

Acoustic properties of vowels in Upper Necaxa Totonac

Michelle García-Vega

University of Alberta
mgarciav@ualberta.ca

Benjamin V. Tucker

University of Alberta
bvtucker@ualberta.ca

Upper Necaxa Totonac is a Totonacan language spoken in the Necaxa River valley in the Sierra Norte of Puebla State, Mexico. While the Totonacan languages historically have three phonemic vowel qualities, the Upper Necaxa system consists of five vowels that contrast length and laryngealization. With acoustic data from six native speakers from the Totonacan communities of Patla and Chicontla, we explore the phonetic properties of vowels with respect to the first and second formant frequencies, quantity (duration), vowel phonation (modal vs. laryngeal), and stress. The data indicate that long, short, modal and laryngeal vowels occupy a similar formant space and that duration is the primary phonetic correlate of phonemic vowel length. A shift in vowel quality and an increase in duration and pitch were shown to be the acoustic characteristics of stress. The study provides a first acoustic analysis of vowels in Upper Necaxa, and contributes to typological descriptions of the properties of vowels connected with quality, quantity, stress, and phonation.

1 Introduction

The Totonacan language family (formerly Totonac-Tepehua) is one of the largest language families in Mesoamerica spoken by roughly 240,000 people (INEGI 2010). While the exact number of Totonacan languages is unknown, to date many Totonacan language varieties remain un-described and unanalyzed. The present study provides a description of the acoustic properties of vowels in Upper Necaxa Totonac (UNT), a member of the Northern branch of the Totonacan language family. We investigate the five-vowel system with contrastive length and laryngealization in relation to F1 and F2, duration, and fundamental frequency (f_0) in UNT. We further test the claims Beck (2004) makes on the phonetic correlates of stress on vowels based on his auditory perception of the sound system, namely that (i) stressed vowels have increased duration, (ii) stressed short vowels have the same duration as long vowels, and (iii) stressed long vowels are not consistently longer than unstressed long vowels. The results demonstrate that claims (i) and (iii) are supported by acoustic measurements, and that slight differences in vowel quality in addition to increased duration is a significant predictor of stress in UNT. This difference in vowel quality and lengthening in prosodic contexts is only found with short stressed vowels but not stressed long vowels, an observation that has been termed target undershoot in Lindblom (1963) as a result of time constraints on vowel transitions. Complex interactions between syllable position and vowel length were also investigated and indicates that syllable-initial and syllable-final positions were predictors of increased vowel duration though positional effects became weaker with every repetition.

Acoustic analysis of the Upper Necaxa Totonac vowel system adds to typological descriptions that have demonstrated the phonetic correlates of vowel quality, quantity, and stress vary among languages. Where some acoustic studies have described vowel length as a property marked primarily by a quantity difference (Kari 1990, Holton 2000), other studies connect vowel length to distinctions in quantity and quality (Henderson 1967, Muehlbauer 2012), or as affected by other cues like formant transitions and syllable positions (Lehnert-LeHouillier 2010). These distinctions in formant structure might apply to only a subset of vowels correlated with vowel length or height (Tuttle, Lovick & Núñez-Ortiz 2011). Gordon (1998) and Gordon & Ladefoged (2001) suggest that duration and formant patterns of vowels are often affected by non-modal phonation, and that non-modal vowels are generally less salient characterized by reduced acoustic intensity and loudness. In Kirk, Ladefoged & Ladefoged (1993), non-modal vowels were described as phonetically much longer than their modal counterpart. In terms of stress, studies since work by Fry (1955, 1958) have shown that while the most effective property for cueing stress is duration, other factors such as intensity and fundamental frequency play a role. Studies have also found that lengthening effects are correlated with certain positions and prosodic domains used to mark phonological junctures or phrasal boundaries (Wightman et al. 1992, Tuttle 1998, 2005). The present study contributes to these investigations on the phonetic correlates of vowel length and stress, and analyzes the role of laryngealization on vowel quality and quantity and its relation to syllable position.

The literature on the phonetics and phonology of Totonacan languages is relatively sparse, the majority focusing on phonological descriptions based on researchers' perception of the sounds. Recent reconstruction efforts in Brown et al. (2011) were based on the phonological inventory of 18 Totonac languages, of which only seven have published grammars or dictionaries compiled by academically-trained linguists. Several phonological accounts of the sound systems in Totonacan are found in Román Lobato (2008), McFarland (2009), and MacKay & Trechsel (2013) to name a few, where some descriptions have gone as far as suggesting perceptual correlates of stress and length, which always ignore potentially complex interactions concerning laryngealization. In Zapotitlán Totonac, long vowels are described as produced at a lower tongue position than corresponding short vowels, and short vowels as more flexible in quality varying freely with a more central position (Aschmann 1946). MacKay & Trechsel (2013) describe stress as affecting both the loudness and duration of vowels in Pisaflores Tepehua. However, acoustic analysis of the phonemic inventory of the Totonacan languages has not appeared in the literature until recently. Phonetic studies on vowel laryngealization (Misantla: Trechsel & Faber 1992 (which remains unpublished); Papantla: Alarcón 2008, Herrera 2009), the system of ejective fricatives (Upper Necaxa: Beck 2006, Puderbaugh & Tucker 2013, Puderbaugh 2015), and the acoustics of obstruents (Huehuetla: Puderbaugh 2016) are now beginning to appear. These studies have contributed important information on the phonetic realization of the sounds in the language family. In a phonetic study of vowel modality in Papantla Totonac, Alarcón Montero (2008) demonstrates that vowel neutralization of laryngealization results in homophonous forms, where minimal pairs such as /tsʲtsi/ 'warm' and /tsʲtsi/ 'grain' are neutralized with [tsʲtsi]. The study further demonstrates that intervocalic consonant voicing, /stapu/ → [stabu] 'chaquiste', may help distinguish potential homophony from other neutralizing items, as in /stapu/ → [stapu] 'bean' (Alarcón Montero 2008). In Upper Necaxa Totonac, the acoustic properties of duration, center of gravity, and voice onset time were measured in the rare set of ejective fricatives in a study presented by Puderbaugh & Tucker (2013). Up until now, however, descriptions of the Upper Necaxa Totonac vowel inventory have been based on auditory and perceptual observation by the field researcher, Beck (2004, 2011), who describes a relatively large system of 20 vowel phonemes.

The Totonacan languages are currently spoken in east-central Mexico in the states of Puebla, Hidalgo, and Veracruz (Map in Figure 1). Upper Necaxa Totonac, part of the Northern branch of the Totonacan language family (Brown et al. 2011), is spoken in four communities: Chicontla, Patla, San Pedro Tlaolantongo, and Cacahuatlán. The largest of the four communities is Chicontla with a population of 3305 people. The second largest and



Figure 1 Location of Totonacan languages.

geographically closest town to Chicontla (about 2.5 km away) is Patla with 1060 people. About 16 kilometers from Patla and Chicontla are the towns of San Pedro Tlaolantongo with a population of 1004 people, and Cacahuatlán, with 247 people; the four communities are located by the dark grey shaded oval on the map in Figure 1 (courtesy of David Beck).

Not all community members in the four villages speak Totonac; it is estimated that overall there are around 3000 speakers of UNT according to the Instituto Nacional de Estadística y Geografía census results (INEGI 2010). For this study, we focused on the two largest communities of Patla and Chicontla where most community members over 40 years of age are native Totonac speakers with high to native fluency in Spanish. Some people over the age of 65 years are monolingual Totonac speakers with relatively little Spanish. Speakers, generally under 30 years, may be considered heritage language speakers of Totonac, and only a small number of children are learning the language as a mother tongue. According to *Ethnologue*, the variety of Totonac spoken in Patla and Chicontla is threatened (category 6b), however, we may have reason to believe the situation is more critical than has been noted (Lam & Beck 2008; Lam 2009, 2012). Case studies in Lam (2009, 2012) indicate that the language is closer to a state of extinction, noting that the language ideology and socio-economic opportunities associated with the national language are the primary reasons for shifting toward a more monolingual Spanish community. Particularly in Patla and Chicontla, the older generation have been purposefully blocking transmission to younger generations, who are consequently no longer learning the language.

The Totonacan languages have complex polysynthetic agglutinative morphology as is common with languages in the Meso-American *sprachbund*. They have been described as having variable word and constituent order that poses challenges to typological theories. At the highest level, the internal genetic structure of the Totonacan language family is divided between Totonac and Tepehua. Totonac-Tepehua vowel inventories historically have three phonemic vowel qualities /i a u/ with contrastive length, where most languages in the Totonac branch make further distinctions in modal and laryngeal phonation creating an inventory of 12 vowel phonemes. In Totonac-Tepehua, with the exception of Spanish loanwords, the two

mid-vowels [e] and [o] are described as allophones of /i/ and /u/. The two vowels [e] and [o] are restricted to positions surrounding uvular or glottal consonants indicating that they emerged from lowering as a result of anticipatory assimilation (Beck 2014). However, the vowels /e/ and /o/ have obtained phonemic status in the Northern branch of the Totonac family constituting a shared innovation which distinguishes the languages of this group from the other languages in the family (Brown et al. 2011). In Upper Necaxa, part of the Northern branch of Totonac, /e/ and /o/ occur in contexts not predicted by their conditioning environments and in contexts where length and phonation are contrastive, which expand the inventory of vowel phonemes to 20. Upper Necaxa Totonac is also notable in being one of the only known Totonacan languages to have lenited the uvular stop, /q/, to a glottal stop, /ʔ/, in all environments, where the distribution of the glottal /ʔ/ parallels that of other stop-consonants (Beck 2006). The diachronic process of lenition affecting uvular stops is also attested in the emergence of a series of three glottalized or ejectives fricatives. The lenition of uvulars has made Upper Necaxa phonologically quite distinct from its sister Totonacan languages.

While there is some dialectal variation between the Totonac spoken in Patla and Chicontla, both communities have the same phonological consonant and vowel inventory, stress patterns, and phonotactics (Kirchner & Varelas 2002). The differences between the language in the two communities are seen in the choice of certain lexical items in nominal reference. For instance, differences occur in the names for certain plants and animals: in Patla *kukxápu'* /kukʰʃapɯ/ is a type of fruit-bearing tree used in making fans, whereas in Chicontla the correct word for this plant is *a'kxápu'* /qkʰʃapɯ/. There are also some common phonological reductions that are particular to the Patla dialect, such as *tawilá* /tawi'la/ > *to:lá* /to:'la/ 'sit down', *lhawá* /ʔa'wa/ > *lho:* /ʔo:/ 'make something', *tayá* /ta'ja/ > *te:* /te:/ 'take something' (Beck 2011). The only other distinctions that have been noted between the two communities are found in the choice of causative morpheme *ma'ha-* /mqʔa-/ or *ma:-* *-ni'* /ma:-ni/ in a limited set of verb stems (Beck 2011). Of the 93 words that were used in this study, there were only five lexical items that varied from one community to the other as noted in the wordlist in the appendix.

1.1 Consonants

Aside from the typologically rare ejective fricatives, /s' ʔ' ʃ'/, the Upper Necaxa Totonac consonant system is fairly typical of the language family. The consonant system has a typical series of voiceless obstruents that do not have contrastive voiced counterparts. Stops, affricates, and fricatives are parallel in places of articulation. Upper Necaxa Totonac also has a series of two nasals, and three approximants as indicated in Table 1.

Table 1 Upper Necaxa Totonac consonant inventory.

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p	t			k	ʔ
Affricate		ts	tʃ	tʃ		
Nasal	m	n				
Fricative		s	ʃ	ʃ	x	
Ejective fricative		s'	ʃ'	ʃ'		
Approximant				j	w	
Lateral approximant			l			

Table 2 Upper Necaxa Totonac vowel inventory.

	Front		Central		Back	
	Laryngeal	Plain	Laryngeal	Plain	Laryngeal	Plain
High	i i:	i i:			u u:	u u:
Mid	e e:	e e:			o o:	o o:
Low			q q:	a a:		

1.2 Vowels

The vowel system exhibits contrasting sets of short and long vowels and plain and laryngealized vowels within a system of five vowel qualities /i e a o u/. The contrast between plain, laryngealized, long and short vowels in Upper Necaxa can be observed in the minimal pairs, *chi*' /tʃi/ 'how?', *chi:*' /tʃi:/ 'tie something', *chi'chi*' /tʃitʃi/ 'hot' and *chichi*' /tʃitʃi/ 'dog'. Minimal pairs, or near minimal pairs, in citation form suggest that the contrast between length and laryngealization is phonemic. The various combinations of vowel length and phonation type are exemplified in Table 2.

All the vowel types are found in initial, final, and medial positions. The mid-vowels appear less frequent and most often occur in Spanish loan words and adjacent to glottal consonants. However, there are sufficient examples where these vowels appear outside of conditioning environments to demonstrate that they are distinct phonemes, e.g. *cho:lhchó:lh* /tʃo:tʃo:/ 'cricket' and *te:taxtú* /te:tʃtu/ 'come out' (see the wordlist in the Appendix).

The 20 phonemic vowels also appear in stressed and unstressed environments. Stress in UNT plays a grammatical function as it is correlated with a shift in the lexical class of an item and may indicate certain aspectual distinctions. Stress patterns are predictable and regular notwithstanding a few exceptions (Beck 2011). Stress falls on the final syllable of words that end in a long vowel, e.g. *pi'ta:jú:* /pi'ta:xu/ 'crush', or a closed syllable, like *chojót* /tʃo'xot/ 'saliva', or the penultimate syllable of words ending in a short vowel as in *pó:ti* /po:ti/ 'boat'. Ideophones, an onomatopoeic or synesthetic class of words related to adverbs in UNT (Beck 2008), are derived by reduplicative patterns and have no fixed primary stress, though out of context ideophones are generally stressed according to regular patterns: *texetexe* /teje'teje/ 'water bursting out' and *su:ksu:k* /su:k'su:k/ 'flapping or humming object'.

2 Method

2.1 Participants

The data for this research was collected in the villages of Patla and Chicontla in Puebla State, Mexico, from August to September 2012. Six native Upper Necaxa Totonac speakers participated in the study. Three female and two male speakers were from the village of Patla and one male speaker from the village of Chicontla. The youngest speakers included two of the females who were in their early to mid-thirties; the other four speakers were over the age of 60. All participants were native speakers of Totonac and Spanish and had good literacy skills.

2.2 Data collection

The wordlist in the appendix is comprised of 93 lexical items that were compiled from the *Upper Necaxa Totonac Dictionary* (Beck 2011). Participants were asked to put the word in the carrier phrase *xla wánli' ___ chuwá:* /ʃla 'wanli ___ tʃu'wa:/ 's/he said ___ now'. Speakers first practiced saying the target word in the carrier phrase, and then repeated the carrier phrase at least three times. The speakers were encouraged to read and talk about

each lexical entry in order to ensure the target word was prompted. Some few lexical items of which speakers were unsure or did not know the meaning were excluded. None of the speakers had difficulties identifying the words in the wordlist written in the UNT practical orthography since all participants were completely literate in Spanish and have had access and experience with written Totonac materials, like bible translations and the UNT dictionary.

The wordlist in the appendix includes lexical items one to five syllables long from the category of nouns, verbs, and adjectives. The selected lexical items in the construction of the wordlist were limited to words with vowels that appeared between plain voiceless stops /p t s ts ʃ tʃ ʎ/ in order to avoid any non-modal and vowel lowering effects from the ejective fricatives and glottal stops, or vowel lengthening effects from surrounding approximants or glides as suggested in Warner et al. (2004) and Kavitskaya (2014). The wordlist also included lexical tokens of ideophones because these items provided good environments for segmentation. Of the 93 items in the wordlist, only five lexical items differed for the Chicontla variety; the variation in the choice of lexical item was noted and transcribed accordingly.

Speakers were equipped with a head-mounted condenser microphone, and the recordings were made using a Marantz PMD 661 solid-state recorder with a modified pre-amp optimized for voice recording. Sound files were saved in 48 or 96 kHz 24 bit WAV format in an SDHC memory card. Vowels were segmented for the target word in all three productions of the phrase, except for some few word tokens where aliasing of the sound file or unexpected background noise (i.e. chickens crowing) occurred. Some speakers gave more repetitions of the phrase (with eight being the maximum number of repetitions for two items), and vowel tokens in those target words were measured as well.

2.3 Measurements

All vowels were coded for length, laryngeal, and length and laryngeal distinctions, as well as stress and syllable position (i.e. first, second, third, and less represented, fourth and fifth syllable). A total of 4151 vowel tokens were segmented (2076 vowel tokens from the females and 2075 from the males). The wordlist had representatives of all five vowel qualities /i e a o u/ and their variant length and phonation types for the exception of /ɔ:/ which is rarely found outside of conditioning environments which alter the vowels' phonation. Other discrepancies include the lack of long laryngeal /a/ and /u/ which were not found in the testing environments in the present study. The distribution of vowel tokens in both stressed and unstressed positions is illustrated in Table 3.

Table 3 indicates that there is an uneven distribution of vowel type represented in the data, which is likely representative of the distribution of vowels in the language generally. Each vowel type is represented in the study, with modal (plain), unstressed, and short vowels being the most representative of the sample.

Table 3 Number of vowel tokens for vowel type.

Vowel	Unstressed	Stressed	Vowel	Unstressed	Stressed	Vowel	Unstressed	Stressed
i	115	44	u	175	95	a	343	106
i:	96	99	u:	182	76	a:	181	40
ɨ	656	194	ɥ	228	75	ɔ	379	128
ɨ:	17	11	ɥ:	0	22	ɔ:	13	0
Vowel	Unstressed	Stressed	Vowel	Unstressed	Stressed			
e	192	38	o	153	35			
e:	81	51	o:	20	165			
ɛ	23	3	ɔ	62	32			
ɛ:	21	0	ɔ:	0	0			

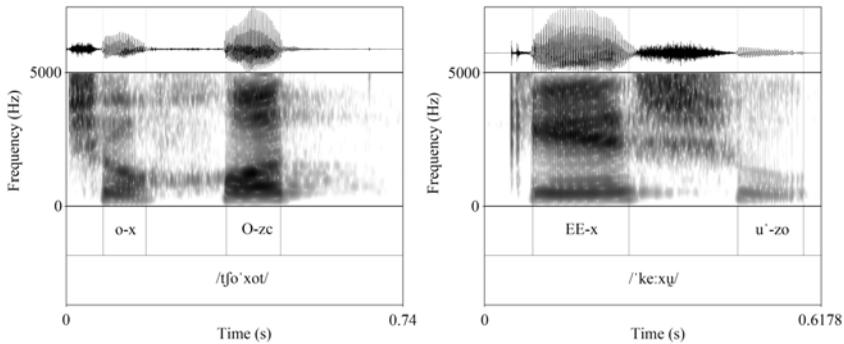


Figure 2 Spectrogram of *chojót* /tʃoˈxot/ and *ké:xu* /ˈkeːʃu/ illustrating the markup of the vowels from the recorded wordlists. The vowels in the first tier were marked up using the following code (as described in Section 2.3): `_x` signifies that the vowel appears in syllable-initial position, `_zc` marks that the vowel appears in a closed final-syllable position, `_zo` in a final open syllable, a capital vowel indicates a stressed vowel, two vowels indicate a stressed vowel, and an apostrophe indicates a laryngeal vowel.

Vowel tokens were segmented using Praat (Boersma & Weenink 2013) and the acoustic measures were extracted using a customized Praat script. Though each vowel token was segmented between plain voiceless stops, some stop consonants were voiced as a result of voicing assimilation. The onset of the vocalic nuclei was specified by a combination of the voice bar, the F2 band, auditory cues, and temporal cues in the waveform. The offset of the vocalic nuclei was selected at the point of lowest energy, where the F2 band is still clearly visible, and the repetitive pattern on the waveform appear to fade. A spectrogram of the word *chojót* /tʃoˈxot/ ‘saliva’ and *ké:xu* /ˈkeːʃu/ ‘cheese’ by one speaker is exhibited in Figure 2.

Figure 2 demonstrates that vowel segmentation was marked by a combination of the F2 band and the spectral peak display. The spectrograms further illustrate how the vowel tokens were coded. The `_x` signifies that the vowel appears in syllable-initial position with a consonant onset (CV) as do all words in the list in the appendix. Since a phonetic glottal stop is inserted in vowel-initial words, these environments were avoided due to the potential effects of vowel lowering in consonantal glottalic environments. Vowels labelled with `_zc` marks that the vowel appears in a closed final-syllable position, and `_zo` in a final open syllable position. Capitalization marks a stressed syllable, double vowels mark length, and the straight apostrophe marks laryngealization. For some vowel tokens that exhibited non-modal effects at the onset of the vocalic nuclei, boundaries were drawn to the grey F2 band in combination with a prominent vocalic burst on the spectral peak display. For vowel tokens which appeared in word-final open syllable position, the offset of the vowel boundaries were pushed further out to capture any residual voicing and non-modal effects. For accurate positioning, the start and end selections were zero-crossed providing a standard onset and offset point for all vowels close to the original selection.

The formant tracking and f_0 were adjusted such that a best-fit setting was found for individual male or female voices. These settings were used for automatic extractions of the acoustic measures and did not exceed 5500 Hz at five formants for males and 5000 Hz at five formants for females. The duration, fundamental frequency (f_0), F1, F2, and F3 values were measured. Formant and f_0 values were also extracted at the 25%, 33%, 50%, and 75% percent points throughout the vowel. For the purposes of the present paper, we use the f_0 and formant measures collected at the mid-point of each vowel.

2.4 Analysis procedures

Vowel plots were generated in R using phonTools (Barreda 2013). The frequency data was normalized using the Nearey (1978) extrinsic method, which minimizes the variation between

speakers due to social, physiological or anatomical differences, as implemented in the phonTools package. The Nearey extrinsic method was selected because there were unequal numbers of each vowel in a category and because it is vowel-extrinsic and formant-intrinsic, which reduces the physiological variation while maintaining sociological and dialectal variation. All f_0 values were converted into semitones (Nolan 2003, Shih & Lu 2015). The semitone conversion has been used to normalize f_0 values across speakers of different sizes and genders. The conversion applied follows Shih & Lu (2015), which uses the individual speakers' median f_0 as the base and converts the individual frequency measures into semitones. Statistical analyses of the formant measures, durational data and f_0 values are described in the results section.

All data were analyzed using linear mixed-effects models as implemented in the lme4 package (Bates et al. 2015) in the R statistical programming language (R Core Team 2017). Subject and item were included as random-effects in all models. The individual model structure is reported for each model below. Each model originally included the following variables of interest: stress (yes, no), laryngeal (yes, no), and length (long, short). Additionally, the control variables include: sex (male, female), position (initial, middle, final), repetition (1–8), vowel quality (a, e, i, o, u), and number of characters (5–13), which is a count of the characters in ASCII based on the phonemic transcription of the word where long vowels and affricates are represented with two characters. This character count is an imperfect measure but provides a close approximation of word length in number of phonemes. The control variables have been included to account for possible effects of position, repetition, and word length, which are not a focus of this paper but can influence the acoustic characteristics analyzed below. The dependent variables were fit using a backwards step-wise fitting procedure starting with all variables and removing predictors that were not significant. All possible random-slopes were investigated using a forward-fitting procedure, keeping only those slopes that significantly improved the fit of the model. Two-way interactions with variables of interest were also explored in the modeling and retained when found to significantly improve the fit of the model. Model comparisons were performed using the ANOVA function to compare the fit of models to the data. Predictor significance was determined using the t -distribution; values with a t -distribution greater than 2 were considered significant (Baayen, Davidson & Bates 2008). For the purpose of publication, p -values are approximated using the Satterthwaite method as implemented in the lmerTest package (Kuznetsova, Brockhoff, & Christensen 2016).¹ In all tables reporting statistical results significant effects have been marked with an asterisk in a separate column.

3 Results

3.1 Vowel quality

To explore the nature of the formant space in Upper Necaxa, separate linear mixed-effects analyses for F1 and F2 were performed. Figure 3 plots vowel spaces for the modal and laryngeal vowels split by vowel length using the mean values of F1 and F2. All frequency measurements were made at the midpoint of the vowel. The plots in Figure 3 suggest that the five-vowel system creates a V-shaped system rather typical of languages like Spanish with five vowels. The vowel types are relatively evenly distributed maximizing the space between vowels as much as possible. The normalized values in the vowel plots demonstrate that long and short modal and laryngeal vowels occupy roughly the same space acoustically. Stressed vowels (represented with capital letters), however, vary in position from the unstressed vowels suggesting possible differences in F1 and F2 for stress.

¹ We would like to note that Bates et al. (2015: 34) indicate that any method for approximating degrees of freedom is 'at best *ad hoc*'.

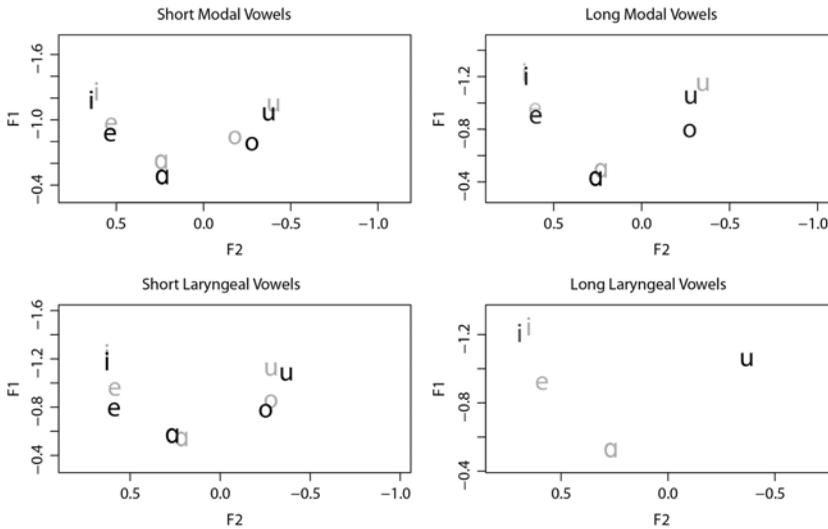


Figure 3 Comparison of F1-F2 values of Nearey normalized vowels from all six participants calculated at the midpoint of the vowel, with short vowels on the left and long vowels on the right. Laryngeal vowels are on the bottom row and modal vowels are on the top row. Black font indicates stressed vowels and grey font unstressed vowels.

The first model investigated the Nearey normalized frequency of F1 as the dependent variable with the independent variables: stress, laryngeal, length, sex, and position. This model included random effects for word, subject, and vowel quality, with random slopes for stress by subject. The final model is summarized in Table 4. The intercept values for this model are: unstressed, long, modal, female, and medial.

The model summary in Table 4 indicates that F1 may vary depending on stress, position, or repetition of the target word. The results indicate that when the vowel is stressed, the frequency of F1 is significantly higher and hence articulated at a lower tongue position. Additionally, Table 4 suggests that the participants' F1 values are lower with more repetitions, so that with each repetition tongue height generally increases. The other predictors all interact with syllable position. A complex interaction between F1 and syllable position is

Table 4 Results table of linear mixed-effects analysis of F1.

Fixed effects	Estimate	Std. Error	df	<i>t</i> -value	Pr(> <i>t</i>)	
(Intercept)	-0.88530	0.11915	4	-7.880	0.0010	*
Stress: yes	0.06882	0.01735	5	3.966	0.0097	*
Length: short	-0.05290	0.00793	4050	-6.679	0.0000	*
Laryngeal: yes	-0.00930	0.00686	4084	-1.350	0.1771	
Sex: male	-0.04540	0.01885	5	-2.406	0.0656	*
Position: initial	-0.01990	0.00891	4055	0.880	0.3791	
Position: final	-0.01360	0.00999	4027	3.994	0.0001	*
Repetition	-0.00570	0.00128	3734	-4.486	0.0000	*
Sex: male × Position: initial	0.01293	0.00744	3975	1.738	0.0823	
Sex: male × Position: final	-0.04920	0.00748	3831	-6.582	0.0000	*
Laryngeal: yes × Position: initial	0.00699	0.00859	4077	0.814	0.4156	
Laryngeal: yes × Position: final	0.03403	0.00902	4058	3.773	0.0002	*
Length: short × Position: initial	0.02647	0.01046	4078	-2.531	0.0114	*
Length: short × Position: final	0.04578	0.01084	4083	-4.223	0.0000	*

* = significant *t*-values (with an absolute value greater than 2).

Table 5 Linear mixed-effects model of F2.

Fixed effects	Estimate	Std. Error	df	t-value	Pr(> t)
(Intercept)	0.12683	0.20039	4	0.633	0.5631
Stress: yes	-0.02260	0.00498	4089	-4.535	0.0000 *
Length: short	0.05872	0.01589	10	3.696	0.0044 *
Laryngeal: yes	-0.01860	0.00818	4073	-2.272	0.0231 *
Sex: male	-0.01590	0.01286	42	-1.239	0.2225
Position: initial	0.01044	0.01064	4058	0.981	0.3268
Position: final	0.02026	0.01193	4087	1.699	0.0894
Position: initial × Length: short	-0.03230	0.01250	4086	-2.581	0.0099 *
Position: final × Length: short	-0.01920	0.01294	4067	-1.481	0.1386
Position: initial × Laryngeal: yes	0.03379	0.01026	4087	3.294	0.0010 *
Position: final × Laryngeal: yes	-0.00220	0.01080	4084	-0.206	0.8367
Position: initial × Sex: male	0.00048	0.00876	3986	0.055	0.9561
Position: final × Sex: male	0.02382	0.00883	3954	2.698	0.0070 *

* = significant t-values (with an absolute value greater than 2).

found in this case with male speakers having lower F1 values (a higher tongue position) when the vowel is at the end of the word. Laryngeal vowels in final position have a higher F1 than in medial environments, and short vowels tend to have higher F1 in initial and final positions as opposed to medial position where F1 is lower.

Table 5 summarizes the final model analyzing the Nearey normalized F2, which includes random effects for word, speaker and vowel quality. Significant model improvements were found with random slopes for length and position by speaker. In this analysis, repetition does not turn out to be a significant predictor and has been removed from the final model. The intercept values are: unstressed, long, modal, female and medial.

The model reported in **Table 5** suggests that stress is again a significant predictor, however, in this case the F2 of the stressed vowels are lower indicating that the tongue is further back in articulation. There are also interactions between position and three other predictors: length, laryngeal, and sex. For laryngeal vowels, F2 is higher in initial position indicating that laryngeal vowels in this environment are produced with the tongue in a more front position. For long vowels, F2 is lower in initial position as compared to the medial position suggesting that syllable-initially long vowels are produced with the tongue further back. Male speakers also have higher F2 values, and thus a more front tongue position, in word-final environments.

The linear mixed-effects regression performed for both Nearey normalized frequency values indicates that the qualitative difference between stress and unstressed vowels is significant, where stressed vowels are produced lower and further back than unstressed vowels; this pattern seems to be consistent in the data. We now turn towards duration to investigate the distinction between vowel quantity and stress.

3.2 Duration

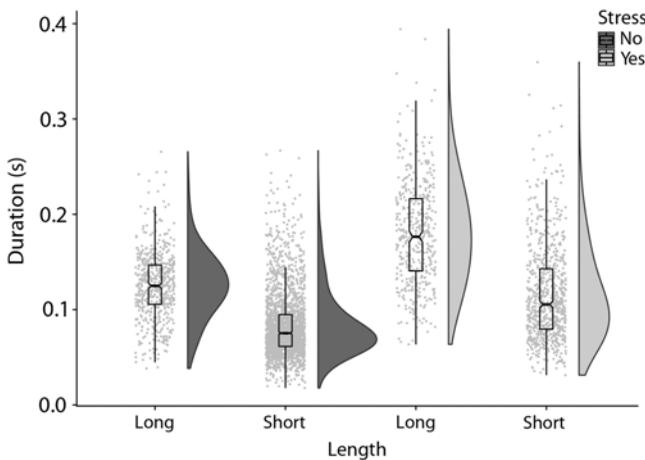
In this part of our analysis, we aim to verify the duration difference between long and short vowels and to test Beck's characterization of stressed vowels in UNT with phonetic analysis. **Table 6** provides the mean duration of each short and long vowel split across stressed and unstressed vowels. The table demonstrates that vowel duration is a phonetic correlate of long versus short vowels, and a phonetic correlate of stress, with a consistent pattern of stressed vowels being longer than unstressed vowels in both their short and long counterparts.

The raincloud plot (Allen et al. 2018) in **Figure 4** summarizes the general patterns described in **Table 6** with long-unstressed, short-unstressed, long-stressed and short-stressed vowels. The raincloud plot includes the data distribution, the individual data points, and a boxplot, which indicates the 25th and 75th percentiles with the notch indicating the estimated 95% confidence interval. **Figure 4** indicates that there are duration differences between

Table 6 Mean duration in seconds of short, long, stressed and laryngealized vowels from all six participants.

Vowel	Duration		Vowel	Duration		Vowel	Duration	
	Unstressed	Stressed		Unstressed	Stressed		Unstressed	Stressed
i	0.065	0.147	u	0.088	0.141	ɑ	0.082	0.117
i:	0.112	0.144	u:	0.121	0.179	ɑ:	0.132	0.170
ĩ	0.077	0.091	ũ	0.086	0.113	ɔ	0.090	0.111
ĩ:	0.110	0.124	ũ:	—	0.129	ɔ:	0.138	—

Vowel	Duration		Vowel	Duration	
	Unstressed	Stressed		Unstressed	Stressed
e	0.085	0.127	o	0.086	0.145
e:	0.131	0.181	o:	—	0.218
ẽ	0.075	0.152	õ	0.104	0.113
ẽ:	0.148	—	õ:	—	—

**Figure 4** Raincloud plot (Allen et al. 2018) demonstrating the data distribution (in seconds), the individual data points, and a boxplot of long and short, stressed and unstressed vowels in UNT. The boxplot indicates the 25th and 75th percentiles with the notch indicating the estimated 95% confidence interval.

phonemically long and short vowels and that there may also be duration effects due to stress in Upper Necaxa.

Figure 4 indicates that there is a lot of variation and a wide overlapping range between both vowel quantities, where most long vowels clearly show a distinction in length compared to the short vowels. The general effects of vowel length and stress were tested using linear mixed-effects regression with the logged duration as the dependent variable and the predictors described above as control variables. The model was fitted using a backward step-wise fitting process, removing non-significant comparisons. Table 7 demonstrates the final model which included subject and item as random effects, and stress, laryngeal, and repetition by subject as random slopes, with the intercept values: unstressed, long, modal, medial, female, and /a/.

The results of the analysis presented in the model summary in Table 7 indicate that there is a significant interaction between stress and vowel length such that long vowels are significantly longer than short vowels, and short stressed vowels are significantly longer than short unstressed vowels. The interaction indicates that the lengthening effect for stressed vowels does not occur for the long vowels. In other words, the lengthening of stressed vowels is only significant for short vowels. The final model was relevelled to check comparisons between the levels of length and stress. The resulting models indicated that all comparisons of length and stress were significantly different except for the comparison between unstressed and stressed

Table 7 Linear mixed-effects model of the vowel duration.

Fixed effects	Estimate	Std. Error	df	t-value	Pr(> t)	
(Intercept)	-2.15400	0.06898	1136	-31.226	0.0000	*
Stress: yes	0.03544	0.05802	6	0.611	0.5621	
Length: short	-0.53720	0.01673	3973	-32.105	0.0000	*
Laryngeal: yes	0.00685	0.02567	7	0.267	0.7977	
Number of Characters	0.00478	0.00722	2048	0.662	0.5077	
Repetition	0.03521	0.01858	177	1.895	0.0597	
Position: initial	0.47613	0.05257	3989	9.058	0.0000	*
Position: final	0.32971	0.05313	3899	6.205	0.0000	*
Sex: male	0.05043	0.02706	44	1.863	0.0691	
Vowel Quality: /e/	0.05343	0.02305	3036	2.318	0.0205	*
Vowel Quality: /i/	-0.19450	0.01750	2453	-11.116	0.0000	*
Vowel Quality: /o/	-0.00050	0.02362	2324	-0.019	0.9846	
Vowel Quality: /u/	-0.10150	0.01600	3818	-6.345	0.0000	*
Stress: yes × Length: short	0.29724	0.02864	3931	10.377	0.0000	*
Stress: yes × Laryngeal: yes	-0.15990	0.02472	4075	-6.466	0.0000	*
Number of Characters × Repetition	-0.00450	0.00188	3640	-2.398	0.0165	*
Number of Characters × Position: initial	-0.04480	0.00525	4011	-8.528	0.0000	*
Number of Characters × Position: final	0.01693	0.00542	3918	3.124	0.0018	*
Position: initial × Sex: male	-0.06820	0.02160	3897	-3.156	0.0016	*
Position: final closed × Sex: male	0.09503	0.02171	3684	4.378	0.0000	*
Repetition × Position: initial	-0.02020	0.00912	3574	-2.215	0.0268	*
Repetition × Position: final	-0.04350	0.00916	3565	-4.755	0.0000	*

* = significant t-values (with an absolute value greater than 2).

long vowels (as seen in Table 7). The results suggest that Beck's auditory perception of duration and the correlation between duration and stress are largely supported by the acoustic analysis. Stressed vowels in general have increased duration, though his phonological description falls short on the durational distinctions between short stressed vowels and long unstressed vowels in that long unstressed vowels were found to be consistently longer than short stressed vowels. Additionally, there was an interaction between the predictors laryngeal and stress, which indicates that while unstressed modal vowels were shorter than the stressed modal vowels, the stressed modal vowels were significantly longer than the stressed laryngeal vowels.

Vowel quality was also a significant predictor of vowel duration with /e/ having the longest average duration and /i/ the shortest average duration. Other observations include a general effect of position in the word with vowels in initial and final position occurring with longer durations than in medial positions. Position further interacts with sex and repetition, and the difference between the positions become weaker as the word is repeated. The duration of vowels in male speakers are also significantly longer in final position than the female speakers, whereas in initial position, this duration effect is reversed. All other effects are considered control effects and are not discussed in detail but are reported in Table 7.

3.3 Fundamental frequency

In this section, the role of fundamental frequency (f_0) is investigated in relation to the differences in stress. The median f_0 was calculated for each speaker and this value was converted to semitones (using each speaker's average f_0 value). A linear mixed-effect analysis was then performed with the semitone values as the dependent variable and stress (yes, no), laryngeal (yes, no), length (long, short). Additionally, the control variables include: sex (male, female), position (initial, middle, final closed, final open), repetition (1–8), vowel quality (a, e, i, o, u), and number of characters (5–13) as the independent variables. Figure 5, also summarized in

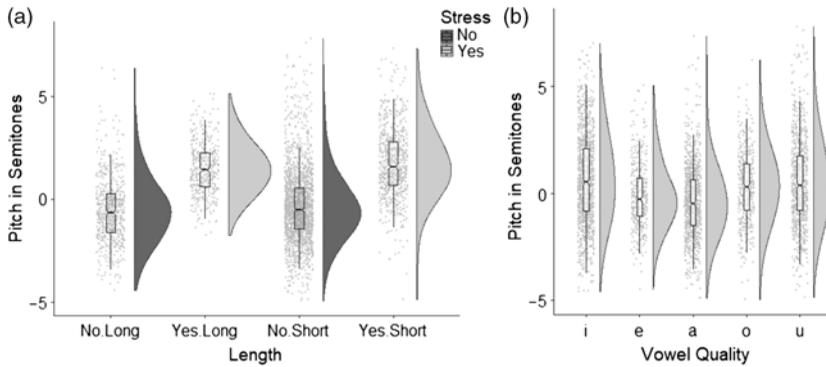


Figure 5 (a) Raincloud plot (Allen et al. 2018) demonstrating the data distribution (in semitones) the individual data points, and a boxplot of long and short, stressed and unstressed vowels in UNT. The boxplot indicates the 25th and 75th percentiles with the notch indicating the estimated 95% confidence interval. Bar plot demonstrating mean values in semitones based on individual speakers' median f_0 of long and short, stressed (Yes) and unstressed (No) vowels in UNT. (b) Raincloud plot illustrating the effect of intrinsic f_0 in the vowels of UNT. In both plots extreme outliers were excluded (3.3% of the data).

Table 8 Linear mixed-effects model of f_0 .

Fixed effects	Estimate	Std. Error	df	<i>t</i> -value	Pr(> <i>t</i>)
(Intercept)	-0.62440	0.64381	57	-0.970	0.3362
Stress: yes	1.70365	0.47446	10	3.591	0.0051 *
Position: initial	0.48277	0.57541	99	0.839	0.4035
Position: final	-0.17760	0.58436	308	-0.304	0.7614
Length: short	0.69794	0.21307	3913	3.276	0.0011 *
Number of Characters	-0.05710	0.05791	2364	-0.985	0.3246
Repetition	-0.04570	0.13260	3639	-0.345	0.7304
Sex: male	-0.18600	0.24329	47	-0.764	0.4484
Laryngeal: yes	-0.76590	0.25138	13	-3.047	0.0094 *
Vowel Quality: /e/	0.51366	0.19662	1844	2.612	0.0091 *
Vowel Quality: /i/	1.29647	0.14772	1416	8.777	0.0000 *
Vowel Quality: /o/	0.27034	0.19250	1558	1.404	0.1604
Vowel Quality: /u/	1.07091	0.13984	2784	7.658	0.0000 *
Stress: yes × Length: short	1.08514	0.31057	2540	3.494	0.0005 *
Position: initial × Length: short	-1.01290	0.27999	3879	-3.618	0.0003 *
Position: final × Length: short	-1.88460	0.36363	3480	-5.183	0.0000 *
Position: initial × Number of Characters	0.02232	0.04896	3915	0.456	0.6485
Position: final × Number of Characters	0.16739	0.05024	3931	3.332	0.0009 *
Number of Characters × Repetition	-0.04110	0.01623	3642	-2.534	0.0113 *
Sex: male × repetition	0.26244	0.06802	3625	3.859	0.0001 *
Position: initial × Laryngeal: yes	0.22649	0.22459	3910	1.008	0.3133
Position: final × Laryngeal: yes	0.84306	0.24030	3924	3.508	0.0005 *

* = significant *t*-values (with an absolute value greater than 2).

Table 8, illustrates the final linear mixed-effects model of f_0 , which included subject and item as random effects and random slopes for stress, laryngeal, and position with subject. All other random slopes did not improve the fit of the model and were excluded from the final model. The data was also analyzed in Hertz (not reported here), which indicates that the analysis of the untransformed data has largely the same results.

The results of the analysis of f_0 in semitones indicate that there is an interaction between stress and length. The effect of stress on f_0 is strongest for the short vowels and weakest

for the long vowels. All comparisons were found to be significant after releveling except for the comparison between unstressed short and long vowels, which were found not to have significant f_0 differences. Other effects of note were that vowel quality significantly predicted the f_0 of the vowel in further support of the intrinsic f_0 research in, for example, Whalen & Levitt 1995, with the high vowels having the highest f_0 and lower vowels having decreasing f_0 values. We also found an interaction between f_0 and position. The model was relevelled to compare partial effects of this interaction. We found that for the medial (reported in Table 8) and initial ($\beta = -0.539$, $SE = 0.239$, $p < .046$) positions pitch was significantly lower for the laryngealized vowels and that there was no difference for the vowels in final position.

4 Discussion

The present study provided an acoustic description of the vowel system with contrasting length and laryngealization, and the effects of stress in Upper Necaxa Totonac. While the Totonacan languages generally and historically have three vowel qualities /i u a/ with length and laryngeal distinctions, the study provides empirical support for a five-vowel system in UNT, which could have come about as part of the natural progression of the language and certain conditioning environments, or influenced by Spanish, the socio-economically dominant language of the region. The acoustic measurements demonstrate that the system is typical of languages with five vowels in that the vowels occupy the perimeters of the vowel space with the mid-vowels spaced at intermediate distances, much like the Spanish vowel system.

The present study has further demonstrated that the different vowel types, short, long, plain, and laryngeal all occupy a similar acoustic formant space. The fact that we find no difference in the formant frequencies across long and short vowels is contrary to Aschmann's (1946) findings that patterns vary in quality based around length for Zapotitlán Totonac. Additionally, f_0 was a significant factor in vowel quality where high vowels were found to have higher f_0 values and lower vowels decreasing f_0 . While not one of the main goals of the present study, this result replicates the intrinsic f_0 findings in Whalen & Levitt (1995) and adds one more empirical data point to that body of research. Fundamental frequency also predictably shows a difference between modal and laryngeal vowels, where laryngeal vowels showed significantly lower f_0 values than the modal vowels.

Additionally, the distinction between phonemic long and short vowels is durational with long vowels being about one and a half times longer than short vowels. Previous studies have shown that duration is the most salient feature for identifying the contrast between long and short vowels, though other cues, like formant transitions, syllable position, and the fundamental frequency may play secondary roles in length perception (see Lehnert-LeHouillier 2010). The present study shows that there were no significant effects on short versus long vowels in terms of f_0 or formant frequencies, and that vowel length seems to play a dominant role in phonemic vowel length.

The results further show that there is a slight quality distinction associated with stress not previously observed by the field researcher. While stressed vowels have increased duration, there are also slight distinctions in vowel height and backness. The acoustic realization of stress is indicated by many cues, where in general the F1 is higher and the F2 is lower in stressed vowels, indicating that stressed vowels are produced at a lower and further back tongue position than unstressed vowels. In terms of duration, stress is indicated by longer vowels, but unstressed long vowels are not significantly different from stressed long vowels. The acoustic evidence for the three claims made in Beck (2004) on the relation of stress and duration of vowels in UNT is summarized in (i)–(iii). The data demonstrate that the impressionistic observation about stress and duration were correct in two of the three statements.

- (i) Stressed vowels have increased duration. (TRUE)
- (ii) Stressed short vowels have the same duration as long vowels. (FALSE)
- (iii) Stressed long vowels are not consistently longer than unstressed long vowels. (TRUE)

That vowel duration is consistently longer in stressed vowels supports that duration is a significant cue in stress even though it may not be the only one. Furthermore, while laryngeal vowels made no difference in the formant analysis, there were some notable differences in the duration in interactions with stress, which indicates that stressed modal vowels are significantly longer than the stressed laryngeal vowels, an observation that is counter to other phonetic studies concerning non-modal phonation (Kirk et al. 1993).

One observation concerning the lengthening effect of stress on short vowels, but not on long vowels, may be explained by the differences in vowel quality in stressed environments overall. This observation, which has been termed target undershoot in Lindblom (1963), indicates that the shorter the vowel, the less time the articulators have in transitioning and therefore the faster the articulators must move. Flemming (2006) further notes that target undershoot is a consequence of minimizing the effort it takes to move the articulators quickly at the expense of not reaching the vowel target. If stressed vowels in general are produced with higher F1 and lower F2 values, stressed short vowels may require more time for transitioning from a lower and further back tongue position which may result in vowel lengthening. Since stressed long vowels are also produced with a lower and further back tongue position, there may be enough time in the articulation of long vowels for formant transition, and no need for further vowel lengthening in prosodic contexts.

Even though stressed vowels are acoustically different, it would be of notable value to conduct perceptual experiments that could elucidate the cues listeners are using to differentiate stress from other unstressed (long) vowels. For instance, while there is a lengthening effect for stress, that effect does not occur for long vowels. One observation from a perceptual perspective is that if there were no other cues differentiating long stressed from long unstressed vowels, then the stressed long vowels would be perceptually indistinguishable with the unstressed long vowels making this aspect of the vowel system confusing for a listener. However, the results suggest that distinguishing long stressed vowels from long unstressed vowels may be reliant on correlations with f_0 . Fundamental frequency was found to be a significant cue for stress, in which stressed vowels have higher f_0 values overall though this effect was weakest for the long stressed vowels.

5 Conclusion

The Totonacan languages have a number of interesting typological properties including vowel systems that contrast in both length and phonation. This study provides a first investigation of the frequency and durational characteristics of modal and non-modal vowels in Upper Necaxa Totonac. The acoustic measurements support previous descriptions of the quality and quantity for long, short, modal and laryngeal vowels. The results indicate that the short, long, modal, and laryngeal vowels occupy the same space and distribution in the vowel system, with some differences in formant frequency values due to the effect of stress. Vowel duration was also a significant predictor of stress but only for short vowels, indicating that duration is not a significant cue for long stressed vowels. The phonetic analysis was also used to test Beck's (2004, 2011) observations between stress and duration in Upper Necaxa Totonac. The study demonstrated that Beck's observations were correct for two out of three predictions on his perceptual correlates of stress and length. Stressed vowels were significantly longer than unstressed vowels, but long-vowels appeared with greater duration than short-stressed vowels. Long-stressed vowels were not consistently longer than long-unstressed vowels. These results imply that speakers might use duration as a distinguishing factor between these categories, but must also use a variety of other cues like format frequency and f_0 to distinguish stress. The study further investigates some of the interactions concerning f_0 , syllable position and repetitions in relation to quality and quantity distinctions, such that for example, vowel duration was longer in syllable-initial and syllable-final positions; however, effects became weaker with every repetition. Future research needs to investigate how vowel and syllable length plays a role in prosody and phrase-level marking, as has been investigated

in two Apachean languages, which demonstrated that morphological factors can be used to predict certain lengthening environments (Tuttle 2005). Upper Necaxa, like most indigenous languages of the world, is still under-described, with technical acoustic descriptions of the sound inventories representing an especially glaring lacuna. Acoustic studies of these languages represent an underdeveloped but critical source of typological data on the properties of sound and primary cues in sound perception and production. Phonetic descriptions of this type not only contribute to linguistic typology and the acoustic signals associated with vowel sounds, but also the prosodic factors that guide native-like pronunciations important to language learning and language revitalization.

Acknowledgements

Our gratitude goes to our participants in Patla and Chicontla for their engagement and enthusiasm with the research project. Many thanks to David Beck for his support and connections to community members, and to Conor Snoek and the anonymous reviewers for their helpful comments and feedback. This project was funded by a grant awarded to the first author from the Jacobs Research Fund and in part by the Social Sciences and Humanities Research Council for the Upper Necaxa Totonac Field Project.

Appendix. Wordlist for Patla and Chicontla

Orthography	International Phonetic Alphabet	Part of speech	Gloss	Chicontla
<i>chejé:t</i>	[tʃeˈxe:t]	<i>n</i>	hail	—
<i>chí'xí'xi'</i>	[tʃiˈʃiːʃiː]	<i>adj</i>	furry	—
<i>cho:lhchó:lh</i>	[tʃo:ʔtʃo:ʔ]	<i>n</i>	cricket	<i>matzolhtzólh</i>
<i>chojót</i>	[tʃoˈxot]	<i>n</i>	saliva	—
<i>chú'kupu:pi'tzi</i>	[tʃukupu:piˈtʃiː]	<i>vt</i>	saw in half	—
<i>ji:chí'ji:chi'</i>	[xi:tʃiːxi:tʃiː]	<i>idph</i>	rumble	—
<i>ké:xu'</i>	[ˈke:ʃu]	<i>n</i>	cheese	—
<i>kí'plhkí'plh</i>	[kɪpʔkɪpʔ]	<i>idph</i>	rustling	—
<i>kí'pski'ps</i>	[kɪpskɪps]	<i>idph</i>	skipping	—
<i>kí'px</i>	[kɪpʃ]	<i>idph</i>	clicking	—
<i>kí'sít</i>	[kɪˈsit]	<i>n</i>	lice	—
<i>kit</i>	[kit]	<i>prn</i>	I, me	—
<i>kitz'ks</i>	[kiˈtʃiːks]	<i>adv</i>	crammed	—
<i>ku'k</i>	[kuk]	<i>n</i>	uncle	—
<i>kuká't</i>	[kuˈkat]	<i>n</i>	oak	—
<i>kukú:x</i>	[kuˈku:ʃ]	<i>n</i>	sapote	—
<i>kukxápu'</i>	[kukˈʃapu]	<i>n</i>	type of tree	<i>a'kxápu'</i>
<i>lakaka:xtú</i>	[lakaka:ʃtu]	<i>vt</i>	chop	—
<i>lakapu'pu'kú:n</i>	[lakapupuˈku:n]	<i>n</i>	type of bee	—
<i>latojó:</i>	[latoˈxo:]	<i>vi</i>	be wet	—
<i>lhpa'pa'pa'</i>	[ʔpapapa]	<i>idph</i>	vibrating	—
<i>lhpoj</i>	[ʔpox]	<i>n</i>	type of plant	<i>lhpaúj</i>
<i>lojój</i>	[loˈxox]	<i>adv</i>	loose	<i>lojónh</i>
<i>lunlhó'lh</i>	[lunˈʔoʔ]	<i>adv</i>	soaked	—
<i>pá'xni'</i>	[ˈpaχniː]	<i>n</i>	pig	—

Continued.

Orthography	International Phonetic Alphabet	Part of speech	Gloss	Chicontla
<i>pi'tajú:</i>	[pi'ta'xu:]	<i>vt</i>	crush	—
<i>pi'xi'</i>	[pi'xi]	<i>n</i>	wart	—
<i>pi:tzi'ki'</i>	[pi:tsi'ki]	<i>n</i>	nipple	—
<i>pó:ti</i>	[po:ti]	<i>n</i>	pot	—
<i>po'jó'jo'</i>	[po'xo:xo]	<i>adj</i>	leafy	—
<i>pu:pejpe'h</i>	[pu:pex'peʔ]	<i>n</i>	type of toy	—
<i>pu:chú:t</i>	[pu:'tju:t]	<i>n</i>	type of tree	—
<i>pu:sa:tojó:</i>	[pu:sa:to'xo:]	<i>vt</i>	get stuck	—
<i>pu:tekáte</i>	[pu:te'kate]	<i>n</i>	beer can	—
<i>sa:pu:tojó:</i>	[sa:pu:to'xo:]	<i>vt</i>	get stuck	—
<i>sá:sti'</i>	[sa:sti]	<i>adj</i>	new	—
<i>sa:tojó:</i>	[sa:to'xo:]	<i>vi</i>	get stuck	—
<i>sapa'sapa'</i>	[sapsapa]	<i>idph</i>	crunching	—
<i>sapatu</i>	[sapatu]	<i>n</i>	shoe	—
<i>se:táj</i>	[se:'tax]	<i>n</i>	thread	—
<i>si'pi:</i>	[si'pi:]	<i>vt</i>	grind	—
<i>si'pi:lhtu'ki'ta'</i>	[si'pi:ltu'ki'ta]	<i>n</i>	type of atole	—
<i>sipéj</i>	[si'pex]	<i>n</i>	hill	—
<i>sipejchichi'</i>	[sipextʃi'tʃi]	<i>n</i>	coyote	—
<i>ska'tán</i>	[ska'tan]	<i>n</i>	type of deer	—
<i>ski'tí</i>	[ski'ti]	<i>vi</i>	make tortillas	—
<i>ski:'ti'</i>	[ski'ti]	<i>n</i>	fish	—
<i>skukij</i>	[sku'kux]	<i>n</i>	type of plant	—
<i>slajája'</i>	[sla'xaxa]	<i>adj</i>	sharp, prickly	—
<i>spa'ta'ta'</i>	[spatata]	<i>idph</i>	oozing	—
<i>spalhalha</i>	[spa'la'la]	<i>idph</i>	water flowing	—
<i>spé:ju</i>	[spe:xu]	<i>n</i>	window, mirror	—
<i>spi'pi'spi'pi'</i>	[spipispipi]	<i>idph</i>	tremble	—
<i>spujúju'</i>	[spu'xuxu]	<i>adj</i>	sharp	—
<i>su:ksu:k</i>	[su:ksu:k]	<i>idph</i>	flap, hum	—
<i>su'ku'ku'</i>	[su'kuku]	<i>adj</i>	perforated	—
<i>sú'tsu't</i>	[sutsut]	<i>idph</i>	tear	—
<i>ta:lhto'jó'lu'</i>	[ta:ltu'xo:lu]	<i>adj</i>	hanging	—
<i>ta'ja'ja'</i>	[taxaxa]	<i>idph</i>	light moving	—
<i>ta'jatát</i>	[taxa'tat]	<i>n</i>	sickness	—
<i>ta'pú'j</i>	[tax'puj]	<i>adv</i>	overcast	—
<i>ta'satí:</i>	[taxa'ti:]	<i>vt</i>	call someone	—
<i>ta:'pa:ti:</i>	[tax'pa:ti:]	<i>vdt</i>	suffering	—
<i>tatojó:</i>	[tato'xo:]	<i>vi</i>	immerse	—
<i>tatu:jú:</i>	[tatu'xu:]	<i>vi</i>	put down	—
<i>tatu'kxni'</i>	[taxuk'ni]	<i>vt</i>	break	—
<i>tatzá'pa'</i>	[tax'paxa]	<i>adj</i>	sewn	—
<i>tatzá'psni'</i>	[tax'paxni]	<i>adj</i>	piled up	—
<i>tatzi'tá</i>	[taxi'ta]	<i>vi</i>	wrinkle	—
<i>taxká:t</i>	[tax'ka:t]	<i>n</i>	type of bee	—

Continued.

Orthography	International	Part of speech	Gloss	Chicontla
	Phonetic Alphabet			
<i>taxpá'ta'</i>	[taʃ ^h paʃtə]	<i>n</i>	salsa	—
<i>te:'lh</i>	[te:'lʰ]	<i>adv</i>	strewn	—
<i>te:kú'</i>	[te:'ku]	<i>n</i>	spirit	—
<i>te:taxtú</i>	[te:taʃ ^h tu]	<i>vi</i>	come out	—
<i>teje:xká:n</i>	[teje:ʃ ^h ka:n]	<i>n</i>	water running	—
<i>textexe</i>	[teʃeteʃe]	<i>idph</i>	water sprouting	—
<i>tí'lhí:t</i>	[tí ^h ʰi:t]	<i>adv</i>	dense (hair)	—
<i>tí'lhí'lhí'</i>	[tí ^h ʰi ^h ʰi]	<i>adj</i>	dense (hair)	—
<i>tí'sí:t</i>	[tí ^h ʰsi:t]	<i>adv</i>	sparse (hair)	—
<i>tí'sí'sí'</i>	[tí ^h ʰsi:si]	<i>adj</i>	sparse (hair)	—
<i>tí'xí:t</i>	[tí ^h ʰxi:t]	<i>adv</i>	dense (hair)	—
<i>tí'xí'xí'</i>	[tí ^h ʰxi:xi]	<i>adj</i>	dense (hair)	—
<i>tí'xtí'x</i>	[tí ^h ʰxi:xi]	<i>adj</i>	squirting (liquid)	—
<i>tinta'kú:'j</i>	[tinta ^h ku:x]	<i>adv</i>	all day	—
<i>tojó:</i>	[to'xo:]	<i>vi</i>	be wet	—
<i>tú:pu'</i>	[^h tu:pu]	<i>n</i>	sprout	—
<i>tu:ta:'tín</i>	[tu:ta: ^h tin]	<i>n</i>	step sister	—
<i>tuta:tá'</i>	[tu:ta: ^h ta]	<i>n</i>	step father	—
<i>tzi'sít</i>	[tzi ^h ʰsi:t]	<i>n</i>	feathers	—
<i>tzi'tzi:ks</i>	[tzi ^h ʰtzi:ks]	<i>n</i>	type of plant	—
<i>wa'xá'xa'</i>	[wa ^h ʰxa:xa]	<i>n</i>	type of tree	—
<i>xke'jék</i>	[ʃke ^h ʰxek]	<i>n</i>	type of herb	<i>xke'jék</i>

References

- Alarcón Montero, Rafael. 2008. Indicios acústicos de las vocales rechinadas del totonaco [Acoustic properties of laryngealized vowels in Totonac]. In Esther Herrera Zendejas & Pedro Martín Butragueño (eds.), *Fonología instrumental: patrones fónicos y variación* [Instrumental phonology: Phonetic patterns and variation], 120–138. Mexico City: El Colegio de México.
- Allen, Micah, Davide Poggiali, Kirstie Whitaker, Tom R. Marshall & Rogier Kievit. 2018. Raincloud plots: A multi-platform tool for robust data visualization. *PeerJ Preprints* 6:e27137v1. <https://doi.org/10.7287/peerj.preprints.27137v1> (accessed October 2018).
- Aschmann, Herman P. 1946. Totonaco phonemes. *International Journal of American Linguistics* 12, 34–43.
- Baayen, R. Harald, Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412.
- Barreda, Santiago. 2013. phonTools: Functions for phonetics in R. R package (version 0.1-6). <http://www2.uaem.mx/r-mirror/web/packages/phonTools/index.html>.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Beck, David. 2004. *A grammatical sketch of Upper Necaxa Totonac*. Munich: Lincom Europa.
- Beck, David. 2006. The emergence of ejective fricatives in Upper Necaxa Totonac. *University of Alberta Working Papers in Linguistics* 1, 1–18.
- Beck, David. 2008. Ideophones, adverbs, and predicate qualification in Upper Necaxa Totonac. *International Journal of American Linguistics* 74, 1–46.

- Beck, David. 2011. *Upper Necaxa Totonac dictionary*. Berlin: Mouton de Gruyter.
- Beck, David. 2014. Totonacan languages. Presented by invitation at Workshop on the State of the Arts in Mesoamerican Languages, Max Planck Institute for Evolutionary Anthropology, Leipzig, 6–7 December.
- Boersma, Paul & David Weenink. 2013. Praat: Doing phonetics by computer (version 5.3.59). <http://www.praat.org/> (accessed 6 April 2013).
- Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters & Søren Wichmann. 2011. Totozoquean. *International Journal of American Linguistics* 77, 323–372.
- Flemming, Edward. 2006. A phonetically-based model of phonological vowel reduction. Ms., MIT.
- Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27, 765–768.
- Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1, 126–152.
- Gordon, Matthew. 1998. The phonetics and phonology of non-modal vowels: A cross-linguistic perspective. *Berkeley Linguistic Society* 24 (BLS 24), 93–105.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.
- Henderson, Eugenie. 1967. Vowel length and vowel quality in Khasi. *Bulletin of the School of Oriental and African Studies* 30, 564–588.
- Herrera Zendejas, Esther. 2009. *Formas sonoras: mapa fónico de las lenguas mexicanas* [Sound forms: A phonetic map of Mexican languages]. México: El Colegio de México.
- Holton, Gary. 2000. *The phonology and morphology of the Tanacross Athabaskan language*. Ph.D. dissertation, University of California, Santa Barbara.
- INEGI [Instituto Nacional de Estadística y Geografía]. 2010. Censo de Población y Vivienda de 2010 [Census of Population and Housing Units]. <http://www.inegi.org.mx/est/contenidos/proyectos/ccpv/cpv2010/Default.aspx> (accessed 21 December 2011).
- Kari, James M. 1990. *Ahtna Athabaskan dictionary*. Fairbanks, AL: Alaska Native Language Center.
- Kavitskaya, Darya. 2014. *Compensatory lengthening: Phonetics, phonology, diachrony*. New York: Routledge.
- Kirchner, Robert & Eleni Varelas. 2002. A cue-based approach to the phonotactics of Upper Necaxa Totonac. Presented at 7th Workshop on Structure and Constituency in the Languages of the Americas, University of Alberta, Edmonton, Alberta.
- Kirk, Paul, Jenny Ladefoged & Peter Ladefoged. 1993. Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. In Anthony Mattina & Timothy Montler (eds.), *American Indian linguistics and ethnography in honor of Laurence C. Thompson*, 435–450. Missoula, MT: University of Montana Press.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H.B. Christensen. 2016. *ImerTest*: Tests in linear mixed effects models. R package (version 2. 0–33). <https://CRAN.R-project.org/package=ImerTest> (accessed October 2018).
- Lam, Yvonne. 2009. The straw that broke the language's back: Language shift in the Upper Necaxa Valley of Mexico. *International Journal of the Sociology of Language* 195, 219–233.
- Lam, Yvonne. 2012. Oportunidad, ideología y la pérdida del totonaco del río Necaxa [Opportunity, ideology and the loss of Upper Necaxa Totonac]. In Paulette Levy & David Beck (eds.), *Las lenguas totonacas y tepehuas: Textos y otros materiales para su estudio* [The Totonac and Tepehua languages: Texts and other reference materials], 519–543. Mexico City: Universidad Nacional Autónoma de México.
- Lam, Yvonne & David Beck. 2008. Language loss and linguistic suicide: A case study from the Sierra Norte de Puebla. In Sarah Cummins, Brigit Janoski & Patricia A. Shaw (eds.), *All the things you are: A Festschrift for Jack Chambers*, 5–16. Toronto: Toronto Working Papers in Linguistics.
- Lehnert-LeHouillier, Heike. 2010. A cross-linguistic investigation of cues to vowel length perception. *Journal of Phonetics* 38, 472–482.
- Lindblom, Bjorn. 1963. Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America* 35, 1773–1781.

- MacKay, Carolyn J. & Frank R. Trechsel. 2013. A sketch of Pisaflores Tepehua Phonology 1. *International Journal of American Linguistics* 79, 189–218.
- McFarland, Teresa A. 2009. *The phonology and morphology of Filomeno Mata Totonac*. Ph.D. dissertation, University of California Berkeley.
- Muehlbauer, Jeffrey. 2012. Vowel spaces in Plains Cree. *Journal of the International Phonetic Association* 42, 91–105.
- Nearey, Terrance M. 1978. *Phonetic feature systems for vowels*. Bloomington, IN: Indiana University Linguistics Club.
- Nolan, Francis. 2003. Intonational equivalence: An experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS XV)*, Barcelona, 771–774.
- Puderbaugh, Rebekka. 2015. Contextual effects on the duration of ejective fricatives in Upper Necaxa Totonac. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII)*, Glasgow, 931–934.
- Puderbaugh, Rebekka. 2016. Acoustic characteristics of obstruents in Huehuetla Tepehua. *Journal of the Canadian Acoustical Association* 44(3), 140–141.
- Puderbaugh, Rebekka & Benjamin V. Tucker. 2013. Acoustic features of Upper Necaxa Totonac ejective fricatives. Presented at the 166th Meeting of the Acoustical Society of America, San Francisco, CA.
- R Core Team. 2017. R: A language and environment for statistical computing. <https://www.R-project.org/>.
- Román Lobato, Gabriela. 2008. La juntura fonológica en el totonaco de Ozelonacaxtla, Huehuetla, Puebla [Phonological junctures of Ozelonacaxtla Totonac]. Master's thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Shih, Chilin & Hsin Yi Dora Lu. 2015. Effects of talker-to-listener distance on tone. *Journal of Phonetics* 51, 6–35.
- Trechsel, Frank R. & Alice Faber. 1992. Acoustic properties of plain and laryngealized vowels in the Misantla dialect of Totonac. Ms., Universidad de las Américas-Puebla, Mexico.
- Tuttle, Siri G. 1998. *Metrical and tonal structures in Tanana Athabaskan*. Ph.D. dissertation, University of Washington.
- Tuttle, Siri G. 2005. Duration, intonation and prominence in Apache. In Sharon Hargus & Keren Rice (eds.), *Athabaskan prosody*, 319–344. Amsterdam: John Benjamins.
- Tuttle, Siri G., Olga Lovick & Isabel Núñez-Ortiz. 2011. Vowels of Upper Tanana Athabaskan. *Journal of the International Phonetic Association* 41, 283–312.
- Warner, Natasha, Allard Jongman, Joan Sereno & Rachel Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32, 251–276.
- Whalen, Douglas H. & Andrea G. Levitt. 1995. The universality of intrinsic f0 of vowels. *Journal of Phonetics* 23(3), 349–366.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91, 1707–1717.