CrossMark

Phonetic structure in Yoloxóchitl Mixtec consonants

Christian T. DiCanio
University at Buffalo & Haskins Laboratories
cdicanio@buffalo.edu

Caicai Zhang
Hong Kong Polytechnic University
caicai.zhang@polyu.edu.hk

Douglas H. Whalen
Haskins Laboratories & CUNY Graduate Center
whalen@haskins.yale.edu

Rey Castillo García
Secretaria de educación pública, Guerrero (Mexico)
reyyoloxochitl@gmail.com

While Mixtec languages are well-known for their tonal systems, there remains relatively little work focusing on their consonant inventories. This paper provides an in-depth phonetic description of the consonant system of the Yoloxóchitl Mixtec language (Oto-Manguean, ISO 639-3 code *xty*), a Guerrero Mixtec variety. The language possesses a number of contrasts common among Mixtec languages, such as voiceless unaspirated stops, prenasalized stops, and a strong tendency for words to conform to a minimally bimoraic structure. Using a controlled set of data, we focus on how WORD SIZE and WORD POSITION influence the acoustic properties of different consonant types. We examine closure duration, VOT, and formant transitions with the stop series, spectral moments with the fricative series, the timing between oral and nasal closure with the prenasalized stop series, and both formant transitions and qualitative variability with the glide series. The general effect of WORD SIZE is discussed in relation to work on POLYSYLLABIC SHORTENING (Turk & Shattuck-Hufnagel 2000) and demonstrates the importance of prosodic templates in Mixtec languages (Macken & Salmons 1997). The prenasalized stop data provides evidence that such stops are best analyzed as allophones of nasal consonants preceding oral vowels (as per Marlett 1992) and not as hypervoiced variants of voiced stops (as per Iverson & Salmons 1996).

1 Introduction

The process of language description frequently focuses on aspects of a language that are notable from an areal, typological, or theoretical perspective. Although Mixtec languages have relatively less complex consonantal inventories than many languages from other Oto-Manguean families, e.g. Zapotecan, Popolocan, Otomian, and Tlapanecan, several aspects of Mixtec consonantal phonology have been the focus of debate in the literature. For instance, the phonological status of prenasalized stops throughout different Mixtec languages is unresolved. Marlett (1992) argues that they are best analyzed as allophones of

Journal of the International Phonetic Association (2020) 50/3
doi:10.1017/S0025100318000294
© International Phonetic Association
First published online 12 February 2019
https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

nasal consonants, while Iverson & Salmons (1996) argue that they are best analyzed as voiced stops with hypervoicing (see Section 5.1.1 below). Moreover, in Coatzospan Mixtec, there is evidence for nasalized fricatives (C. Gerfen 2001). This stands counter to a claim in the phonetics literature arguing that it is aerodynamically impossible to maintain nasal airflow while simultaneously producing sufficiently high intra-oral air pressure to maintain turbulence in the oral cavity for fricative production (Ohala & Ohala 1993, Huffman & Krakow 1993).

Debates like these can certainly be informed with instrumental phonetic data from Mixtec languages, but there are also several independent motivations for pursuing descriptive phonetic work on undescribed or under-described languages. First, it is by no means certain that contrasts which have been labeled identically across languages, even at the phonetic level, e.g. 'velar ejective,' are produced similarly across languages. While some of these differences may arise from errors in transcription, others arise from the coarseness inherent to phonetic transcription itself. For instance, the fundamental frequency (f0) on the vowel following the stop release is dependent on the phonological voicing of the stop (Keating 1984, Kingston & Diehl 1994). Thus, in a language like Hindi, where a phonetically voiceless unaspirated stop is an allophone of a voiceless stop series, e.g. [t] is an allophone of /t/, f0 raises upon release. In a language like English, where a phonetically voiceless unaspirated stop is an allophone of a VOICED stop series in utterance-initial position (Davidson 2016), e.g. [t] is an allophone of /d/, f0 LOWERS upon release. These stops are identically, narrowly transcribed as [t] in the IPA, but they differ in phonetic detail. Second, the establishment of these phonetic details in an individual language is the baseline from which one may test more targeted theories of speech production. Research on well-studied languages often relies on fundamental work in phonetic science which assumes existing knowledge of the descriptive phonetic properties of the language, i.e. the extensive research on voice onset time (VOT; Lisker & Abramson 1964, Abramson & Whalen 2017) was made possible because early work investigated the descriptive properties of languages like English, French, and Thai. Third, careful phonetic detail is useful for the purposes of historical reconstruction (Ohala 2001, Blevins 2004) and for examining how a language's phonological contrasts compare with those from well-studied languages. Finally, a majority of the published phonetic literature on endangered languages has not been based on large samples, partly due to limited use of automatic methods for acoustic analysis. The current paper relies heavily on automated methods for the analysis of different acoustic properties, and we will discuss the limitations as well as the advances inherent in the methods. The scripts used in this study are listed in the Appendix A. We hope that this study can serve as an accessible model for modern descriptive phonetic research.

The focus of the current paper is an acoustic examination of the different consonant types in Yoloxóchitl Mixtec, a Mixtec language in the Guerrero Mixtec branch (Josserand 1983) and the focus of major documentation work (Amith & Castillo García, n.d.). In the remainder of this section, we provide a background discussion of the language and its phonological inventory. In each of the following sections, the acoustics of a particular consonant type are investigated. The phonetic data consists of field recordings of different consonant types produced in carrier phrases by eight native speakers. Like similar studies on the phonetic structure of a language (Silverman et al. 1995, Maddieson, Avelino & O'Connor 2009), the current study is exploratory and not driven by an over-arching set of theoretical hypotheses. However, individual theoretical concerns pertinent to specific sound contrasts are explored throughout the paper.

1.1 Language background

The genetic affiliation of the Mixtecan languages is provided in Figure 1. While Cuicatec and Triqui have notably little internal diversification, Mixtec has extensive internal diversification, possessing roughly sixty distinct varieties (recognized as separate languages in the ISO-639 standard) spoken in twelve pan-dialectal regions (Josserand 1983).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

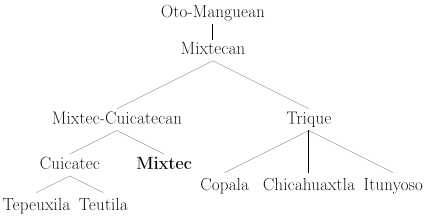


Figure 1 Mixtecan languages (Rensch 1976).

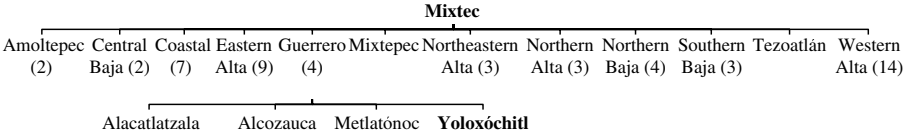


Figure 2 Mixtec languages (Josserand 1983).

As a result of this, there are a large number of languages, each of which is labelled 'Mixtec', but many of which are as distinct as modern-day Italian and Portuguese. The internal diversification of Mixtec ostensibly began in the late Preclassical period in Mexico (Josserand 1983: 458), giving it roughly the same time depth as the diversification of Romance languages, beginning 1800–2000 years ago (Adams 2007).[1] Yoloxóchitl Mixtec, the focus of the current study, is an endangered language spoken in the towns of Yoloxóchitl, Cuanacaxtitlán, Buena Vista, and Arroyo Cumiapa (Castillo García 2007), located approximately 20 miles north of the town of Marquelia, Guerrero, along the southeastern coast 'la costa chica' of Guerrero, Mexico. Within Cuanacaxtitlán and Buena Vista, the shift to Spanish is nearly complete among the younger generation, while in Yoloxóchitl speakers of all generations have maintained the language. There are approximately 4,000 speakers remaining, though many younger speakers are more dominant in Spanish than in Yoloxóchitl Mixtec. Figure 2 shows the diversification of Mixtec and its sub-family, Guerrero Mixtec, the larger grouping into which Yoloxóchitl Mixtec belongs.[2] The numbers next to each lower branch reflect the number of language variants within that branch. There is reasonably good mutual intelligibility across most of the Guerrero Mixtec languages (Castillo García 2007), but only approximately 30% mutual intelligibility between Guerrero and Southern Baja varieties spoken nearby (Lewis, Simons & Fenning 2013).[3]

Linguistic work on the Mixtec languages has a long history, with the first dictionary appearing in the colonial period (de Alvarado 1593). Even more than four hundred years ago, de Alvarado remarked on the extensive diversity of Mixtec languages. It was not until the mid-20th century, however, that a number of phonological descriptions of different Mixtec varieties were produced, among them work in Huajuapan Mixtec (Pike & Cowan 1967), Silacayoapan Mixtec (North & Shields 1977), Diuxi Mixtec (Pike & Oram 1976), Alacatlatzala Mixtec (Zylstra 1980), San Juan Mixtepec Mixtec (Pike & Ibach 1978), Ayutla Mixtec (Pankratz & Pike 1967, Herrera Zendejas 2009), Jicaltepec Mixtec (Bradley 1970), and Molinos Mixtec (Hunter & Pike 1969). Castillo García (2007) provides a phonological description of Yoloxóchitl Mixtec and is the basis of some of the phonological generalizations made here. With few exceptions (H. Gerfen 1996, C. Gerfen 2001, C. Gerfen & Baker

[1] The rationale for this specific time depth in Josserand (1983) is perhaps a bit unclear. Note that extensive internal diversification was already discussed in de Los Reyes (1596).

[2] Note that *Guerrero Mixtec* here refers to a genetic classification, not a geographic one, i.e. there are Mixtec varieties spoken in the state of Guerrero that are not part of this classification.

[3] Despite being spoken in Guerrero, the varieties which have less intelligibility with Guerrero Mixtec, according to Castillo García (2007: 11), belong to Southern Baja Mixtec (Josserand 1983).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Table 1 Yoloxóchitl Mixtec consonant inventory.

	Bilabial	Dental	Alveolar	Post-alveolar	Palatal	Velar	Labialized velar
Plosive	p	t̪				k	k^w
Pre-nasalized plosive	(^{m}b)		(^{n}d)			^{ŋ}g	
Affricate				tʃ			
Nasal	m		n				
Tap			(ɾ)				
Fricative		s̪		ʃ		(x)	
Approximant	β				j		
Lateral approximant			l				

2005, and Herrera Zendejas 2009), there are no studies examining the phonetics of Mixtec consonants.[4]

1.2 Phonological inventory

There are only 14 consonant segments in Yoloxóchitl Mixtec, shown in Table 1. Two segments appear in parentheses (/^{m}b ^{n}d/). We argue in Section 5.1 that these are not phonemic in the language, but allophonic. They are included in Table 1 because they are extremely common allophones and appear in a number of examples throughout the paper. The glottal stop is not included in the consonant inventory either. The dental fricative is represented as /s̪/ and not as /θ/, since the sound is always strictly post-dental (and not interdental) in Yoloxóchitl Mixtec and, to our ears, more similar to /s/ in other languages than it is to /θ/. Glottalization is phonologically-contrastive and considered to be a couplet/foot-level autosegmental feature (see below) which obligatorily occurs either intervocalically in monosyllabic words, e.g. /^{n}doʔ^{1}o^4/ 'basket', or preceding a medial voiced consonant, e.g. /jaʔ^{1}βi^3/ 'market, price' in disyllabic words. In our transcriptions, tone is written as a superscript number, where /1/ is low and /4/ is high.

While the focus of the current study is not a comprehensive phonological description of Yoloxóchitl Mixtec consonants, a few patterns are worthy of mention here. First, like most Mixtec languages, Yoloxóchitl Mixtec has (C)V(V) syllable structure, though utterance-initial words without an onset receive an obligatory and predictable [ʔ] onset. No consonant clusters or codas are permitted. Second, as in many other Oto-Manguean languages, bilabial plosives (prenasalized or voiceless unaspirated) are quite rare in Yoloxóchitl Mixtec. Third, the alveolar tap /ɾ/ occurs only in a few functional morphemes, such as the 3rd person masculine enclitic /=ɾa^3/ and the discourse particle indicating agreement, /ɾã^4/. Fourth, the vowel system is relatively simple in Yoloxóchitl Mixtec. There is a system of five vowels, each of which may be contrastively nasal or oral, shown in Table 2. The vowel system is explored in DiCanio et al. (2015).

The feature for which Yoloxóchitl Mixtec differs most from certain other Mixtec languages is in word-level constraints on phonological structure. In his early work on Mixtec, Kenneth Pike argued that the grammar of Mixtec as a whole is best understood with reference to a unit called the COUPLET (Pike 1944). A couplet is a phonological unit

[4] Moreover, given the extensive diversity of the Mixtec family, there is perhaps an inappropriate tendency to generalize the findings from work on specific dialects to other variants simply due to the shared use of the name *Mixtec*.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Table 2 Yoloxóchitl Mixtec vowel inventory.

	Front	Central	Back
Close	i ĩ		u ũ
Mid-close	e ẽ		o õ
Open		a ã	

consisting of either two syllables or morae. Pike coined this term in order to capture the generalization that most Mixtec words in the dialects he examined consisted of either CVV roots or CVCV roots.[5] Most words in Mixtec languages consist of units of this size. The couplet has been a useful unit in the analysis of phonological features in Mixtec, such as nasalization and glottalization (Hinton 1991, Macaulay & Salmons 1995, Macaulay 1996), as well as a useful template for explaining processes of sound change (Macken & Salmons 1997). While many words in Yoloxóchitl Mixtec consist of a bimoraic couplet and this is the MINIMAL size for content words, the language also possesses a fair number of trimoraic roots. In Yoloxóchitl Mixtec, words may be of the maximal size/shape CVCVCV, CVCVV, or CVCVʔV. Note that *CVCVCVV is not a possible shape, suggesting that the mora is the appropriate unit for measuring word size.

2 Stress and the phonetics of the word

Like many Oto-Manguean languages, words in Yoloxóchitl Mixtec have stress in addition to lexical tone. The typical Mixtec pattern is for stress to appear in the penultimate syllable of disyllabic lexical roots. In certain Mixtec varieties, stress shifts rightward when an enclitic is attached, suggesting that stress is assigned at the phonological word level, not at the couplet level. In Yoloxóchitl Mixtec, stress falls on the final syllable of the couplet (the lexical root), not on the penult. The main acoustic realization of stress is increased duration. Final vowels are approximately 30–50% longer than penultimate vowels in words produced in isolation (DiCanio, Amith & Castillo García 2012, DiCanio, Benn & Castillo García 2018). Moreover, there is a distributional asymmetry with respect to tone on roots, with far more tonal contrasts occurring on root-final than penultimate syllables (DiCanio, Amith & Castillo García 2014). These two factors suggest that root-final syllables are stressed compared to non-final syllables.

In several Mixtec languages, researchers have also observed a process of consonantal lengthening in couplet-medial position, e.g. CV**C**V, yet it is unclear if this process is related to stress at all. Longacre cites this process in his early work on Proto-Mixtecan (Longacre 1957). Later work on individual languages with penultimate stress has reported a similar pattern. For instance, consonantal lengthening occurs on the post-tonic (ultimate) syllable in Acatlán Mixtec (Pike & Wistrand 1974) and in Ayutla Mixtec (Pankratz & Pike 1967). In Ayutla Mixtec, medial stop consonants may also be optionally pre-aspirated. It is notable that both these varieties of Mixtec are spoken in Guerrero. Use of lengthening and preaspiration is also a characteristic of final stressed syllables in Itunyoso Triqui (DiCanio 2010), suggesting that this could be an areal/family feature. Given its proximity to these languages, Yoloxóchitl Mixtec should be examined to determine whether its medial consonants are also lengthened. Since there is phonological evidence that ultimate syllables carry stress in Yoloxóchitl Mixtec, it needs to be determined whether vowel lengthening co-occurs with onset consonant lengthening. In order to assess these issues, we examined the durational aspects of consonants in Yoloxóchitl Mixtec in a variety of contexts in carrier phrases.

As is common in phonetic descriptions, the terms ‘consonant duration’ and ‘vowel duration’ will be used as terms of convenience for the acoustic segments to be measured. Consonant duration will apply to stop closures, fricative noise, and resonances as appropriate.

[5] The Mixtec couplet is functionally equivalent to a bimoraic foot.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

This is not intended to deny the existence of consonantal information in vocalic formant transitions, which could be classified as belonging to the consonant and vowel equally. Similarly, the vowel duration will be the duration of the vocalic segment, even though that includes some consonantal information and excludes some vowel information (such as the shaping of fricatives noises; Soli 1981, Whalen 1983).

2.1 Method

Consonant duration was examined in three contexts in Yoloxóchitl Mixtec: in word-initial position in monosyllabic words, e.g. /**k**a^{3}a^{2}/ 'metal', in word-initial position in disyllabic words, e.g. /**k**a^{3}ka^{2}/ 'lime (stone)', and in word-medial position in disyllabic words, e.g. /ja^{3}**k**a^{2}/ 'granary'. With a few exceptions, one word for each of the three contexts was chosen for each consonant in the language. The exceptions were the consonant /^{ŋ}g/, which does not occur in word-initial position, and the consonant /ɾ/, which was excluded from analysis owing to its limited distribution in content words. In addition to this set of 43 words, an additional seven words were included in case the speakers were unfamiliar with particular target words. Each of these 50 words was elicited in a carrier sentence /ni^{1}-^{n}da^{1}ʔju^{1} = ɾa^{1} TARGET ka^{1}a^{3}/ 'PERF-shout=3s TARGET here', 'He/she shouted TARGET here'. Each carrier phrase was produced 5 times, for a total of 250 sentences per speaker. The vowels adjacent to the consonant in the target words were always /a/.

Speakers were recruited from within the Yoloxóchitl community and recorded in a quiet room in the nearby town of San Luis Acatlán. Eight male native speakers (mean age = 47.1 years) were recruited. It was not possible to recruit female speakers for this study. The research team was all male and no women were able to participate during the few days of recording. In total, 2000 utterances were analyzed (250 sentences × 8 speakers). The recordings were made on a Marantz PMD 670 solid-state recorder with a Shure SM10a head-mounted microphone at a sampling rate of 48 kHz. The sentences were produced in Yoloxóchitl Mixtec by the first author or fourth author as a prompt to the speaker. A repetition task was used here since the population is almost entirely illiterate in Yoloxóchitl Mixtec, and very little functional literacy in Spanish exists among the native speaker population as well. One limitation of this method is that speakers may imitate certain phonetic characteristics that are produced by the speaker rather than speak freely. An alternative would be to ask speakers to translate spoken Spanish sentences into Mixtec. However, this method involves the speaker switching language modes and, as a result, their responses may shift when compared to interactions composed entirely in their native language (see Drager, Hay & Walker 2010, Olson 2013). This is especially relevant in the context of research on endangered languages where accommodation towards a higher prestige language (Spanish) may be strong (see Bourhis & Giles 1977, Babel 2009).

Each of the target words was acoustically segmented and labeled in Praat (Boersma & Weenink 2016) by the second author. Five percent of the target consonants chosen at random were segmented by the first author and checked against the second author's segmentation. Consonant boundaries for all words fell within ±5 ms between the segmenters. For stop consonants, two separate portions were segmented: closure and VOT. Glides were assumed to begin/end at the midpoint between the onset/offset of the steady-state of the adjacent vowel's formants. The segmentation of the phonetic components of all other consonants (nasals, fricatives, prenasalized stops, affricates) was based on typical acoustic patterns: nasal murmur, frication, and closure.

The duration data were extracted automatically from the segmented sound files using a script written for Praat. The durational data were statistically analyzed with two separate repeated measures analyses of variance. In the first analysis, only the disyllabic words were examined. CONSONANT DURATION was treated as the dependent variable, while CONSONANT CLASS (stop, fricative, affricate, approximant, nasal, prenasalized stop) and POSITION (initial, medial) were treated as independent variables. CONSONANT DURATION here included the sum of the closure duration and VOT, i.e. all stop components. The model

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

evaluates the effect of position and consonant type on consonant duration. In the second analysis, only consonants in word-initial position were examined. CONSONANT DURATION was treated as the dependent variable, while CONSONANT CLASS and WORD SIZE (monosyllabic, disyllabic) were treated as independent variables. This model evaluates the effect of word size and consonant type on consonant duration. The interaction of each set of independent variables was included in the models. In both models, SPEAKER was treated as an error term. All statistical analyses were made using R version 3.33 (R Development Core Team 2016).

2.2 Results

There was a marginal main effect of POSITION on consonant duration in disyllabic words ($F(1, 2) = 13.2$, $p = .07$). Consonants produced in word-medial position were slightly longer than those produced in word-initial position (132.1 ms vs. 109.4 ms). The pattern was most apparent for obstruents (stops, fricatives, affricates) but weaker for sonorants (prenasalized stops, nasals, approximants). The effect of CONSONANT CLASS on CONSONANT DURATION was much stronger ($F(5, 35) = 15.9$, $p < .001$). A significant interaction between POSITION and CONSONANT CLASS was also found ($F(5, 35) = 4.3$, $p < .01$). Post-hoc Tukey tests showed a significant effect of POSITION for each consonant class ($p < .01$) except the prenasalized stops. Figure 3 shows the duration data by consonant class and context.

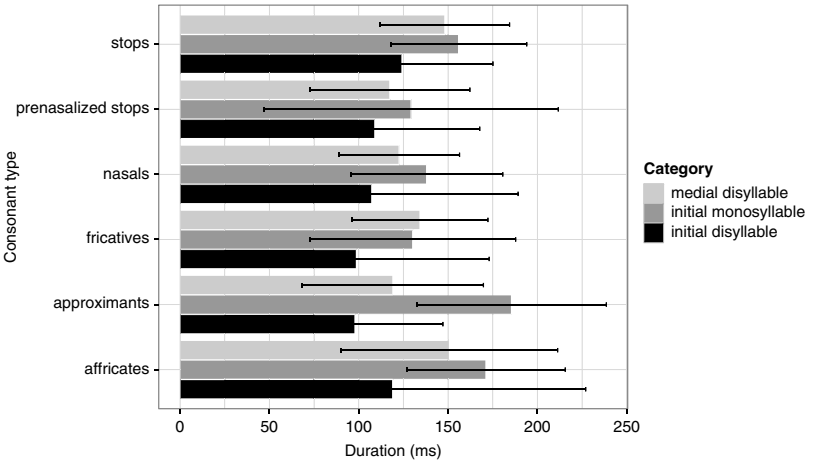


Figure 3 Consonant duration across context and word size (error bars reflect standard error).

For the model examining durational differences in word-initial position (the second statistical model described above), there was a significant main effect of WORD SIZE on CONSONANT DURATION ($F(1, 2) = 542.1$, $p < .01$). Consonants produced in monosyllabic words were longer than those produced in disyllabic words (151.7 ms vs. 109.4 ms). The durational difference here is approximately twice the size of the durational difference observed in the previous analysis examining the effect of POSITION (42.3 ms compared to 22.7 ms). The main effect of CONSONANT CLASS also emerged as significant within this model ($F(5, 30) = 16.2$, $p < .001$). A significant interaction between CONSONANT CLASS and WORD SIZE was also found ($F(5, 34) = 8.7$, $p < .001$). Similar to the findings above, post-hoc Tukey tests showed a significant effect of size for each consonant class ($p < .01$) except the prenasalized stops.

2.3 Discussion

The effect of WORD SIZE on consonant lengthening in the Yoloxóchitl Mixtec data is more robust than that of WORD POSITION. The pattern here reflects the well-known inverse

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

relationship between word size and syllable duration; the duration of individual syllables shortens as the number of syllables in the word increases (Jones 1942–1943, Lehiste 1970, Lindblom, Lyberg & Holmgren 1981), also known as POLYSYLLABIC SHORTENING (Turk & Shattuck-Hufnagel 2000, White & Turk 2010). The magnitude of the effect is unexpectedly large, with disyllabic onsets approximately 39% shorter than onsets in monosyllables. In contrast, Swedish shows a reduction of approximately 15% in consonant duration in stressed syllables for each additional syllable added to the word (Lindblom & Rapp 1973), while Hungarian shows a decrease in vowel duration of just 17–25% for each additional syllable added to the word (Tarnóczy 1965, cited in Lehiste 1970). Considering that the syllable nucleus typically undergoes greater length contraction in polysyllabic shortening than the onset (Turk & Shattuck-Hufnagel 2000), one anticipates an even smaller contraction for onset duration with polysyllabic shortening here. Thus, a 39% reduction in consonant duration in Yoloxóchitl Mixtec might seem large by comparison.

On the other hand, another study investigating the influence of stress and accent on polysyllabic shortening in English yielded findings similar to those observed for Yoloxóchitl Mixtec (White & Turk 2010). Those authors compared the degree of consonantal shortening for left-headed keywords (***mace*** > ***mas****on* > ***mas****onry*) and right-headed words (***port*** > *re****port*** > *misre****port***) where such words were either prosodically accented or unaccented. Syllable onset consonant duration was measured in this English study in a way similar to that of the current study. While White and Turk found little evidence of polysyllabic shortening in left-headed words, right-headed words showed a similar pattern to that observed for Yoloxóchitl Mixtec. The two languages are compared in Figure 4, where the onset duration for the stressed syllable is shown for two different word contexts.[6] Only consonant duration from the initial consonant in monosyllables and the medial consonant in disyllables is included as a comparison here. The appropriate comparison of English data against Yoloxóchitl Mixtec data is with the right-headed English words, as both the medial disyllable and initial monosyllables here are in the onset of a stressed syllable. Recall that words in Yoloxóchitl Mixtec are also obligatorily right-headed.

Examined this way, the magnitude of the effect of polysyllabic shortening on consonants in English varies substantially. In left-headed words, the onset consonant of the stressed syllable is just 6.6% longer in monosyllables than in disyllables. In right-headed words, this difference is 32%. For Yoloxóchitl Mixtec, this difference is 15%. The observed difference of 39% mentioned above reflects the durational difference between onsets in unstressed syllables in Yoloxóchitl Mixtec and monosyllabic words. Yet, the fairer comparison here is the duration of the onset of only the stressed syllable across words of different sizes.

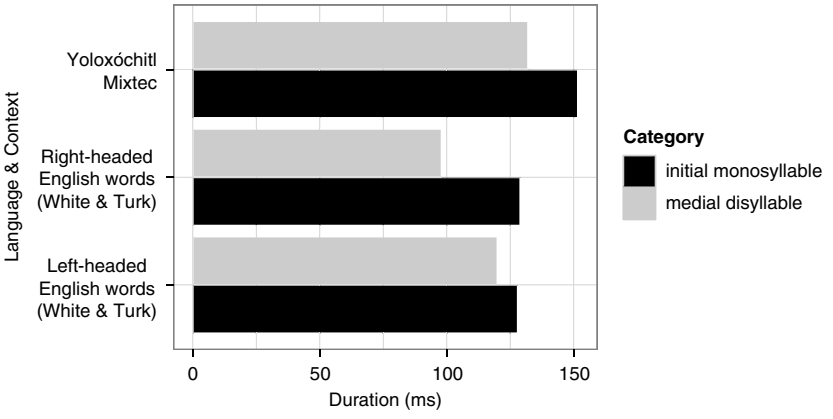


Figure 4 Consonant duration across language and word size.

[6] The English data here reflects words in the accented condition.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

The data from White & Turk (2010) show that the position of stress in the word mediates the effect of polysyllabic shortening. They also help clarify how final stress in Yoloxóchitl Mixtec influences consonant duration. There is a relatively large difference between penultimate and final consonant durations in disyllabic words (109 ms vs. 132 ms, or 21%), but a separate process of polysyllabic shortening on stressed syllables (comparing monosyllabic and disyllabic words) with a smaller magnitude from that observed for comparable English (right-headed) words. In sum, the polysyllabic shortening in Yoloxóchitl Mixtec is best compared to English when the position of prosodic prominence is made consistent (final) across languages. Seen this way, polysyllabic shortening in Yoloxóchitl Mixtec is not as sizable as one might expect from older findings in the literature, i.e. Lehiste (1970).

3 Voiceless stops

The stop inventory of Mixtec languages is consistent across much of the family (Josserand 1983). Yoloxóchitl Mixtec possesses four voiceless unaspirated stops at bilabial, dental, velar, and labialized velar places of articulation. There is no voicing distinction, although there is a series of prenasalized stops. There is one post-alveolar affricate in the language.

Many Mixtec variants also have patterns of stop lenition. In Yoloxóchitl Mixtec, voiceless velar stops undergo variable lenition and may be produced as voiced velar fricatives or frictionless continuants/approximants, [ɣ ɣʷ], or, in the case of the non-labialized velar, they may be deleted altogether (Castillo García 2007). However, it should be noted that these lenited variants are rare in the elicited data here. A similar pattern of velar lenition is discussed in Acatlán Mixtec (Pike & Wistrand 1974), Silacayoapan Mixtec (North & Shields 1977), San Juan Mixtepec Mixtec (Pike & Ibach 1978), Ayutla Mixtec (Pankratz & Pike 1967), Jicaltepec Mixtec (Bradley 1970), San Miguel el Grande Mixtec (Pike 1944), and in Alacatlatzala Mixtec (Zylstra 1980). This process of variable lenition in Yoloxóchitl Mixtec may be driven by tendencies in the production of velar stops, much in the way that variable lenition asymmetrically influences stops/affricates with particularly short closure duration more so than those with longer closure durations. For instance, in Itunyoso Triqui, post-alveolar affricates have the shortest closure duration among the lenis stops and affricates and, as a result, undergo greater lenition than other stops and affricates (DiCanio 2012).

In the phonetic description of stop contrasts, researchers typically focus on the durational aspects of the different components of the stop (closure and VOT) and evidence for place of articulation (via an examination of formant trajectories). It is these aspects of the stop and affricate system that we will examine here with a descriptive analysis. As duration varies with word size in Yoloxóchitl Mixtec (see Section 2 above), this factor is also considered.

3.1 Method

For the analysis of closure duration and VOT, the same data and methods discussed in Section 2 are applied here. CLOSURE DURATION was measured as the time from the abrupt cessation of formant amplitude and voicing on the preceding vowel to the burst release of the stop or affricate. VOT was measured as the duration of the burst in addition to aspiration. Figure 5 provides an example of how each stop was labeled. The label 'bs' here stands for 'burst'. While both bursts and a possible short period of aspiration were coded here below, they were combined as a single measure 'VOT' in the analysis.

For the analysis of place of articulation, the formant transitions from the stop release into the following vowel were analyzed. The first three formant values were extracted over a 100 ms duration at 10 ms intervals starting from the end of the aspiration duration (marked 'asp' in Figure 5). The window length was 25 ms, five formants were specified, and there was a 6.25 ms step size. Formant values were extracted automatically using a script written for Praat (Boersma & Weenink 2016) and statistically analyzed using R (R Development Core Team 2016).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

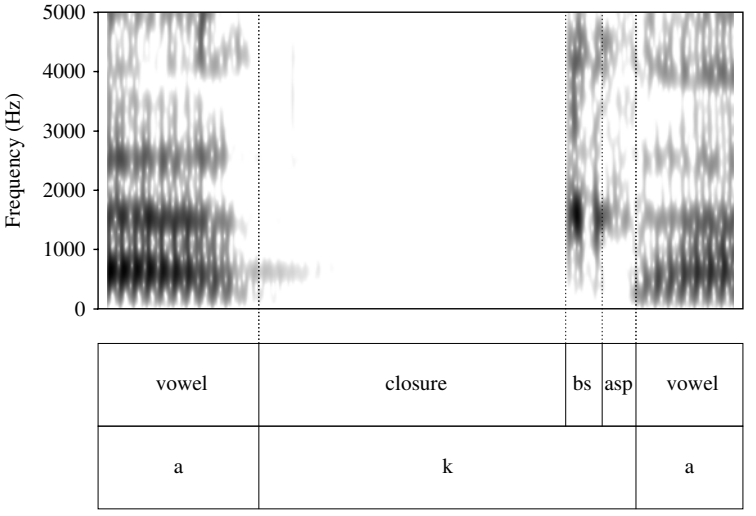


Figure 5 Labeling of stop consonant components (bs = burst, asp = aspiration duration).

Ten of the words in the data set contained stop consonants varying by position, and they were the items examined here. Following the methods described in Section 2.1, each word occurred in a carrier sentence that was repeated five times (though some speakers produced only four usable repetitions). A total of 427 tokens were examined across eight speakers. For the duration data, two statistical tests were run. In the first, the overall duration of the stops was compared. A single factor ANOVA was used with TOTAL DURATION treated as the dependent variable and STOP POA (place of articulation) as the independent variable. SPEAKER was treated as an error term. In the second, the duration of individual components of the stops (closure, VOT) were compared. A two-factor ANOVA was run with PERCENT DURATION treated as the dependent variable and with STOP POA and COMPONENT (with two levels: closure, VOT) treated as independent variables. SPEAKER was treated as an error term. Estimating the duration results in terms of percentages provides an additional way of normalizing the data across different speakers. For the statistical analysis of formant trajectories, three sets of two factor analyses of variance were used, one for each of the first three formants. In each model, the FORMANT VALUE was treated as the dependent variable, while the STOP POA and TIME INTERVAL (10 samples of 10 ms) treated as independent variables. SPEAKER was treated as an error term in these models as well.

3.2 Results

3.2.1 Duration

The analysis of the TOTAL DURATION of stops (closure plus VOT) did not reveal a significant main effect of STOP POA on duration ($F(3, 21) = 0.94, p = .438$) (see Table 3). For the analysis of the duration of the stop components, we are particularly interested in how stop components (closure, VOT) vary in percent duration with the stop place of articulation, not with how each of the components varies in duration with each other (e.g. VOT is usually shorter than the stop closure duration). So, the durational aspects of each component are examined via the interaction of STOP POA and COMPONENT. The results from the ANOVA found a significant and strong interaction between STOP POA and COMPONENT on the percent duration of the stop components ($F(3, 21) = 25.8, p < .001$). The CLOSURE DURATION of the bilabial stop was longer than that of the dental stops, which was in turn longer than that of the velar and labialized velar stops. As expected from the lack of an effect on total duration, the exact opposite pattern was true with respect to VOT, i.e. velar > dental > bilabial. The possible

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Table 3 Stop duration across contexts (in ms). Numbers in parentheses are standard deviations.

	[p]	[t̤]	[k]	[k^{w}]
Closure duration ($p < .001$)	135.9 (36.1)	124.7 (48.2)	108.4 (42.7)	122.3 (35.4)
VOT ($p < .001$)	11.3 (4.6)	16.9 (6.8)	26.6 (7.8)	31.5 (12.2)
Total duration (*ns*)	147.2	141.6	135.0	153.8

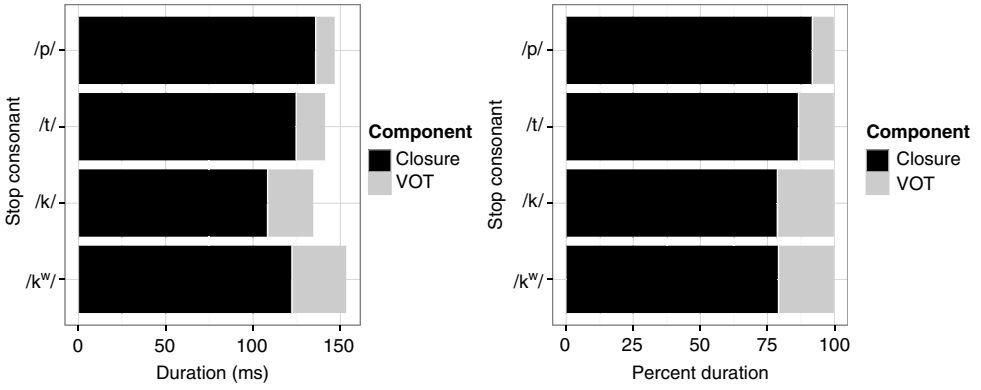


Figure 6 Stop duration and duration percentage.

influence of position (word-initial, word-medial) on component duration percentage was also considered in this statistical model, but it was not significant. Thus, the stop duration results are pooled together across positions in Table 3 and in Figure 6.

A post-hoc Tukey's HSD test was applied to the ANOVA model to examine for which place of articulation closure duration varied the most. Most of the possible pairings ($VOT_{/k/}$–$VOT_{/p/}$, $VOT_{/k/}$–$VOT_{/t/}$, $VOT_{/kw/}$–$VOT_{/p/}$, $VOT_{/kw/}$–$VOT_{/t/}$, $VOT_{/p/}$–$VOT_{/t/}$, $Closure_{/k/}$–$Closure_{/p/}$, $Closure_{/k/}$–$Closure_{/t/}$, $Closure_{/kw/}$–$Closure_{/p/}$, $Closure_{/kw/}$–$Closure_{/t/}$, $Closure_{/t/}$–$Closure_{/p/}$) were significant ($p < .001$), though no significant differences were found for the $VOT_{/k/}$–$VOT_{/kw/}$ and $Closure_{/k/}$–$Closure_{/kw/}$ pairs. These latter two cases are to be expected, though, as [k] and [**k^{w}**] share the same place of articulation.

Pooled across speakers, these results show a strong, largely reciprocal effect of place of articulation on the stop component, i.e. the stops produced at a more anterior place of articulation had a longer closure duration and the shortest VOT, while stops produced at a more posterior place of articulation has a shorter closure duration and a longer VOT. While the place effect on VOT appeared robust across all speakers, there was some variation observed with respect to whether VOT values for /p/ were shorter than those for /t̤/. The between-speaker differences are shown in Figure 7.

In Figure 7, we observe the general trend where the VOT for velar stops is longer than that of anterior stops across most speakers, but the differences between the VOT of bilabial and dental stops are small or restricted to certain speakers. There is virtually no difference observed for three of the eight speakers (EGS, MFG, MSF) and VOT for the dental stop is almost as long as that of the velar stops for three of the eight speakers (FNL, MSF, VRR). For only three of the speakers does the canonical 'bilabial < dental < velar' VOT duration pattern hold (CTB, ECG, RCG). Variation in dental stop VOT duration across speakers appears to be most responsible for the inconsistent pattern here. While bilabial stops and velar stops have more consistent VOT values, dental stops have VOT values either similar to those of the velars (for speakers FNL, VRR), similar to those of the bilabial stop (for speaker EGS), or somewhere in-between (the remaining speakers). Taking an average VOT value across speakers will show the canonical 'bilabial < dental < velar' VOT duration pattern, but this pattern does not seem to be consistent across all speakers.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

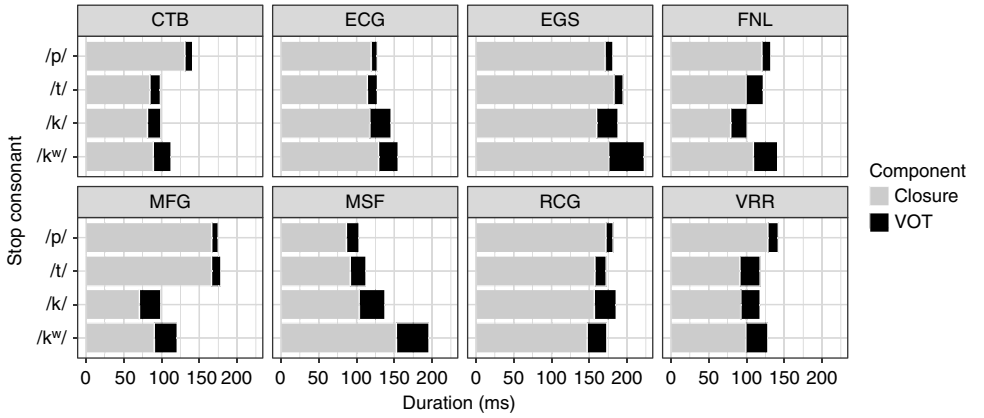


Figure 7 Stop component duration across speakers.

3.2.2 Formant trajectories

Formant transitions between the stop and the following vowel are shown in Figure 8. Analyzing the effect of STOP POA on formant values, a significant main effect was found for F1 ($F(3, 21) = 25.7$, $p < .001$), F2 ($F(3, 21) = 20.5$, $p < .001$), and F3 ($F(3, 21) = 3.4$, $p < .05$). A post-hoc Tukey's HSD test was applied to each of these models to evaluate which place of articulation had greatest influence on formant transitions. While all stops had significantly different F1 values from each other, the largest significant comparison was between the labialized velar stop and the remaining stops. The same pattern was found for F2. The largest significant comparison for F3 was between both labial stops ([p] and [kʷ]) and the non-labial stops. In Figure 8 the robust pattern for the labialized velar stop is not only visible early in the acoustic transition between the stop and the following vowel, but occurs long into the vowel's production. The lowered F2 values following [kʷ] suggest that the following vowel undergoes significant retraction in this context.

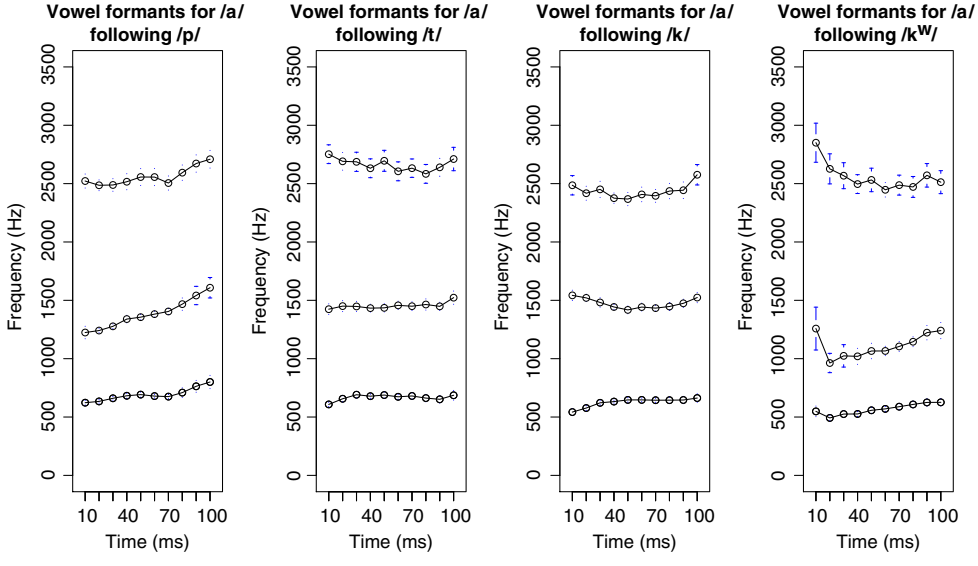


Figure 8 Stop–Vowel Formant Transitions.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Analyzing the effect of TIME on formant values, a significant main effect was also found for F1 ($F(1, 4) = 12.3, p < .05$), F2 ($F(1, 4) = 39.2, p < .01$), and F3 ($F(1, 4) = 38.2, p < .01$). This finding reflects the observation that formants change in time in the transition between the stop and the following vowel. As such, it is an unsurprising finding. Yet, like the examination of the durational differences above, we are particularly interested in the interaction between STOP POA and TIME, as this reflects differences in the temporal extent in which stops of differing place of articulation influence formant values. A significant interaction was found for F1 ($F(3, 21) = 20.6, p < .001$), F2 ($F(3, 21) = 6.1, p < .01$), and F3 ($F(3, 21) = 19.7, p < .001$). Post-hoc Tukey's HSD tests were applied to the ANOVA models to evaluate how the temporal extent of the formant transition varied with place of articulation. In general, formant transitions for both velar stops were longer in duration than those for dental and bilabial stops. The F1 and F2 transitions following [k] extended up to 40 ms into the vowel, while F3 was significantly lowered throughout the 100 ms window. The formant transitions following [k^{w}] extended well into the vowel, but such effects were strongest for F2, as previously mentioned.

3.3 Discussion

3.3.1 Duration

The durational characteristics of stops in Yoloxóchitl Mixtec are similar to those found in many other languages. In particular, stop closure duration increases and VOT decreases as stop place of articulation becomes more anterior. This inverse relationship between aspiration duration and closure duration in relation to place of articulation is well-known (Lisker & Abramson 1964, Umeda 1977, Weismer 1980, Maddieson 1997, Cho & Ladefoged 1999, Gordon et al. 2001). The observation that stop VOT increases with more posterior places of articulation was first observed in classic work by Lisker & Abramson (1964), where they showed the effect across a range of languages. However, the first attempt to link this pattern with that of closure duration was made by Weismer (1980). Citing this later work, Maddieson (1997) suggests that the relation between closure duration and VOT is maintained by a stable abduction–adduction cycle of the vocal folds for all stops. Stops with an earlier release and shorter closure duration only have longer VOT by virtue of there existing a stable duration of devoicing. This has been shown for English (Weismer 1980), Western Apache (Gordon et al. 2001: 427), Chickasaw (Gordon, Munro & Ladefoged 2000), Waima'a (Stevens & Hajek 2004), and for several languages in a cross-linguistic survey on VOT (Cho & Ladefoged 1999).

When compared with the language data in Cho & Ladefoged (1999), the average VOT values for Yoloxóchitl Mixtec appear typical among languages with voiceless unaspirated stops (see Figure 9). The VOT for bilabial stops is approximately 10 ms, the VOT for alveolar or dental stops lies between 10 ms and 25 ms, and the VOT for velar stops lies between 25 ms and 50 ms. Certain languages, like Hupa, Montana Salish, Navajo, and Yapese have slightly longer VOT values for velars than the other languages shown.

The comparative data also shows cross-linguistic variation in whether bilabial stops have significantly different VOT from alveolar/dental stops. Cho & Ladefoged (1999: 220) find no general trend here across languages, but certain individual languages, like Chickasaw, Gaelic, Hupa, and Tsou, suggest it to be a pattern. However, as we observed above with the Yoloxóchitl Mixtec data, this may result as an effect of averaging across those speakers who have particularly long VOT values for alveolar/dental stops with those who have shorter values. If this were true, the results would produce a lack of consistency in the relative VOT difference between bilabial and alveolar/dental stops within a given language, a finding which accords with the observation made by Cho & Ladefoged (1999).

In the Yoloxóchitl Mixtec data, velar stops had the shortest closure durations of all stops. In some rare cases, most notably for speaker MFG, closure was absent and the stop was produced as a voiceless fricative, [x]. These comprised half of the tokens for this speaker

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

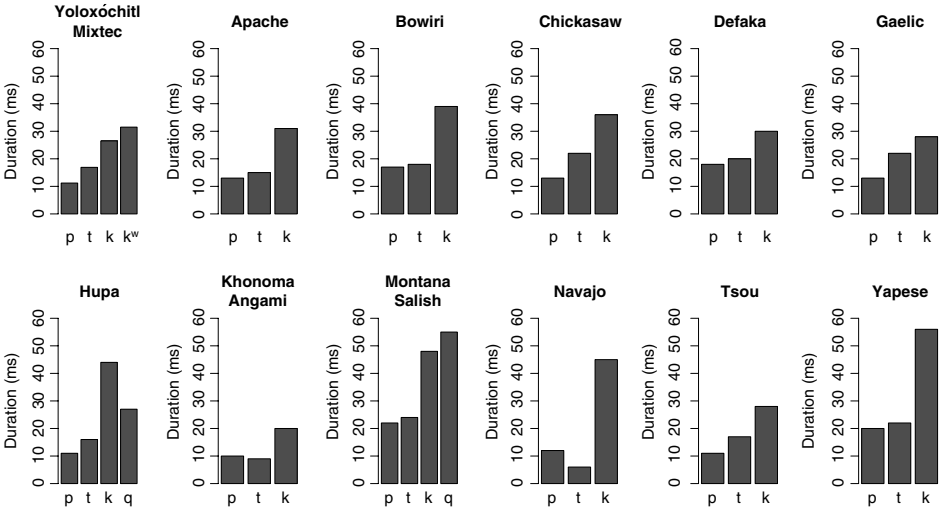


Figure 9 Unaspirated stop VOT values across languages. (Additional, non-Mixtec data from Cho & Ladefoged 1999.)

(7/14 repetitions). This speaker also had the shortest velar stop closure duration of all speakers. This finding fits well with the observation that there is variable lenition of velar stops in Yoloxóchitl Mixtec. As mentioned previously, voiceless velar stops undergo variable lenition or may be deleted altogether in Yoloxóchitl Mixtec (Castillo García 2007) as well as in other Mixtec variants. This tendency may arise out of a general pattern of consonant undershoot (Gay 1981), where the tendency for closure to be shorter in velar stops results in the failure to achieve dorso–velar contact. Stops that are generally produced with shorter overall closure duration are more likely to undergo processes of lenition in certain prosodic contexts or during faster speech rates. A similar type of process occurs in Itunyoso Triqui, a related Mixtecan language. Singleton affricates have the shortest closure duration of all obstruents and undergo both spirantization and passive voicing more often than other stops do (DiCanio 2012).

3.3.2 Stop formant transitions

In the context of low vowels such as /a/, one anticipates different formant transitions for each place of articulation (Stevens 2000). For bilabial stops, all formants should rise in frequency following stop release. For alveolar stops, F2 and F3 should decline, but F1 should rise. For velar stops, F1 and F3 should rise, but F2 should decline. Generally speaking, labialization should cause formant lowering due to the increased constriction at the most anterior point in the front cavity. This would have the effect of causing a rising trajectory for F1 and F2 on a following low vowel (Ladefoged & Maddieson 1996, Suh 2009).

The data in Yoloxóchitl Mixtec show several similarities to these predictions, but a few minor differences as well. The formant transitions following the release of [p] closely match the predicted values. F1 and F2 show a rising trajectory, while F3 is relatively flat. Meanwhile, the formant transitions following the release of [t̪] are somewhat different. F1 rises and F3 declines, as predicted, but F2 is relatively flat and does not exhibit a characteristic falling trajectory predicted from previous work. The formant transitions following [k] are similar to the predicted trajectories, namely F1 rises while F2 falls. However, F3 falls slightly instead of rising out of the velar constriction. Finally, the formant transitions following [kʷ] match the predictions in overall trajectory. F1 and F2 both rise throughout the 100 ms interval following aspiration. The F2 dip in the initial 20 ms of the vocalic segment is notable though. This dip reflects the transition from the velar constriction into the glide portion of the stop.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

A labiovelar glide will contain a particularly low F2 target which will cause an initial fall prior to the rise in F2.

These results suggest that both the formant measurements and the controlled nature of the elicited data in Yoloxóchitl Mixtec match the predictions from the established literature. The few exceptions also have ready explanations. First, the flat F2 following [t̪] reflects the fact that this stop is dental instead of alveolar. Dental consonants have a lower F2 locus, at around 1400 Hz, than alveolar consonants do, where the F2 locus is centered around 1800 Hz (Delattre 1968). The location of F2 on the following vowel [a] in our data is between 1400–1500 Hz. Thus, the stop–vowel formant transition is flat. Second, though one anticipates a rise in F3 for velar stops, the data here show a level or falling pattern. At this time, we have no explanation for this pattern. Finally, the 'dip–rise' pattern for the F2 transition for the labialized velar stop is explained by speaker variability in stop production. Some speakers produced the labialized velar stop as a stop+vowel sequence, e.g. [kŭa]instead of [k^{w}a]. The net effect of such a production would be to delay labialization and lower F2 in the stop–vowel transition.

4 Fricatives

The main differences in the consonant inventories across Mixtec languages are found within their fricatives. Diuxi and Huajuapan Mixtec each have six fricatives (Pike & Cowan 1967, Pike & Oram 1976), while most other variants have three or four fricatives.[7] Yoloxóchitl Mixtec contrasts only two fricatives: /s̪/ and /ʃ/. However, virtually all described Mixtec languages contain a dental or alveolar sibilant, /s/, and a post-alveolar sibilant, /ʃ/. Where the fricative systems differ is in whether they contain an interdental series, /θ ð/ (which may occur in lieu of dental or alveolar fricatives), and whether they contain velar/glottal fricatives as well, /x h/. In Yoloxóchitl Mixtec, there is some degree of free variation between the post-alveolar fricative /ʃ/ and a velar fricative /x/ (Castillo García 2007). Certain speakers freely produce [x] in lieu of [ʃ] or vice versa. However, the more typical production for this contrast is as a post-alveolar fricative (ibid.). Interestingly, the rarity of the velar fricative /x/ is a shared property of Ayutla Mixtec (Pankratz & Pike 1967), Alacatlazala Mixtec (Zylstra 1980), and Huajuapan Mixtec (Pike & Cowan 1967). The first two of these variants are also spoken in Guerrero, Mexico.

There is a large set of acoustic features which may distinguish voiceless fricatives in the languages of the world. With some exceptions, such as Nartey (1982) and Gordon, Barthmaier & Sands (2002), most of the acoustic analyses of fricatives have been based on relatively well-described languages like English. Among the studies that investigate fricative acoustics, many have substantial methodological differences. Such differences have, in part, led to varying conclusions regarding the utility of certain acoustic measures. While Yoloxóchitl Mixtec contrasts only two fricatives, our analysis here has paid careful attention to the methodological aspects when analyzing their different acoustic realizations.

4.1 Method

The fricative data come from the same data set discussed in Section 2. Three target words for each fricative were produced five times in carrier sentences by each speaker, for a total of 240 analyzed fricative durations (3 words × 2 fricatives × 5 repetitions × 8 speakers). Six different measures were collected from each fricative: duration, average intensity, and the

[7] The voiced bilabial fricative, /β/, is excluded here, as it typically patterns with glides or sonorants in Mixtec languages and is rarely produced with noticeable frication. It is better described as a frictionless continuant.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

first four spectral moments. The spectral moments characterize the distribution of aperiodic energy in the spectrum in terms of the center of gravity, standard deviation, skewness, and kurtosis – moments 1–4, respectively.

The adjacent vocalic context may cause substantial differences in the acoustics of fricatives, especially if there is lip-rounding (Whalen 1981, McGowan & Nittrouer 1988). While all the fricatives produced in the carrier sentences had adjacent /a/ vowels, the spectral moments were based on the fricative noise between the initial and final 10% of the fricative duration. Extracting measures from this middle 80% prevents low amplitude frication and V–C and C–V boundaries from being analyzed. Furthermore, spectral moment estimates are sensitive to the frequency range used for the analysis. For instance, the center of gravity of a flat spectrum in an analysis ranging from 0–20 kHz is located at 10 kHz, while the same center would be located at 5 kHz were the range only 0–10 kHz. High amplitude lower frequencies, unrelated to the main source of the fricative energy may also introduce bias in spectral moment estimation. To control for these effects, all speech signals with fricatives were resampled from 48 kHz to 44.1 kHz and high-pass filtered with a cut-off of 300 Hz, as per Maniwa, Jongman & Wade (2009). This method ensures that the frequency range for spectral analysis is not influenced by low frequency energy that is often present for reasons other than the fricative itself.

Previous descriptive work on the acoustics of fricatives has estimated the spectral center of gravity by averaging together single window fast Fourier transforms (FFTs) of fricative spectra across words and speakers (Gordon et al. 2001, Gordon, Barthmaier & Sands 2002). One consequence of this procedure is that single window FFTs extracted from fricative midpoints result in spectral estimates with a large error. As the signal is aperiodic, spurious peaks may occur across the spectrum. Such error is proportional to the mean amplitude at that frequency, and increasing the window length does not, in fact, change the error (Shadle 2012). One solution to this problem is to take several discrete Fourier transforms (DFTs) across a fricative duration, compute their average amplitude distribution, and then calculate the spectral moments (ibid). This procedure, called time-averaging, was performed here automatically with a script written for Praat (Boersma & Weenink 2016). Six 15 ms windows were averaged at equidistant intervals over the center 80% of each fricative in the data. From these average windows, spectral moments were calculated.

As there are six different acoustic measures that may possibly distinguish fricatives, all were statistically analyzed using linear discriminant analysis (LDA). This method ranks the utility of the particular measurement in the categorization of the contrast. Each measurement was also separately analyzed using a repeated measures one-factor ANOVA with fricative (two level) as the independent variable and speaker as the error term. Prior to statistical analysis, all extracted values were normalized using a standard z-score normalization.

4.2 Results

All the acoustic features analyzed were statistically significant. However, the results from the LDA analysis showed that spectral center of gravity was the strongest predictor, followed by standard deviation and skewness. The strength of these features is evaluated by examining the coefficient of the first linear discriminant. Table 4 shows the statistical results for each measurement as well.

Data showing the values for each measure is provided in Figure 10 with spectra in Figure 11. As per previous research on fricative acoustics, we have visualized the non-normalized values (see Gordon et al. 2002, Maniwa et al. 2009), though the normalized values are provided in a table in the appendix. The center of gravity was much higher for the dental [s̪] than for the post-alveolar [ʃ], with mean centers of gravity of 6286 Hz and 4349 Hz, respectively. The distribution of spectral energy for [s̪] was also much more dispersed than for [ʃ], as the former had a much higher average standard deviation (2537 Hz) than the latter

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Table 4 Fricative acoustic cues ranked by strength.

Cue	$F(1,7)$	LDA coefficient
Center of gravity	28.1, $p < .01$	-0.925
Standard deviation	16.1, $p < .01$	-0.799
Kurtosis	5.7, $p < .05$	-0.463
Skewness	20.3, $p < .01$	0.433
Duration	17.3, $p < .01$	0.235
Intensity	0.1, $p = .74$	0.172

(1658 Hz). The shape of the distribution of energy differed between each fricative, with an even average skewness (no skew) for /s̪/ (0.003) and a positive (leftward) average skewness for [ʃ] (2.2). The kurtosis values for each fricative also differed, with notably more peaked spectral energy for [ʃ] (mean = 13.24) than for [s̪] (mean = 1.29). While the post-alveolar fricative was longer and had higher intensity than the dental fricative, these differences were rather small in comparison with the spectral moments in the LDA model.

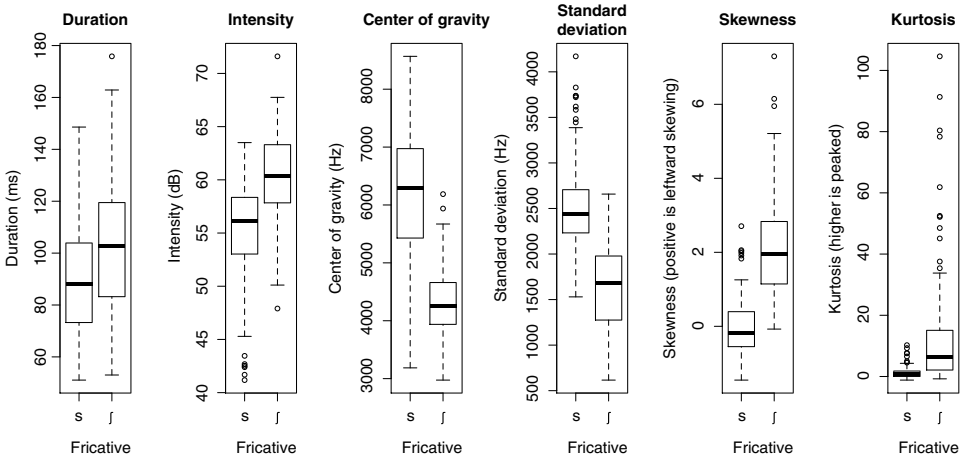


Figure 10 Fricative features. (Quartiles are plotted. Dark lines indicate median values.)

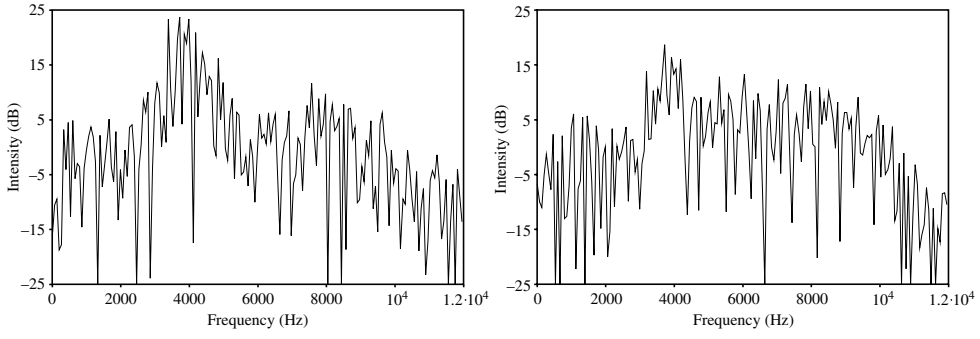


Figure 11 Fricative spectra averaged from one speaker (post-alveolar on left, dental on right).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

4.3 Discussion

The acoustic characteristics of the fricative spectra largely match values found in the literature for other languages. One expects a higher center of gravity for the dental [s̪] than for the post-alveolar [ʃ] given the decreased length of the front cavity for the former (Stevens 2000, Johnson 2012). The increased dispersion of energy for the dental fricative seen in Figure 11 is somewhat different from that of alveolar sibilants, where the resonance of the longer front cavity filters the aperiodic noise and produces greater overall kurtosis. The spectrum here is more like the relatively diffuse spectral energy observed with voiceless interdental and labiodental fricatives. In fact, such a distinction is visible in spectra comparing dental and alveolar sibilants in the literature (Jassem 1968). These observations confirm the transcriptional data that this fricative is dental and not alveolar.

5 Sonorants

5.1 Nasals and prenasalized stops

Yoloxóchitl Mixtec contrasts two underlying nasal consonants, /m/ and /n/, but three at a surface phonetic level. The additional palatal nasal described in Castillo García (2007) is better analyzed as the result of a phonological process in which a glide nasalizes when it precedes a phonologically nasal vowel via leftward nasal assimilation. This phonological argument is supported by two pieces of evidence. First, bilabial and alveolar consonants freely occur in any position in Yoloxóchitl Mixtec words, while the palatal nasal surfaces only in the onset position of word-final syllables, e.g. **C**V(ʔ)V#, CV**C**V#, CVCV**C**V#, or CV**C**V(ʔ)V#. The contrast between oral and nasal vowels only occurs in the final syllable of the word; this is the same context where one finds the palatal glide. Second, there is variability in the degree of oral closure observed in the palatal nasal among speakers. While some produced a nasal stop with complete closure, others produced a nasalized palatal glide, [ɲ̆].

In addition to the nasal stops, there are three prenasalized stops in Yoloxóchitl Mixtec, but each has a different distribution among native words. The alveolar stop, /ⁿd/, surfaces in any position, while the bilabial stop, /ᵐb/, only surfaces in polysyllabic words. The velar stop /ᵑg/ only surfaces word-medially in polysyllabic words and only in a few lexical items in the language. The bilabial and velar prenasalized stops are marginal in many Mixtec languages (Josserand 1983), especially those spoken in the state of Guerrero, e.g. Ayutla Mixtec (Pankratz & Pike 1967) and Alacatlatzala Mixtec (Zylstra 1980).

5.1.1 The phonological status of prenasalized stops

There are two major perspectives on the phonological status of prenasalized stops in the literature. The first view holds that prenasalized stops are surface allophones of nasal consonants (Marlett 1992). This NASAL ALLOPHONE argument is supported by the observation that in many Mixtec variants, only oral vowels surface after prenasalized stops and only nasal vowels surface after nasal stops. There is a contrast between oral and nasal vowels in Mixtec (Pankratz & Pike 1967, Josserand 1983, Castillo García 2007), thus allowing an interpretation in which prenasalized stops are simply conditioned by vowel nasalization. The second view holds that prenasalized stops are surface allophones of voiced stops which have undergone hypervoicing (Iverson & Salmons 1996). The argument here is mainly supported by the observations that, in two closely related Mixtec languages, the distribution of prenasalized stops follows from typical aerodynamic constraints on voicing (Ohala 2001). Continuous voicing requires the cavity behind the constriction to retain lower air pressure (intra-oral pressure) than the cavity below the glottis (sub-glottal pressure). Once the intra-oral air pressure is equivalent to the sub-glottal air pressure, voicing ceases. This cessation of voicing occurs more quickly with smaller back cavities. Since more posterior places of articulation

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

are produced with smaller back cavities than anterior places of articulation, the aerodynamic constraints on voicing for posterior consonants, such as velars, are greater than those for anterior consonants, such as bilabials. One way to maintain voicing during closure is to lower the velum and prenasalize the voiced stop, effectively venting the intra-oral air pressure. This is termed HYPERVOICING. Note that many Mixtec languages possess a prenasalized stop series but no voiced stop series; the former may be considered hypervoiced variants of voiced stops.

These two views have been mainly supported with phonological evidence and transcriptional phonetic observations. However, each makes different predictions in how prenasalized stops are produced at the phonetic level. The Nasal Allophone hypothesis predicts that the oral closure of prenasalized stops reflects an early velar raising prior to the consonant's release in order to maintain vowel orality. In essence, the nasal consonant is 'post-stopped' for the sake of the following oral vowel. Thus, one anticipates that there will be negligible differences in the duration of nasal and oral closure across different places of articulation. After all, under this view, the prenasalized stops are all simply nasal consonants at an underlying level. The hypervoiced allophone hypothesis predicts that the oral closure of prenasalized stops will vary with the ease in which voicing can be maintained with place of articulation. As it is more difficult to maintain voicing for velar consonants than for alveolars and bilabials, one anticipates that the nasal portion of the nasalized velar will be relatively longer than the oral portion. While these two hypotheses are presented as opposing, universal viewpoints, there is nothing that precludes each from occurring within different Mixtec variants if they are, instead, alternative ways in which prenasalized stops can arise.[8] A third hypothesis is also possible – that the nasal and oral stop sequence reflects a consonant cluster. This hypothesis is problematic, however, as it necessitates positing consonant clusters in a language which otherwise lacks them.

The phonological patterning of nasal consonants and prenasalized stops in Yoloxóchitl Mixtec mostly argues in favor of the Nasal Allophone hypothesis. Consider that there is an alternation in Yoloxóchitl Mixtec involving the process of leftward nasal assimilation with clitics, shown in Table 5.[9]

When a vowel-initial clitic containing a nasal vowel attaches to a Yoloxóchitl Mixtec word, it replaces the final non-high vowel of the stem. Front high vowels become palatal glides while back high vowels become labiovelar glides. On /CVʔV/ stems, the entire vowel sequence changes in quality and nasalization to match that of the clitic via a process of

Table 5 Regressive nasal assimilation in Yoloxóchitl Mixtec.

Root	Clitic	Surface form	Interlinearization	Gloss
a. koʔ^{3}o^{3}	= õ^{4}	kõʔ^{3}õ^{4}	drink=2S	'You drink'
b. koʔ^{3}o^{3}	= e^{4}	k^{w}eʔ^{3}e^{4}	drink=1INCL	'We drink'
c. ka^{3}ta^{3}	= õ^{4}	ka^{3}tõ^{4}	dig.out=2S	'You dig out'
d. ka^{3}ta^{3}	= e^{4}	ka^{3}te^{4}	dig.out=1INCL	'We dig out'
e. ka^{3}nã^{3}	= õ^{4}	ka^{3}nõ^{4}	call=2S	'You will call'
f. ka^{3}nã^{3}	= e^{4}	ka^{3n}de^{4}	call=1INCL	'We will call'
g. nã^{3}mã^{3}	= õ^{4}	na^{3}mõ^{4}	change=2S	'You change'
h. nã^{3}mã^{3}	= e^{4}	na^{3m}be^{4}	change=1INCL	'We will change'

[8] Note that Marlett groups together many diverse Mixtec languages, so his view is essentially a pan-Mixtec hypothesis. By contrast, the HYPERVOICED ALLOPHONE perspective is motivated specifically by patterns in Chalcatongo and San Miguel el Grande Mixtec, two closely related variants.

[9] Most clitics have allomorphs conditioned by vowel height/backness, not shown in the table here. Thus, the 2s clitic is /= ũ^{4}/ after high vowels, but /= õ^{4}/ after non-high vowels, see Palancar, Amith & Castillo García (2016).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

leftward assimilation. For clitics with vowel nasalization, like the 2s clitic, the outcome of this process is a change in nasalization in the final vowel on the stem, as seen in the cliticized forms in (a) and (c) in Table 5. Note that all vowels following a nasal consonant are always nasal in Yoloxóchitl Mixtec, as we observe in the stem forms in (e)–(h). When the 2s clitic attaches to these words, the vowel changes only in quality, as in (e) and (g). Yet, when the 1INCL clitic attaches, which contains an oral vowel, it causes a change where the preceding nasal consonant is now a prenasalized stop. Recall that nasal vowels never surface after prenasalized stops in Yoloxóchitl Mixtec. If cliticization targets only vowel quality and vowel nasality, then one must assume that the nasal consonants and prenasalized stops are allophones conditioned by the nasality of the following vowel. This analysis favors the nasal allophone hypothesis.

Yet, despite this alternation, there are still phonological differences between the prenasalized stop series and the nasal consonant series. Whereas the plain nasals have an unrestricted distribution, /^{m}b/ is restricted to polysyllabic words and /^{ŋ}g/ to the final syllable in polysyllabic words (Castillo García 2007). Moreover, unlike the bilabial and alveolar nasal series, there is no plain velar nasal in Yoloxóchitl Mixtec which corresponds with the prenasalized stop. Velar nasals do not undergo this alternation. Thus, while there is a morphophonological alternation with clitics that supports the nasal allophone hypothesis, the distributional evidence is less compelling.

However, both bilabial and velar prenasalized stops are rather rare in Yoloxóchitl Mixtec. Out of an extensively tagged lexicon of 2029 words (Amith & Castillo García, no date), only 11 contain a velar prenasalized stop and 17 contain a bilabial prenasalized stop. Alveolar prenasalized stops are vastly more common, appearing in 542 lexical items. In general, bilabial oral stops, bilabial prenasalized stops, and bilabial nasals are rare in Oto-Manguean languages and none are reconstructed for Proto-Mixtec (Josserand 1983). The same is true for velar prenasalized stops, which occur mainly in loanwords and in historical compounds consisting of a nasal vowel and a following velar onset, e.g. /le^{3ŋ}gu^{3}/ ‘lame’ from Spanish ‘rengo’, and /i^{3ŋ}ga^{2}/ ‘another’ from /ĩĩ^{3}/ ‘one’ + /ka^{2}/ (collocation). While the historically marginal nature of the velar prenasalized stop does not explain why it is exceptional phonologically, it is also the case that contrasts in loanwords and derived from compounds may not fully participate in native phonological alternations.

Phonetic research on nasals and prenasalized stops may offer additional clues as to the nature of this contrast in Yoloxóchitl Mixtec. Two patterns are relevant here: (i) a comparison between nasal consonant duration and prenasalized stop duration, and (ii) a comparison between the nasal and oral closure durations for prenasalized stops across place of articulation. The first investigation examines if prenasalized stops are longer than nasal stops. If prenasalized stops are consonant clusters, then they should be longer in duration than single consonants, based on results from other languages. If they are single segments, then they should have similar durations as single consonants do. The second investigation examines whether the relative timing of nasalization is aerodynamically constrained by place of articulation, as per the hypervoiced allophone hypothesis.

5.1.2 Method

For the nasal allophone analysis, we examined the total duration of nasal and prenasalized consonants using the same data outlined in Section 2 above. For the hypervoicing analysis, subcomponents of the duration were examined. For this, the second author labeled three acoustic events in the production of prenasalized stops: nasal closure duration, oral closure duration, and burst duration. It is often difficult to distinguish between the nasal and oral closure duration using the acoustic signal alone. Both nasal murmurs and voicing during closure have relatively low amplitude. We distinguished between them by looking for the presence of nasal formants above 2000 Hz in the spectrogram as well as a loss in overall amplitude during oral closure in the waveform; when the vocal tract is closed, the voicing

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

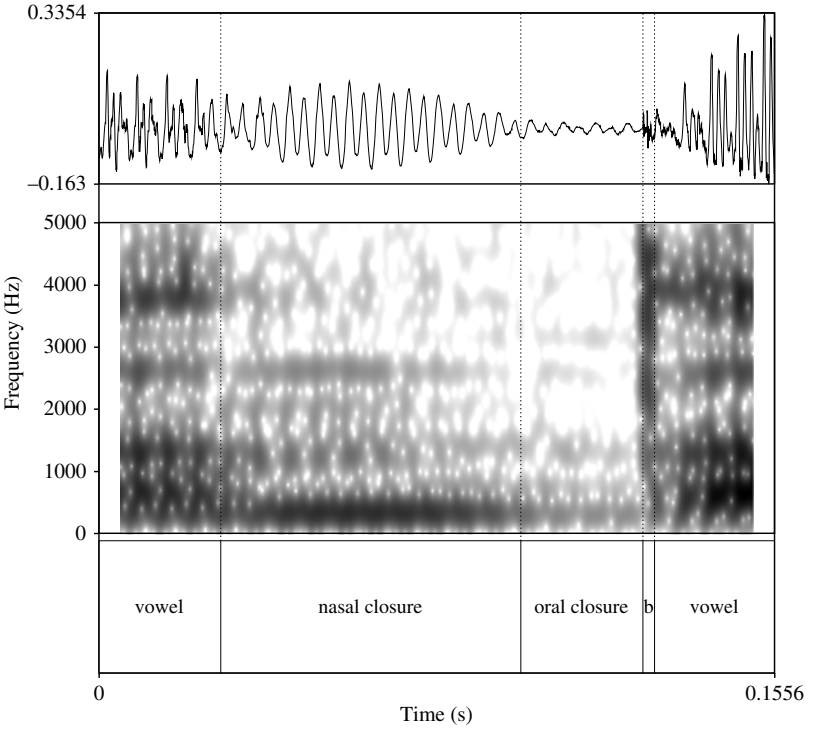


Figure 12 Labeling of prenasalized stops (b = burst).

tends to decline in amplitude, while the open velopharyngeal port of the nasals allows for a larger amplitude. Figure 12 provides an example of how prenasalized stops were labeled.

A total of six words were examined from the same data set described in Section 2. This set comprises the alveolar prenasalized stop produced in all three contexts mentioned at the top of Section 4 above, the bilabial produced in two contexts, and the velar produced in one context. A total of 240 tokens were examined. Since only the alveolar prenasalized stops surface in monosyllabic words, only disyllabic words were examined. The duration of the individual components (nasal closure, stop closure, burst) were analyzed along with whether certain components were missing (count data). For instance, not all velar prenasalized stops contained clear oral closure components, that is, velic raising coincided with the burst release and did not precede it.

Two different types of statistical models were used to analyze this data. The first models examined the influence of POA and WORD POSITION (initial, medial) on the duration of each of the different components. Within these models, speaker was treated as an error term. The second model examines the influence of POA and WORD POSITION on whether an oral closure was present in the prenasalized stop. As the dependent variable here is bimodal, a generalized linear model (instead of ANOVA) was used. SPEAKER was treated as an additional factor.

5.1.3 Results: Duration

The nasal and prenasalized stop durations are shown in Figure 13 alongside voiceless stop duration data from Section 2 (as a visual comparison). Duration was analyzed in a three factor ANOVA with POA, WORD POSITION, and CLASS (prenasalized vs. nasal) as independent variables and speaker treated as an error term. None of these effects were significant.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

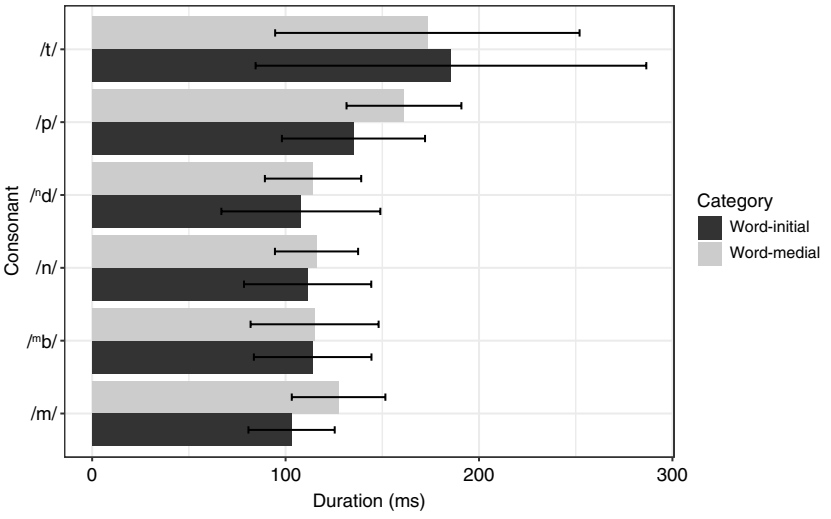


Figure 13 Duration of nasal consonants and stops.

Nasal consonants produced in word-medial position were not significantly longer than those produced in word-initial position. Prenasalized stops did not significantly differ in duration with nasal stops nor was there any influence on place of articulation (alveolar vs. bilabial) on overall duration.

5.1.4 Results: Prenasalized stop components

The duration of the different prenasalized stops and their component durations across word positions are shown in Figure 14. With respect to the duration of the nasal and oral closure duration, no significant main effects were found from the first statistical model. However, a significant main effect of POA was found for burst duration ($F(2, 12) = 6.6$, $p < .05$). The burst duration of velar prenasalized stops was longer than the burst duration for alveolars, which was longer still than that of bilabials. This finding reflects a common cross-linguistic pattern for burst duration to increase with more posterior place of articulation and closely matches the observations made in Section 3 for voiceless unaspirated stops in Yoloxóchitl Mixtec.

The findings here hide a more notable qualitative observation: in many cases there was no clear oral closure duration preceding the burst release, meaning that the nasal portion extended throughout the duration, but a burst was still observed. The absence of an oral closure was examined with respect to both WORD POSITION and POA. The main effect of WORD POSITION was significant ($G^2 = 5.8$, $p < .05$), but not the main effect of POA. The word-medial context contained more examples of words missing an oral closure than the word-initial position. Yet, this observation seems to only be true for the bilabial stops. The alveolar stops show the opposite trend. The lack of balance across position for the velar prenasalized stops may bias the outcome as well. This data is shown in Figure 15.

5.1.5 Discussion

There was no difference in the total duration of nasals and prenasalized stops in Yoloxóchitl Mixtec. In languages with true /ND/ clusters, the cluster duration is frequently longer than that of indisputably unary consonants like nasals (Cohn & Riehl 2008). The fact that prenasalized stops here are of equal duration as simple nasals and, in fact, shorter than plain stops

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

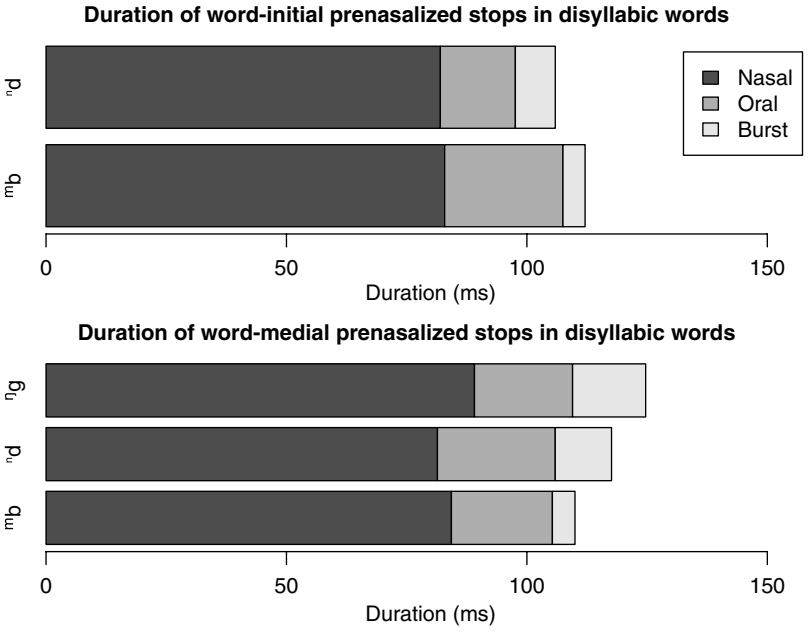


Figure 14 Duration of prenasalized stops and their component durations across word positions.

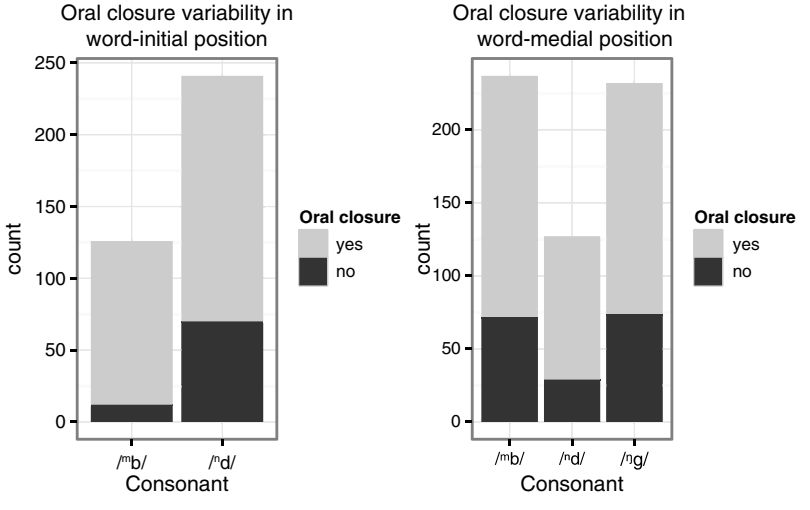


Figure 15 Number of tokens possessing a clear oral closure duration.

suggests that they are unary segments and not clusters in Yoloxóchitl Mixtec.[10] Previous work also argues that there is often no difference in timing between NC consonants analyzed as clusters and those analyzed as unary segments (Browman & Goldstein 1986, Maddieson & Ladefoged 1993, Huffman & Krakow 1993). Moreover, NC sequences that are analyzed as unary segments may in fact be longer than unary segments in certain Bantu languages (Hubbard 1995). As a result, the quantitative data on total duration here may only be suggestive with respect to the phonological status of prenasalized stops. Rather, it seems that the

[10] However, this conclusion also carries some caveats, discussed in Downing (2005).

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

nature of the alternations shown in Table 5 above is more convincing evidence in favor of the NASAL ALLOPHONE hypothesis.

Nevertheless, the data from the quantitative and qualitative acoustic analyses do not support the idea that the prenasalized stops in Yoloxóchitl Mixtec are hypervoiced variants of voiced stops. In the HYPERVOICING perspective, one predicts longer oral closure duration for more anterior stops and shorter oral closure duration for more posterior stops. No differences in oral closure duration were found. Moreover, voicing is typically more difficult to maintain in utterance-initial position than in intervocalic position because the transglottal air pressure differential for voicing initiation is greater than for its continuation (Weismer 1980, Westbury & Keating 1986). This leads to the prediction that venting through the velopharyngeal port in order to maintain voicing would be more likely in word-initial contexts than in word-medial ones. Yet, the data here show a slight tendency for the opposite: velar raising (and closure) is more often coincidental with release in word-medial position than in word-initial position. There is a stronger case for hypervoicing word-medially than word-initially, which is the opposite prediction than one expects from the hypervoiced allophone hypothesis. Note that voicing is not contrastive elsewhere in the phonological system.

What the data do show is that velic raising may either slightly precede or be coincidental with stop burst release. This nasal–oral timing strategy is consistent across speakers and this suggests that the careful control of where nasalization occurs on Yoloxóchitl Mixtec words is the main factor speakers control in the production of prenasalized stops. An analysis where prenasalized stops are considered allophones of nasal consonants, conditioned by the following vowel's nasality, is more congruent with this view. This finding mirrors work on the production of prenasalized stops in other languages (Maddieson & Ladefoged 1993, Huffman & Krakow 1993, Beddor & Onsuwan 2003, Cohn & Riehl 2008). Examining nasal airflow data from several Austronesian languages, Cohn and Riehl found that prenasalized stops, /ND/ sequences, and post-ploded nasals tend to have quite short oral closure duration relative to a long nasal duration. In each of these languages, there is a process of perseverative nasalization on vowels following nasal consonants. Yet, vowels are obligatorily oral following the release of prenasalized stops in each of the languages they examined. Yoloxóchitl Mixtec exhibits the same surface phonetic pattern here.

If prenasalized stops (or, rather, post-stopped nasals) are allophones of nasal consonants before oral vowels, then there is necessarily a contrast in vowel nasalization after nasal consonants. If this is the case, then vowel nasalization is not merely phonetically-redundant after nasal consonants, but contrastive. The contrast is acoustically and perceptually maintained here by the differences between preceding onsets: abrupt velic raising always occurs prior to stop release before oral vowels.

5.2 Glides

Like most Mixtec languages, Yoloxóchitl Mixtec possesses a bilabial glide /β/ and a palatal glide /j/. There is some variability in the production of both glides across Mixtec languages. For instance, certain variants, such as San Juan Mixtepec Mixtec, show variability between bilabial and labiodental productions [β ~ v] of the labial glide (Pike & Ibach 1978). In many Mixtec variants, the palatal glide has become a voiced palatal fricative /ʒ/, as in Huajuapan Mixtec (Pike & Cowan 1967), Silacayoapan Mixtec (North & Shields 1977), and Diuxi Mixtec (Pike & Oram 1976). In at least a few variants, there is variability in production, e.g. [j ~ ʒ], as in Acatlán Mixtec (Pike & Wistrand 1974), San Juan Mixtepec Mixtec (Pike & Ibach 1978), and Jicaltepec Mixtec (Bradley 1970). In Coatzospan Mixtec, cognate words with /y/ have merged entirely with /ʃ/.

The palatal glide in Yoloxóchitl Mixtec is always produced without frication, though there are two variants of the bilabial glide: [b] and [β]; the latter seems impressionistically more common for two of the eight speakers. In addition to the formant values and trajectories for each glide, the frequency of occurrence for each of these bilabial glide variants was

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

examined here. Target formant trajectories for each glide were examined across preceding and following vowels in two different word positions.

5.2.1 Method

Segmentation of glides in the acoustic signal is particularly problematic. They are defined by a target constriction degree (see Maddieson 2008), yet movement towards this constriction frequently involves relatively slow formant trajectories. The process of segmentation therefore often requires either placing a vowel-glide boundary before the onglide or offglide formant trajectory, or at the midpoint of this trajectory. There are advantages and disadvantages to both approaches with respect to acoustic segmentation. The former attempts to capture most of the dynamics of the articulation involved in glide production, but inevitably includes a portion of the adjacent vowel.[11] The latter attempts to evenly distribute the formant transitions to both the vowel and the glide, though estimating a trajectory midpoint is more prone to segmenter error.

One way around both these approaches is to examine formant trajectories in the larger context where the glide is produced. This is the method used here. Glides were selected from the data described in Section 2. We then estimated the onset of the vowel preceding the glide and the offset of the vowel following the glide and extracted formant trajectories from across this duration. Formant values were extracted using a script written by the first author for Praat (Boersma & Weenink 2016) and collected at 20 evenly distributed time points across the VCV interval. As the carrier sentences always included a vowel /a/, both on the preceding and following vowel, vowel quality was controlled. A total of six target words (three per glide) were examined, each repeated five times by eight speakers, for a total of 240 tokens. This data was statistically analyzed with an ANOVA for each of the first three formants. CONSONANT TYPE (/j/, /β/) and TIME (1–20) were treated as independent variables and the formant values as the dependent variable. SPEAKER was treated as an error term.

In the second analysis, each of the bilabial glides was qualitatively examined for degree of constriction. While the consonant varied in its realization between a bilabial glide, a bilabial fricative, and a bilabial stop, only the continuant vs. non-continuant realizations were catalogued. This categorical decision was treated as a dependent variable in a logistic regression model with WORD SIZE (monosyllable, disyllable), POSITION (initial, medial), and SPEAKER as independent variables. This model specifically tests whether the realization of bilabial glides is context or speaker-dependent.

5.2.2 Results

The formant trajectories for both the bilabial and palatal glides are shown in Figure 16. Dotted lines here reflect the estimated location of formant movement pertaining to the glide constriction. In all statistical models, we are more concerned with the change in formant values in time than with an average formant value across the time range. In the statistical model examining F1 trajectory, a main effect of CONSONANT TYPE was found ($F(1, 6) = 20.6$, $p < .01$). A significant interaction between CONSONANT TYPE and TIME was also observed here ($F(1, 6) = 6.0$, $p < .05$). Greater F1 lowering was observed in the context of the palatal glide than in the bilabial glide. Observe also that F1 increases prior to the glide, but decreases with glide constriction. No such pattern was observed for the bilabial.

In the statistical model examining F2 trajectory, a significant main effect of CONSONANT TYPE was found ($F(1, 6) = 51.2$, $p < .001$). A significant interaction between CONSONANT

[11] From an articulatory perspective, the vowel and the glide are co-produced as several gestures coordinated in phase. Thus, while segmentation is frequently necessary for acoustic analysis, it is also somewhat artificial from the perspective of speech production.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

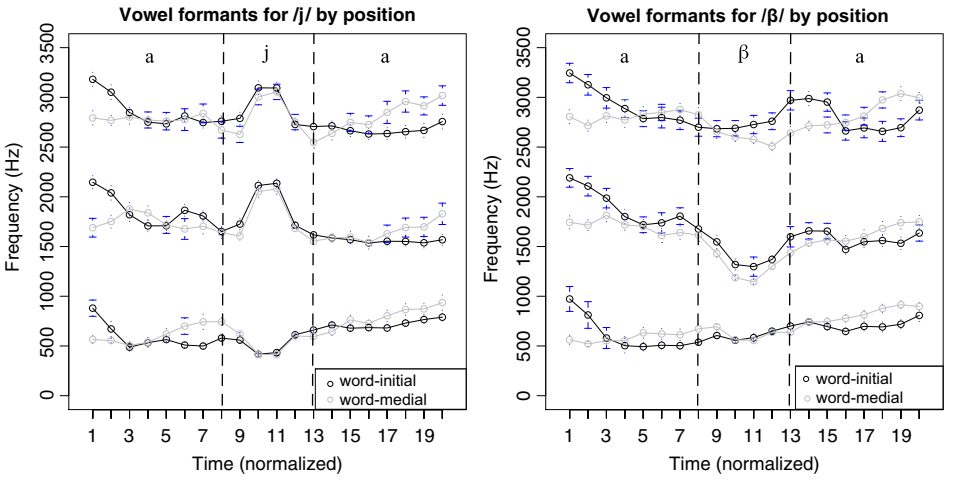


Figure 16 Glide formant transitions.

TYPE and TIME was also observed ($F(1, 6) = 10.4, p < .05$). F2 lowered during the production of the bilabial glide and raised during the production of the palatal glide. Finally, in the statistical model examining F3 trajectory, no main effects were observed. However, there was a significant interaction between CONSONANT TYPE and TIME ($F(1, 6) = 21.3, p < .01$). In the production of the palatal glide, F3 raises, but it appears to lower and subsequently raise upon release in the bilabial glide.

Figure 17 shows qualitative manner differences in the production of the bilabial glide. While stop variants of the glide were more frequent in word-initial contexts than in word-medial contexts, no significant differences were found. The bilabial glide was produced as either a glide or a fricative in a majority of the cases.

5.2.3 Discussion

In general, the formant data show bilabial and palatal glides to be nearly maximally dispersed in the acoustic space. Bilabial glides are realized with lowered F2 and F3 targets, while palatal

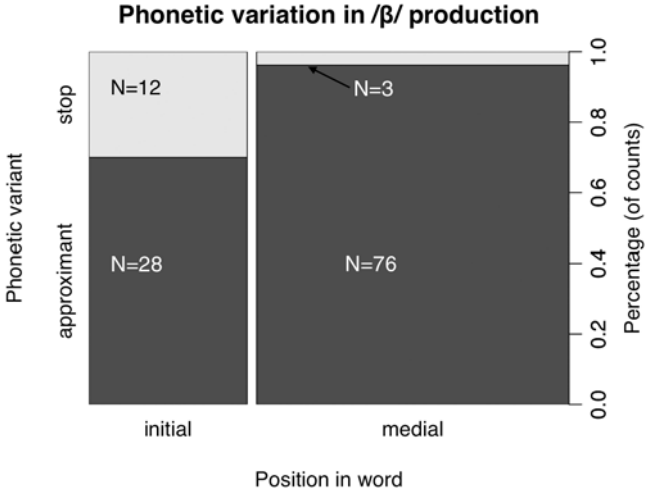


Figure 17 Qualitative manner differences in the production of the bilabial glide.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

glides are realized with raised F2 and F3 targets. The first formant target for the bilabial glide is relatively flat, while it is lowered for the palatal glide. The data also show some surprising patterns. For instance, there is a pattern of F1 raising prior to the production of the palatal glide, but only in the word-medial context. Conversely, there is a pattern of F2 and F3 raising following the production of the bilabial glide. In both cases, these patterns appear to be examples of coarticulatory dissimilation (Xu & Wang 2001, Bye 2011).

The qualitative data on bilabial glides shows that, while there is some variability in their production, speakers produce bilabial glides primarily as frictionless continuants. In fact, for some speakers, there was little notable constriction, so that the glide was barely detectable in the acoustic signal. The stop variants of glides were produced almost entirely in word-initial position. This is notable, as this is the phonological context where bilabial continuants are produced as stops in Spanish (Hualde 2005). The phonetic variation observed here may reflect some transfer of native Yoloxóchitl Mixtec speakers' second language phonology into their first language. As both languages contain a bilabial continuant, produced similarly across languages, such a transfer may be more likely here than for other phonological contrasts.

6 General discussion

6.1 The universal and the language-specific

Phonetic description involves a combination of close attention to the more universal tendencies in speech production and the language-specific patterns which sometimes hide such generalities. Several aspects of Yoloxóchitl Mixtec consonantal phonetics support different universal phonetic tendencies. For instance, the consonants are longer in monosyllabic words than in disyllabic words in Yoloxóchitl Mixtec. This pattern, often called POLYSYLLABIC SHORTENING, is ostensibly a language-universal tendency (Jones 1942–1943, Lehiste 1970, Lindblom et al. 1981, Turk & Shattuck-Hufnagel 2000, White & Turk 2010). At first glance, the Yoloxóchitl Mixtec data would seem subject to the same phonetic constraint, but the degree to which consonantal shortening occurs is greater than predicted from previous research. Yet, once one considers the greater influence of rightmost prominence on polysyllabic shortening, the Yoloxóchitl Mixtec data more closely matches observations from work on English (White & Turk 2010). Polysyllabic shortening is stronger in stress-final words in English, and similarly stronger in a stress-final language like Yoloxóchitl Mixtec when compared to findings from other languages.

Furthermore, Yoloxóchitl Mixtec shows a strong language-general effect of place of articulation on VOT, even though only short-lag VOTs occur (see Cho & Ladefoged 1999). The investigation of individual differences here shows that dental stops tend to vary the most in terms of VOT duration. This observation may explain why, cross-linguistically, bilabial and coronal stops tend to have similar VOT duration in comparison with velar stops: coronal stops are simply more variable. For certain speakers, they have VOT durations similar to bilabial stops, while for others, they have longer VOT durations.

Another near universal pattern is the inverse relationship between consonant closure duration and VOT in voiceless stops (Weismer 1980, Maddieson 1997, Cho & Ladefoged 1999, Gordon et al. 2001). The Yoloxóchitl Mixtec data show evidence of such a pattern, but the shorter closure duration for the velar stops (/k k^w/) compared to the other places of articulation cannot be explained by this principle alone. Instead, it seems that there is a language-specific rule whereby velar consonants are shortened and potentially lenited. The presence of a similar process of velar lenition in many other Mixtec varieties, especially others spoken in Guerrero, suggests that this pattern might also be more of an areal feature of Mixtec languages. Both this and the previous pattern demonstrate how universal phonetic tendencies may be exaggerated by language or family-specific patterns.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

6.2 The phonology of complex segments

In Yoloxóchitl Mixtec, the prenasalized stops are consistently realized with an oral release and a following oral vowel. The relative timing of nasal and oral closure does not vary with place of articulation, contra predictions from the hypervoiced allophone hypothesis (Iverson & Salmons 1996). Furthermore, the production of prenasalized stops frequently involved the absence of any observable oral closure duration (in 20–30% of cases). These observations are more in line with the view that prenasalized stops (or post-occluded nasals) are allophones of nasal consonants in Yoloxóchitl Mixtec than hypervoiced variants of voiced stops.

This finding has ramifications for the phonology of Yoloxóchitl Mixtec. In particular, one must posit that nasal vowels are contrastive after nasal consonants and that an [ORAL] feature on vowels conditions consonantal allophony. This argument fits well with the comparative data from Austronesian (Cohn & Riehl 2008), where nasal and oral vowels have the same distribution as they do in Yoloxóchitl Mixtec. However, it also suggests that a phonological feature like [NASAL] is not privative (as argued by Marlett 1992) in Mixtec because the absence of nasality triggers allophony, as we observed in the data in Table 5.[12]

7 Conclusions

In addition to its descriptive value, the phonetic data examined here have a bearing on several topics addressed in the literature. They include the relationship between phonetic universals and language-specific structural constraints, factors determining variation in VOT, and the phonological status of complex consonant types (prenasalized stops). These findings also pave the way for testing future hypotheses on the phonetic and phonological structure either in the target language or cross-linguistically. For instance, to what extent do the durational differences in VOT here correspond to patterns of consonant undershoot in spontaneous speech? Many voiceless consonants become at least partially voiced in running speech. Castillo García (2007) observes patterns of velar lenition and recent investigations into variation in spontaneous speech have revealed this variation to be related to prosodic structure, i.e. greater lenition in unstressed syllables (see DiCanio et al. 2017).

Furthermore, the acoustic observations regarding nasal and oral stop closure in the prenasalized stops would be more easily interpreted when accompanied by aerodynamic data. While the acoustic signal is suggestive, it does not fully determine the timing of denasalization. Such a future investigation may more conclusively shed light on the nature of prenasalized stops in the language.

Finally, the present paper may also serve as a model for intermediate-length phonetic descriptions of other languages' consonantal systems. There will always be sufficient data of interest for a monograph on any language's phonetic structure (e.g. Olive, Greenwood & Coleman 1993, McDonough 2003), but shorter sketches are of value as well. We propose that their value is enhanced when issues of universality of phonetic effects are directly addressed and when possible insights into phonological questions can be generated.

Acknowledgements

This work was supported by NSF Grant No. 0966411 to Haskins Laboratories and NSF DEL/RI Grant No. 1603323 to the University at Buffalo. The authors wish to acknowledge the generous help and support of Jonathan Amith in this research as well as the commentary provided by two anonymous reviewers and the associate editor.

[12] See Steriade (1993) for a review of the arguments on nasal privativity.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Appendix A. Information on scripts

Three scripts were written for Praat (Boersma & Weenink 2016) to extract acoustic data from the speech signal. The script used to extract duration from the speech signal was *Get_duration_2.0.praat*. The script used to extract formant values for C–V transitions and glides was *Get_Formants_nonnormalized.praat*. The script used to do fricative time-normalization and extract fricative spectral moments was *Time_averaging_for_fricatives_2.0.praat*. Each of these scripts is publicly available at http://www.acsu.buffalo.edu/~cdicanio/scripts.html.

Appendix B. Additional table

Table B1 Normalized (*z*-score) values for acoustic cues for fricatives.

Cue	[s̪]	[ʃ]
Center of gravity	0.857	−0.683
Standard deviation	0.780	−0.617
Kurtosis	−0.430	0.443
Skewness	−0.795	0.671
Duration	−0.364	0.037
Intensity	−0.080	0.047

References

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

Adams, James N. 2007. *The regional diversification of Latin 200 BC – AD 600*. Cambridge: Cambridge University Press.

Amith, Jonathan D. & Rey Castillo García. No date. Corpus and lexicon development: Endangered genres of discourse and domains of cultural knowledge in Tu'un ísaví (Mixtec) of Yoloxóchitl, Guerrero. Jonathan D. Amith Yoloxóchitl Mixtec collection. Archive of the Indigenous Languages of Latin America: www.ailla.utexas.org. Media: audio, text. Access public.

Babel, Molly. 2009. *Phonetic and social selectivity in speech accommodation*. Ph.D. dissertation, University of California, Berkeley.

Beddor, Patricia S. & Chutamanee Onsuwan. 2003. Perception of prenasalized stops. In Maria Josep Solé, Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (ICPhS XV), 407–410. Barcelona: Causal Productions.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 2016. Praat: Doing phonetics by computer [computer program]. Version 6.0.21. http://www.praat.org.

Bourhis, Richard Y. & Howard Giles. 1977. The language of intergroup distinctiveness. In Howard Giles (ed.), *Language, ethnicity, and intergroup relations*, 11–136. London: Academic Press.

Bradley, C. Henry. 1970. *A linguistic sketch of Jicaltepec Mixtec* (Summer Institute of Linguistics Publications in Linguistics and Related Fields 25). Norman, OK: Summer Institute of Linguistics, University of Oklahoma.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3, 219–252.

Bye, Patrick. 2011. Dissimilation. In Marc van Oostendorp, Colin Ewen & Elizabeth Hume (eds.), *The Blackwell companion to phonology*, 1408–1433. Oxford: Wiley-Blackwell.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Castillo García, Rey. 2007. Descripción fonológica, segmental, y tonal del Mixteco de Yoloxóchitl, Guerrero [Phonological, segmental, and tonal description of the Mixtec language of Yoloxóchito, Guerrero]. Master's thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS), México, D.F.

Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27, 207–229.

Cohn, Abigail C. & Anastasia K. Riehl. 2008. The internal structure of nasal–stop sequences: Evidence from Austronesian. Presented at the 11th Conference on Laboratory Phonology, Wellington, NZ.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54, 35–50.

de Alvarado, Fray Francisco. 1593. *Vocabulario en Lengua Mixteca Hecho por los Padres de la Orden de Predicadores* [Vocabulary of the Mixtec language made by the Order of Preachers (Dominican Order)] (reprint, 1962). México, D.F.: Instituto Nacional Indígena e Instituto Nacional de Antropología e Historia.

de Los Reyes, Fray Antonio. 1593. *Arte en Lengua Mixteca* [The art of the Mixtec language]. Casa de Pedro Balli, Mexico: Comte H. de Charencey edition.

Delattre, Pierre. 1968. From acoustic cues to distinctive features. *Phonetica* 18, 198–230.

DiCanio, Christian T. 2010. Illustrations of the IPA: San Martín Itunyoso Trique. *Journal of the International Phonetic Association* 40(2), 227–238.

DiCanio, Christian T. 2012. The phonetics of fortis and lenis consonants in Itunyoso Trique. *International Journal of American Linguistics* 78(2), 239–272.

DiCanio, Christian T., Jonathan Amith & Rey Castillo García. 2012. Phonetic alignment in Yoloxóchitl Mixtec tone. Presented at the Society for the Study of the Indigenous Languages of the Americas Annual Meeting, Portland, OR.

DiCanio, Christian [T.], Jonathan Amith & Rey Castillo García. 2014. The phonetics of moraic alignment in Yoloxóchitl Mixtec. *Proceedings of the 4th Tonal Aspects of Language Symposium*, Nijmegen, the Netherlands, 203–210.

DiCanio, Christian [T.], Joshua Benn & Castillo García, R. 2018. The phonetics of information structure in Yoloxóchitl Mixtec. *Journal of Phonetics* 68, 50–68.

DiCanio, Christian [T.], Wei-Rong Chen, Joshua Benn, Jonathan D. Amith & Rey Castillo García. 2017. Automatic detection of extreme stop allophony in Mixtec spontaneous speech. Presented at the 5th Annual Meeting in Phonology, New York University.

DiCanio, Christian [T.], Hosung Nam, Jonathan D. Amith, Rey Castillo García & Douglas H. Whalen. 2015. Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics* 48, 45–59.

Downing, Laura J. 2005. On the ambiguous segmental status of nasals in homorganic NC sequences. In Marc van Oostendorp & Jeroen M. van de Weijer (eds.), *The internal organization of phonological segments*, 219–252. Berlin: Mouton de Gruyter.

Drager, Katie, Jennifer Hay & Abby Walker. 2010. Pronounced rivalries: Attitudes and speech production. *Te Reo* 53, 27–53.

Gay, Thomas. 1981. Mechanisms of control in speech rate. *Phonetica* 38, 148–158.

Gerfen, Chip. 2001. Nasalized fricatives in Coatzospan Mixtec. *International Journal of American Linguistics* 67(4), 449–466.

Gerfen, Chip & Kirk Baker. 2005. The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics* 33, 311–334.

Gerfen, Henry J. 1996. *Topics in the phonology and phonetics of Coatzospan Mixtec*. Ph.D. dissertation, University of Arizona.

Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32(2), 141–174.

Gordon, Matthew, Pamela Munro & Peter Ladefoged. 2000. Some phonetic structures of Chickasaw. *Anthropological Linguistics* 42, 366–400.

Gordon, Matthew, Brian Potter, John Dawson, Willem de Reuse & Peter Ladefoged. 2001. Phonetic structures of Western Apache. *International Journal of American Linguistics* 67(4), 415–448.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Herrera Zendejas, Esther. 2009. *Formas sonoras: mapa fónico de las lenguas mexicanas* [Voiced forms: A phonic map of Mexican languages] (Estudios de Linguística 6). México D.F.: Colegio de México.
Hinton, Leanne. 1991. An accentual analysis of tone in Chalcatongo Mixtec. In James E. Redden (ed.), *Papers from the American Indian Languages Conferences held at the University of California, Santa Cruz* (Occasional Papers on Linguistics 16), 173–182. Carbondale, IL: Southern Illinois University.
Hualde, José Ignacio. 2005. *The sounds of Spanish*. Cambridge: Cambridge University Press.
Hubbard, Kathleen. 1995. Toward a theory of phonological and phonetic timing: Evidence from Bantu. In Bruce Connell & Amalia Arvaniti (eds.), *Papers in Laboratory Phonology IV: Phonology and phonetic evidence*, 168–187. Cambridge: Cambridge University Press.
Huffman, Marie K. & Rena A. Krakow (eds.). 1993. *Phonetics and phonology: Nasals, nasalization, and the velum*, vol. 5. San Diego, CA: Academic Press.
Hunter, Georgia G. & Eunice V. Pike. 1969. The phonology and tone sandhi of Molinos Mixtec. *Journal of Linguistics* 47, 24–40.
Iverson, Gregory K. & Joseph C. Salmons. 1996. Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics* 62(2), 165–175.
Jassem, Wiktor. 1968. Acoustic description of voiceless fricatives in terms of spectral parameters. In Wiktor Jassem (ed.), *Speech analysis and synthesis*, 189–206. Warsaw: Państwowe Wydawnictwo Naukowe.
Johnson, Keith. 2012. *Acoustic & auditory phonetics*, 3rd edn. Hoboken, NJ: Wiley-Blackwell.
Jones, Daniel. 1942–1943. Chronemes and tonemes. *Acta Linguistica* 3, 1–10.
Josserand, Judy K. 1983. *Mixtec dialect history*. Ph.D. dissertation, Tulane University.
Keating, Patricia A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286–319.
Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3), 419–454.
Ladefoged, Peter & Ian Maddieson. 1996. *Sounds of the world's languages*. Oxford: Blackwell.
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
Lewis, M. Paul, Gary F. Simons & Charles D. Fennig. 2013. *Ethnologue: Languages of the world*, 17th edn. Dallas, TX: SIL International. http://www.ethnologue.com.
Lindblom, Björn, Bertil Lyberg & Karin Holmgren. 1981. Durational patterns of Swedish phonology: Do they reflect short-term motor memory processes? Bloomington, IN: Indiana University Linguistics Club.
Lindblom, Björn & Karin Rapp. 1973. Some temporal regularities of spoken Swedish. *Papers from the Instititue of Linguistics at the University of Stockholm* 21, 1–59.
Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 384–422.
Longacre, Robert E. 1957. Proto-Mixtecan. *International Journal of American Linguistics* 23(4), 1–195.
Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec* (University of California Publications in Linguistics 127). Berkeley, CA: University of California Press.
Macaulay, Monica & Joseph C. Salmons. 1995. The phonology of glottalization in Mixtec. *International Journal of American Linguistics* 61(1), 38–61.
Macken, Marlys A. & Joseph C. Salmons. 1997. Prosodic templates in sound change. *Diachronica* 14(1), 31–66.
Maddieson, Ian. 1997. Phonetic universals. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 619–639. Hoboken, NJ: Wiley-Blackwell.
Maddieson, Ian. 2008. Glides and gemination. *Lingua* 118, 1926–1936.
Maddieson, Ian, Heriberto Avelino & Loretta O'Connor. 2009. The phonetic structures of Oaxaca Chontal. *International Journal of American Linguistics* 75(1), 69–103.
Maddieson, Ian & Peter Ladefoged. 1993. Phonetics of partially nasal consonants. In Huffman & Krakow (eds.), 251–301.
Maniwa, Kazumi, Allard Jongman & Travis Wade. 2009. Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America* 125(6), 3962–3973.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Marlett, Stephen A. 1992. Nasalization in Mixtec languages. *International Journal of American Linguistics* 58(4), 425–435.
McDonough, Joyce M. 2003. *The Navajo sound system*. Dordrecht: Kluwer.
McGowan, Richard S. & Susan Nittrouer. 1988. Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/. *The Journal of the Acoustical Society of America* 83(1), 229–236.
Nartey, Jonas. 1982. *On fricative phones and phonemes*. Ph.D. dissertation, UCLA.
North, Joanne & Jäna Shields. 1977. Silacayoapan Mixtec phonology. In William. R. Merrifield (ed.), *Studies in Otomanguean phonology*, 21–33. Arlington, TX: Summer Institute of Linguistics, University of Texas at Arlington.
Ohala, John J. 2001. The phonetics of sound change. In Charles W. Kreidler (ed.), *Phonology: Critical concepts in linguistics*, vol. 4, 44–81. London & New York: Routledge.
Ohala, John J. & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In Huffman & Krakow (eds.), 225–249.
Olive, Joseph P., Alice Greenwood & John Coleman. 1993. *Acoustics of American English speech: A dynamic approach*. New York: Springer.
Olson, Daniel J. 2013. Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41, 407–420.
Palancar, Enrique L., Jonathan D. Amith & Rey Castillo García. 2016. Verbal inflection in Yoloxóchitl Mixtec. In Enrique L. Palancar & Jean-Léo Léonard (eds.), *Tone and inflection: New facts and new perspectives*, 295–336. Berlin: Mouton de Gruyter.
Pankratz, Leo & Eunice V. Pike. 1967. Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics* 33(4), 287–299.
Pike, Eunice V. & John H. Cowan. 1967. Mixtec phonology and morphophonemics. *Anthropological Linguistics* 9(5), 1–15.
Pike, Eunice V. & Thomas Ibach. 1978. The phonology of the Mixtepec dialect of Mixtec. In Mohammad Ali Jazayery, Edgar C. Polomé & Werner Winter (eds.), *Linguistic and literary studies in honor of Archibald A. Hill*, vol. 2: *Descriptive linguistics*, 271–285. The Hague: Mouton.
Pike, Eunice V. & Joy Oram. 1976. Stress and tone in the phonology of Diuxi Mixtec. *Phonetica* 33, 321–333.
Pike, Eunice V. & Kent Wistrand. 1974. Step-up terrace tone in Acatlán Mixtec. In Ruth M. Brend (ed.), *Advances in tagmemics*, 81–104. Amsterdam: North-Holland.
Pike, Kenneth L. 1944. Analysis of a Mixteco text. *International Journal of American Linguistics* 10(4), 113–138.
R Development Core Team. 2016. R: A language and environment for statistical computing [computer program]. Vienna: R Foundation for Statistical Computing. http://www.R-project.org.
Rensch, Calvin R. 1976. *Comparative Otomanguean phonology* (Language Science Monograph 14). Bloomington, IN: Indiana University.
Shadle, Christine H. 2012. On the acoustics and aerodynamics of fricatives. In Abigail C. Cohn, Cécile Fougeron, Marie K. Huffman & Margaret E. L. Renwick (eds.), *The Oxford handbook of laboratory phonology*, 511–526. Oxford: Oxford University Press.
Silverman, Daniel, Barbara Blankenship, Paul Kirk & Peter Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* 37(1), 70–88.
Soli, Sigfrid D. 1981. Second formants in fricatives: Acoustic consequences of fricative–vowel coarticulation. *The Journal of the Acoustical Society of America* 70, 976–984.
Steriade, Donca. 1993. Orality and markedness. *Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society* (BLS 19): *General Session and Parasession on Semantic Typology and Semantic Universals*, 334–347.
Stevens, Kenneth N. 2000. *Acoustic phonetics*. Cambridge, MA: MIT Press.
Stevens, Mary & John Hajek. 2004. A preliminary investigation of some acoustic characteristics of ejectives in Waima'a: VOT and closure duration *Proceedings of the 10th Australian International Conference on Speech Science & Technology*, Macquarie University, Syndey, 277–282.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press

Suh, Yunju. 2009. *Phonological and phonetic asymmetries of Cw combinations*. Ph.D. dissertation, Stony Brook University.
Tarnóczy, Tamás. 1965. Can the problem of automatic speech recognition be solved by analysis alone? *Rapports de 5e Congrés International d'Acoustique*, vol. II, 371–387. Liége: D. E. Commins.
Turk, Alice & Stefanie Shattuck-Hufnagel. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28, 397–440.
Umeda, Noriko. 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America* 61(3), 846–858.
Weismer, Gary. 1980. Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics* 8, 427–438.
Westbury, John R. & Patricia A. Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145–166.
Whalen, Douglas H. 1981. Effects of vocalic formant transitions and vowel quality on the English [s]–[š] boundary. *The Journal of the Acoustical Society of America* 69, 275–282.
Whalen, Douglas H. 1983. Vowel information in postvocalic fricative noises. *Language and Speech* 26, 91–100.
White, Laurence & Alice Turk. 2010. English words on the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics*, 38, 459–471.
Xu, Yi & Q. Emily Wang. 2001. Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication* 33, 319–337.
Zylstra, Carol F. 1980. Phonology and morphophonemics of the Mixtec of Alacatlatzala, Guerrero. *SIL–Mexico Workpapers* 4, 15–42.

https://doi.org/10.1017/S0025100318000294 Published online by Cambridge University Press