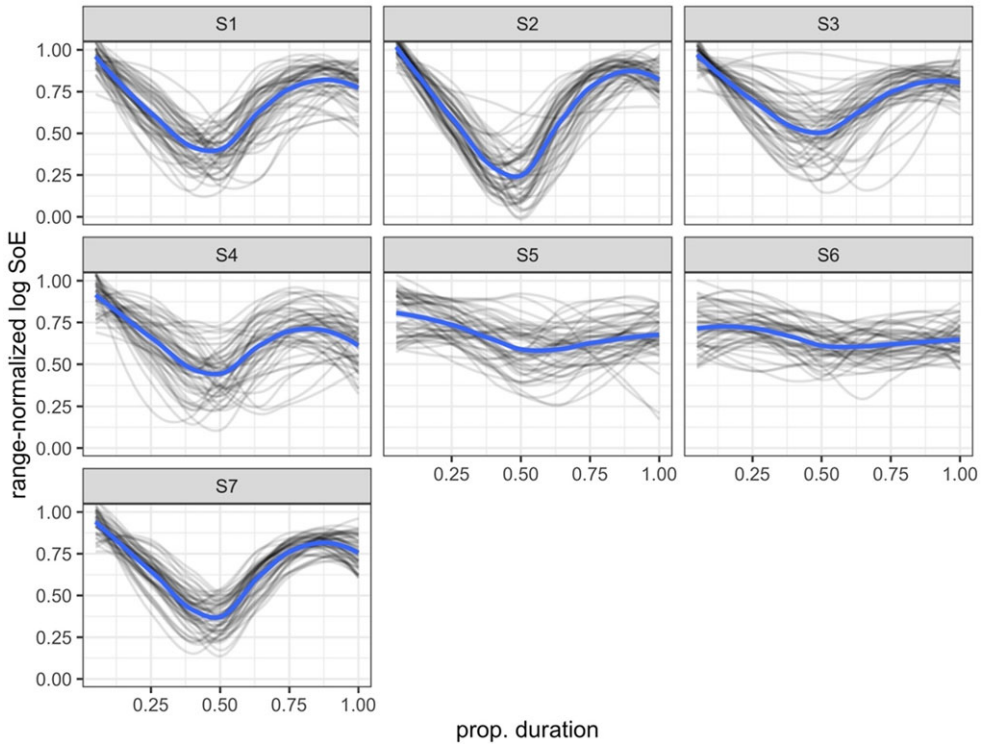


**Figure 6** (Color online) Distribution of the realization of a sustained glottal stop across speakers where S1-S7 refers to speakers 1 through 7 (top) and split by speaker and vowel quality (bottom).

### Nasalization

Vowel nasalization in Mixtec languages is contrastive in the language as evident in the minimal pairs shown above, and it is also the result of the process of nasal spreading, which we leave aside for future work on contextual nasalization. In this paper we focus on describing contrastive nasalization, but we refer the reader to work by Marlett (1992) and DiCano et al. (2020) for more information on nasalization in other varieties of Mixtec.

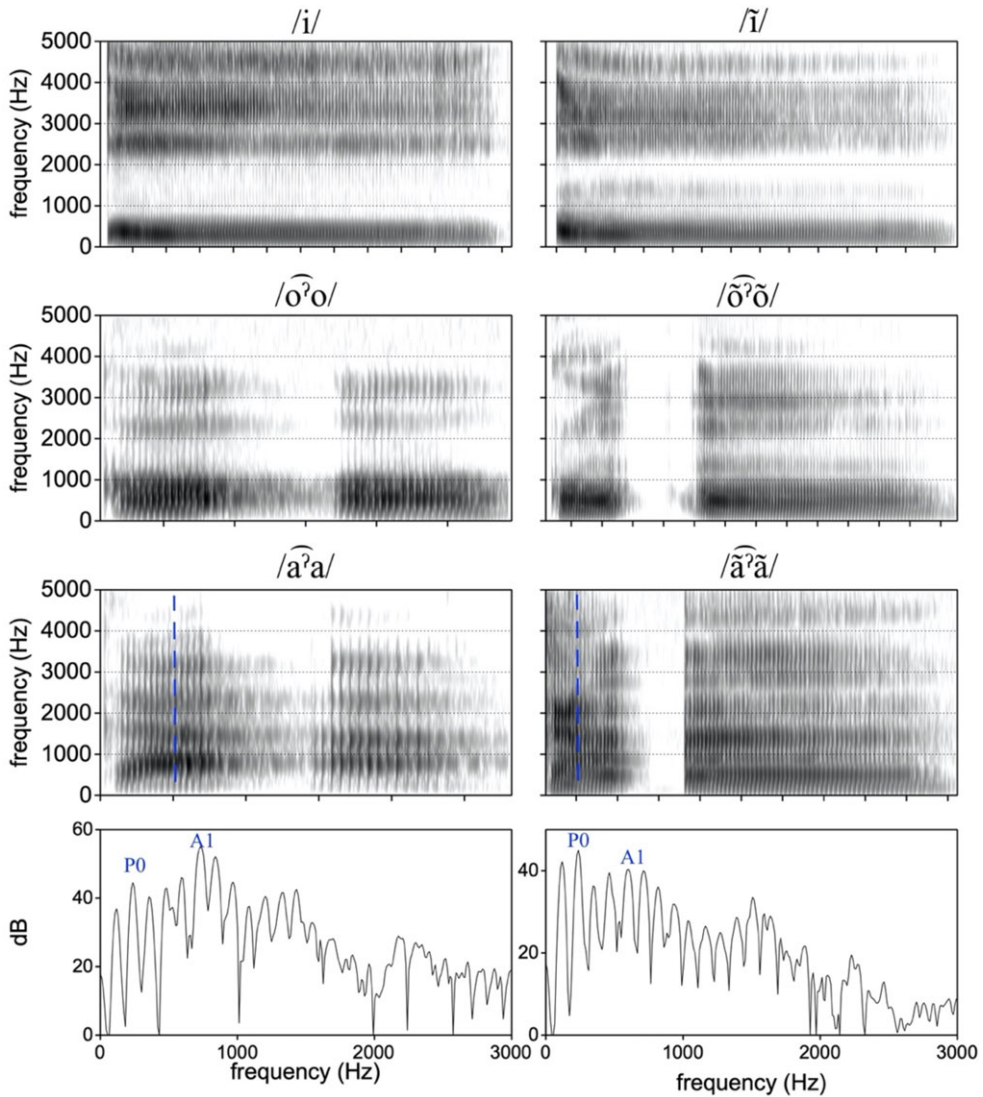
Similarly to other varieties of Mixtec, not all oral vowels have a phonemic nasal pair in the language. We based our conclusion on the minimal and near minimal pairs available, as reported in (26)–(30).



**Figure 7** (Color online) SoE measured over the course of normalized time for each of the seven speakers. Transparent black lines are loess smooths fit to each individual production, thicker blue lines are smooths fit to aggregate speaker data.

- (26) a. ka'à /ka'ǎ/ 'buttock'
- b. ka'an /kǎ'ǎ/ 'speak.POT'
- (27) a. ko'o /ko'o/ 'drink.POT'
- b. ko'on /kõ'õ/ 'go.POT'
- (28) a. ìì /ìì/ 'delicate'
- b. ììn /ììn/ 'nine'
- (29) a. sé'è /sè'è/ 'trash'
- b. seen /sè'è/ 'buy.POT'
- (30) kúú /kúú/ 'be.POT'

As a result, we can conclude that while /a/, /o/, /i/ and /e/ present a nasal counterpart, namely /ǎ/, /õ/, /ĩ/ and /ẽ/, the oral vowel /u/ does not have a nasal counterpart in SSM.



**Figure 8** Spectrograms of oral (left column) and nasal (right column) vowels for three vowel qualities (labeled above each panel). Ticks on the x axis of the spectrogram mark 100 ms intervals. The spectra in the bottom row are from  $\widehat{a}^{\text{?}}\widehat{a}/$  and  $\widehat{\tilde{a}}^{\text{?}}\widehat{\tilde{a}}/$ , at the point of the vowel indexed with the dashed line. The examples are from three minimal pairs:  $/\widehat{i}/$  'delicate',  $/\widehat{i}\tilde{i}/$  'nine',  $/\widehat{o}^{\text{?}}\widehat{o}/$  'to drink',  $/\widehat{o}\tilde{o}^{\text{?}}\widehat{o}/$  'to go',  $/\widehat{ka}^{\text{?}}\widehat{a}/$  'buttock', and  $/\widehat{ka}\tilde{a}^{\text{?}}\widehat{a}/$  'to speak'.

To exemplify how contrastive nasalization in vowels is realized, we present spectrograms and spectra from several pairs of nasal and oral vowels in Figure 8. As shown the figure, nasal vowels  $\widehat{i}\tilde{i}/$  and  $\widehat{o}\tilde{o}^{\text{?}}\widehat{o}/$  show the presence of a resonance in the frequency range between 1000 Hz and 2000 Hz (lacking in their oral counterparts). The formant values of F1 and F2 in these two nasal vowels are otherwise fairly similar to their oral counterparts. The spectrograms and spectra for  $\widehat{a}^{\text{?}}\widehat{a}/$  and  $\widehat{\tilde{a}}^{\text{?}}\widehat{\tilde{a}}/$  in the bottommost two rows of Figure 8 also show clear evidence for an effect of nasalization. First, we see that the second harmonic is boosted in amplitude in  $\widehat{\tilde{a}}^{\text{?}}\widehat{\tilde{a}}/$ , as is typical for nasal vowels. This may be attributed to nasality as an increase in amplitude of the first nasal pole, P0 (Chen 1997). This can be considered in comparison to A1 (the amplitude of the harmonic closest to F1), which is relatively reduced in intensity as



compared to P0. In the oral counterpart /ã<sup>h</sup>a/, A1 is clearly higher in intensity, and P0 is not boosted. We can also note a change in F1 and F2 in nasalized /ã<sup>h</sup>ã/, with F1 lowered and F2 raised slightly (centralizing the vowel), as is visually clear in both the spectrograms and spectra.

## Tones

There are three tones in SSM, a high tone <ã>, a mid tone <a>, and a low tone <ã>. As in other Mixtec varieties, each mora carries a separate tone. Table 3 shows how monomoraic short vowels contrast in tone.

Tonal contours, which can occur in bimoraic CVV syllables, are analyzed as sequences of level tones. In the literature (León Vázquez 2017), seven possible contour tones have been described. Examples of words showing these seven contours are given in Table 4.

In this section we present an acoustic analysis of tonal contrasts in SSM, where we examine how f0 varies as a function of tonal category. For this analysis we used different tokens from the same corpus of speech, recorded from seven speakers described in the ‘Voicing in stops’ section above. In addition to the tone-matched items which we elicited for the purpose of measuring F1 and F2, we elicited vowel-matched words contrasting in tone. These items are shown in Table 4. The words had either /a/ or /o/ as vowels, though note that only /a/ contained all possible tones. A total of six productions of each tone-contrasting word was elicited from each speaker during the randomized elicitations, described above.

The vowel in these words was segmented as described in the ‘Vowels’ section. We subsequently extracted time-normalized f0 information using ProsodyPro (Xu 2013), at 30 time points across each vowel. After inspection of each speakers’ productions, we noted that two speakers produced reliably creaky phonation throughout the elicitation, which was particularly marked on words with low tones. Accurate f0 measurements were difficult to obtain from these speakers for almost all productions, including non-low tones. As such, we report on just five speakers for the tonal data, for whom f0 was reliably tracked. The reported measurements are thus based on a total of 360 tokens (5 speakers × 12 words × 6 repetitions). Figure 9 below plots time-normalized f0 trajectories for each tonal category.

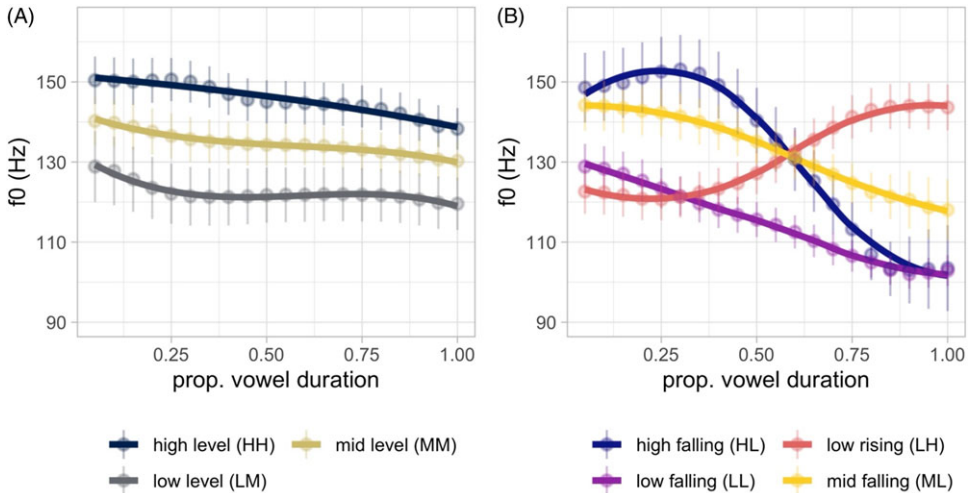
As shown in Figure 9 panel A, three tones are primarily distinguished by the overall height of f0 across the vowel. These are the high level (HH), mid-level (MM), and low level (LM) categories. It is notable that the so-called LM tone is relatively non-dynamic, contrary to what might be expected. This can be compared to the low falling (LL) tonal category shown in panel B. The LL tone is best characterized phonetically as a low falling tone, which is distinguished from the LM category in the latter portion of the vowel (León Vázquez 2017). Turning to the more dynamic tonal categories in panel B, we can see that low rising (LH) and high falling (HL) contours generally behave in the expected way: low rising tones start relatively low in a speaker’s pitch range and rise towards the top. The opposite is true for high

**Table 3** Level tone examples in monomoraic syllables.

High	Mid	Low
<sup>n</sup> dáva ‘fly.CONT’	<sup>n</sup> dava ‘fly.POT’	<sup>n</sup> dàva ‘fly.COMP’
xóko ‘placenta’	xòko ‘hunger’	xòkò ‘shoulder’

**Table 4** Words contrasting in tone, used in the f0 analysis.

Tone	HH	MM	LL	LH	HL	ML	LM
Word	tʃáá	kaa	ʒàà	xàá	xáà	kaà	kàa
Gloss	few	to be abundant	ash	then	hot	there	metal
	ʒóó	koo	kòò	kòó		ʒòò	
	we	to sit	viper	no		pitcher	



**Figure 9** (Color online) Mean f0 values for tones, with panel A showing relatively level tones, and panel B showing dynamic tones. The y axis plots f0, averaged across speakers (four male, one female). Tones are labeled below each plot. Points indicate means for each tone, with errorbars around each mean representing 1 standard error. Lines show a smoothed trajectory with a loess smooth fit to each tone.

falling tones. Mid falling tones start slightly lower than high falling tones, though notably seem to fall to a slightly higher target than high falling and low falling categories – thus it seems that mid falling tones operate in a slightly more compressed pitch range. In sum, we can see that each of the seven tonal categories is well distinguished by f0, including the distinction between the level (LM) and falling (LL) categories.

**Transcription of a recorded passage**

Note that this story is not a traditional Mixtec story and as such one of the authors had to first read it in Spanish, and then translate into San Sebastián del Monte Mixtec; however, the transcription was done after the story was recorded. We included the tones in the phonemic transcription without accounting for tone sandhi or lack of it, as more research is necessary to better understand the tonal system.

**Broad phonemic transcription**

nató<sup>ʔ</sup> tātŋí sǐ<sup>ʔ</sup> n<sup>h</sup>dika<sup>n</sup>diì || nò<sup>h</sup> n<sup>h</sup>di<sup>n</sup>divè<sup>h</sup> n<sup>h</sup>djà<sup>h</sup> ja kúú ja n<sup>h</sup>dako kà ||  
 tjà<sup>n</sup>ò zika nižà<sup>ʔ</sup> ñ tjà sika itŋí || kuà<sup>ʔ</sup> n<sup>h</sup>dísi xò<sup>h</sup> n<sup>h</sup>deé ||  
 xà<sup>n</sup> kè<sup>n</sup>dótò<sup>ʔ</sup> || n<sup>h</sup>djà<sup>h</sup> ja jinapó<sup>ʔ</sup> kaxa ja na tava mee tjà<sup>h</sup> xò<sup>h</sup> n<sup>h</sup>dísi rà || kúú ja n<sup>h</sup>dako kà ||  
 xaá kàsà<sup>ʔ</sup> kánà n<sup>h</sup>deé kini tātŋí || xò<sup>h</sup> tŋí nižà<sup>ʔ</sup> n<sup>h</sup>deé || n<sup>h</sup>díja kí<sup>ʔ</sup>va n<sup>h</sup>deé kà kánà ||  
 ví<sup>ʔ</sup> kà ná<sup>n</sup>kò<sup>n</sup>disi va<sup>ʔ</sup> kà rà xò<sup>h</sup> || nò<sup>h</sup> n<sup>h</sup>dí<sup>ʔ</sup> zó xàá || sa<sup>n</sup>d<sup>h</sup>ava<sup>ʔ</sup> inè ja kò n<sup>h</sup>ikevè<sup>ʔ</sup> ||  
 n<sup>h</sup>dìkò kàsà<sup>ʔ</sup> nándiì n<sup>h</sup>deé va<sup>ʔ</sup> n<sup>h</sup>dika<sup>n</sup>diì || tivit<sup>ʔ</sup>o nitava rà xò<sup>h</sup> n<sup>h</sup>dísi rà ||  
 xìká kùu xàá || n<sup>h</sup>dìsà nakóni tātŋí ja || n<sup>h</sup>dika<sup>n</sup>diì kúú ja n<sup>h</sup>dakò kà nò<sup>h</sup> n<sup>h</sup>di<sup>n</sup>divè

**Orthographic representation**

Nató'ón tàtyì sí'ín ndikandiì nòò ndindivi èn ndiaà ña kúú ña ndako kà, tiàñò yika niyà'a ìin tiàa síka ityì kuà`àn ndísi jòò ndeé. Jàa kèndótò'on òn, ndiaà ña xinañó'ón kaja ña na tava mee tiàa jòò ndísi rà, kúú ña ndako kà. Jaá kàsà'á kánà ndeé kini tàtyì, jòòn tyí niyà'o ndeé, ndixa kì'va ndeé kà kána àn, ví'í kà nákòndisi va'a kà rà jòò; nòò ndi'i yó'o jàa sa'ntiava'a ini èn ña kò nikevà'a àn. Ndikò kàsà'á nándiì ndeé va'a ndikandiì, Tivitio nitava rà jòò ndísi rà; jìkà kùu jàà ndisà nakóni tàtyì ña ndikandiì kúú ña ndakò kà nòò ndindivi èn.

**Glossed orthographic representation**

Nató'ón	tàtyì	sí'ín	ndikandiì	nòò	ndindivi=èn	ndiaà=ña
argue.COMP	wind	and	sun	on	both=3THING.OBJ	some=3THING
kúú=ña	ndako=kà,	tiàñò	yika	ni-yà'a	ìin	[tiàa síka
be.COMP=3THING	strong=GA	while	there	COMP-go	one	man walk.CONT
ityì]	kuà`àn	ndísi	jòò	ndeé.	Jàa	kèndótò'on=òn,
path	go.CONT	cover.CONT	blanket	thick	then	agree.COMP=3THING.OBJ
ndiaà=ña	xinañó'ón	kaja=ña	natava	mee	tiàa	jòò
who=3THING	first	have.POT=3THING	take.off.POT	BASE	man	blanket
ndísi=rà,	kúú	ña-ndako=kà.	Jaá	kàsà'á		
cover.CONT=3HUM.M	be.CONT	CLF:3THING=strong=GA	then	begin.COMP		
kánà	ndeé	kini	tàtyì,	jòòn	tyí	niyà'o
blow.CONT	strong	much	wind	but	because	exceed.COMP
kì'va	ndeé=kà	kána = àn,	ví'í=kà	nákòndisi		
measure	strong=GA	blow.CONT=3THING.OBJ	better=GA	cover.up.CONT		

va'a = kà = rà	jòd;	nòó	ndì'i	yó'o	jàá	[sa'ntia-va'a
well=GA=3HUM.M	blanket	where	finish.COMP	here	then	[cut.COMP-well
ini] = èn	ña-kò-nikev'à = àn.					Ndikò kàsà'á
mind]=3THING.OBJ	CLF:3THING-NEG-win.COMP=3THING.OBJ	soon	begin.COMP			
nándiì	ndeé	va'a	ndikandiì,	tivitio	nitava=rà	jòd
shine.COMP	strong	well	sun	quickly	get.COMP=3HUM.M	blanket
ndísi=rà;	jìká	kùu	jàá	ndisà	nakóni	tàtyì
put.COMP=3HUM.M	so	be.COMP	then	strongly	recognize.COMP	wind
ña-ndikandiì	kúú	ña-ndakò=kà			nòd	ndindivi = èn.
CLF:3THING-sun	be.CONT	CLF:3THING-strong=GA	on	both=3THING.OBJ		

### Spanish translation

El viento y el sol estaban discutiendo sobre cuál de ellos era el más fuerte, cuando pasó un viajero envuelto en una gruesa capa. Quedaron de acuerdo en que quien primero logrará que aquel viajero se quitará la capa sería considerado el más poderoso. Entonces el viento sopló tan fuerte como pudo, pero entre más soplaba, más se cobijaba en su capa el viajero; al fin el viento del norte se rindió. Entonces el sol brilló intensamente, e inmediatamente el viajero se quitó la capa; así el viento se vio obligado a reconocer que el sol era el más poderoso de los dos.

### English translation (translated back from Mixtec and Spanish)

The wind and the sun were arguing about which of the two was stronger. Then there came a traveller wearing a thick blanket. They agreed that whoever of the two would be able to make the man take off his blanket would be considered the stronger of the two. So, then the wind blew with all its might. The stronger the wind blew, the more tightly the traveller held his blanket, until the wind grew tired and stopped blowing. Then the sun shone strongly, making it become very hot. Suddenly the traveller took off his blanket from the heat. And so, the wind agreed that the sun was truly the stronger of the two.

### Acknowledgments

We are grateful to Mr. Miguel Cortés, Ms. María del Carmen Ortiz, Mr. Ignacio Pimentel, Mr. Donaciano Cortés, Ms. María Rojas and Ms. Reyna Torres for sharing their language with us. We thank the Club Proyecto Alma Mixteca, the Agencia Municipal of the town of San Sebastián del Monte and the escuela telesecundaria of San Sebastián del Monte for their support. We are very thankful for the helpful feedback and comments from the two anonymous reviewers and from the editors: Dr. Marc Garellek and Dr. Marija Tabain. Many thanks are due to Jae Weller, Qingxia Guo and Bryan Gonzalez for their help with data annotation and to Octavio León Vázquez for comments and suggestions. This research was supported by the Ladefoged Scholarship Award granted by the UCLA Department of Linguistics and by California State University Dominguez Hills Professional Development Fund. Any mistakes are our own.

## Supplementary material

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0025100322000226>

## References

- Bickel, Balthasar, Bernard Comrie & Martin Haspelmath. 2015. The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. Retrieved from <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>, 1 November 2020.
- Boersma, Paul & David Weenink. 2020. Praat: Doing phonetics by computer (version 6.1.09). Retrieved from <http://www.praat.org>, 1 November 2020.
- Bradley, C. Henry & Barbara E. Hollenbach (eds.). 1988. *Studies in the syntax of Mixtecan languages*, vol. 1 (Summer Institute of Linguistics Publication No. 83). Dallas, TX: SIL & Arlington, TX: The University of Texas at Arlington.
- Bradley, C. Henry & Barbara E. Hollenbach (eds.). 1990. *Studies in the syntax of Mixtecan languages*, vol. 2 (Summer Institute of Linguistics Publication No. 90). Dallas, TX: SIL & Arlington, TX: The University of Texas at Arlington.
- Bradley, C. Henry & Barbara E. Hollenbach (eds.). 1991. *Studies in the syntax of Mixtecan languages*, vol. 3 (Summer Institute of Linguistics Publication No. 105). Dallas, TX: SIL & Arlington, TX: The University of Texas at Arlington.
- Bradley, C. Henry & Barbara E. Hollenbach (eds.). 1992. *Studies in the syntax of Mixtecan languages*, vol. 4 (Summer Institute of Linguistics Publication No. 111). Dallas, TX: SIL & Arlington, TX: The University of Texas at Arlington.
- Campbell, Eric W. 2016. Tone and inflection in Zenzontepec Chatino. In Palancar & Léonard (eds.), 141–162.
- Chen, Marilyn Y. 1997. Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America* 102(4), 2360–2370.
- Daly, John P. & Larry M. Hyman. 2007. On the representation of tone in Peñoles Mixtec. *International Journal of American Linguistics* 73(2), 165–207.
- DiCano, Christian T., Joshua Benn & Rey Castillo García. 2018. The phonetics of information structure in Yoloxóchitl Mixtec. *Journal of Phonetics* 68, 50–68.
- DiCano, Christian T. & Ryan Bennett. 2021. Mesoamerica. In Carlos Gussenhoven & Aojun Chen (eds.), *The Oxford handbook of language prosody*, 408–427. Oxford: Oxford University Press.
- DiCano, Christian T., Caicai Zhang, Douglas H. Whalen & Rey Castillo García. 2020. Phonetic structure in Yoloxóchitl Mixtec consonants. *Journal of the International Phonetic Association* 50(3), 333–365.
- Dilley, Laura, Stefanie Shattuck-Hufnagel & Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4), 423–444.
- Eischens, Ben. 2021. Laryngealization and word reduction in Mixtec: Phonetics in phonology. Qualifying paper, University of California Santa Cruz.
- Garellek, Marc. 2013. *Production and perception of glottal stops*. Ph.D. dissertation, UCLA.
- Garellek, Marc, Yuan Chai, Yaqian Huang & Maxine Van Doren. Voicing of glottal consonants and non-modal vowels. *Journal of the International Phonetic Association*, <https://doi.org/10.1017/S0025100321000116>. Published online by Cambridge University Press, 19 July 2021.
- Gerfen, Chip. 1999. *Phonology and phonetics in Coatzacoapan Mixtec* (Studies in Natural Language and Linguistic Theory 48). Dordrecht: Springer. <https://doi.org/10.1007/978-94-017-2620-7>.
- Gerfen, Chip & Kirk Baker. 2005. The production and perception of laryngealized vowels in Coatzacoapan Mixtec. *Journal of Phonetics* 33(3), 311–334.
- Hillenbrand, James M. & Robert A. Houde. 1996. Role of f0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech, Language, and Hearing Research* 39(6), 1182–1190.
- Hollenbach, Barbara E. 2015. Notes on tense, mood, and negation in Mixtec languages. Ms., SIL International.

- Instituto Nacional de Lenguas Indígenas. 2022. *Norma de escritura del Tu'un Savi (idioma mixteco)*. Mexico City: Secretaría de Cultura Instituto Nacional de Lenguas Indígenas.
- Josserand, Judy K. 1983. *Mixtec dialect history*. Ph.D. dissertation, Tulane University.
- Kahle, David & Hadley Wickham. 2013. ggmap: Spatial Visualization with ggplot2. *The R Journal* 5(1), 144–161.
- Keating, Patricia A., Daniel Wymark & Ryan Sharif. 2019. Proposal for superscript diacritics for prenasalization, preglottalization and preaspiration. *Journal of the International Phonetic Association* 51(1), 75–90.
- León Vázquez, Octavio. 2017. Sandhi tonal en el mixteco de Yucuquimi de Ocampo, Oaxaca. Masters thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Longacre, Robert E. 1957. *Proto-Mixtecan* (Indiana University Research Center in Anthropology, Folklore, and Linguistics 5). Bloomington, IN: Indiana University.
- Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec*. Los Angeles, CA: University of California Press.
- Macaulay, Monica & Joseph Salmons. 1995. The phonology of glottalization in Mixtec. *International Journal of American Linguistics* 61, 38–61.
- Mantenuto, Iara. 2020. 'Other' and 'more' in San Sebastián del Monte Mixtec and beyond. Ph.D. dissertation, UCLA.
- Marlett, Stephen A. 1992. Nasalization in Mixtec languages. *International Journal of American Linguistics* 58(4), 425–435.
- Mendoza, Iní G. 2020. Syntactic sketch of San Martín Peras Tu'un Savi. Bachelor of Arts Honors Thesis, University of California, Santa Barbara.
- Murty, Sri Rama & Bayya Yegnanarayana. 2008. Epoch extraction from speech signals. In *IEEE Transactions on Audio, Speech, and Language Processing* 16, 1602–1613.
- Palancar, Enrique L. 2016. A typology of tone and inflection: A view from the Oto-Manguean languages of Mexico. In Palancar & Léonard (eds.), 109–140.
- Palancar, Enrique L. & Jean Léo Léonard (eds.). 2016. *Tone and inflection: New facts and new perspectives*. Berlin: De Gruyter Mouton.
- Paster, Mary & Rosemary Beam de Azcona. 2004. A phonological sketch of the Yucunany dialect of Mixtepec Mixtec. In Carmen Jany (ed.), *Proceedings of the 7th Annual Workshop on American Indian Languages*, 61–76.
- Penner, Kevin. 2019. *Prosodic structure in Ixtayutla Mixtec: Evidence for the foot*. Ph.D. dissertation, 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. Masters thesis, University of California, Santa Barbara.
- Pierrehumbert, Janet & David Talkin. 1992. Lenition of /h/ and glottal stop. In Gerard Docherty & D. Robert Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody*, 90–117. Cambridge: Cambridge University Press.
- Pike, Kenneth L. 1944. Analysis of a Mixteco text. *International Journal of American Linguistics* 10(4), 113–138.
- Pike, Kenneth L. 1948. *Tone languages: A technique for determining the number and type of Pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor, MI: University of Michigan Press.
- R Core Team 2020. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>, 1 November 2020.
- Rensch, Calvin R. 1976. *Comparative Otomanguean phonology*. Bloomington, IN: Indiana University.
- Seid, Huessien, Bayya Yegnanarayana & S. Rajendran. 2012. Spotting glottal stop in Amharic in continuous speech. *Computer Speech & Language* 26(4), 293–305.
- Shue, Yen-Liang, Patricia A. Keating, Chad Vicenik, & Kristine Yu. 2011. VoiceSauce: A program for voice analysis. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, 1846–1849.
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer.
- Xu, Yi. 2013. ProsodyPro: A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013).