

# Investigating the timing of tone and phonation in Santiago Laxopa Zapotec<sup>1</sup>

Mykel Loren Brinkerhoff

## 1 Introduction

Most of the work on the interaction of tone and phonation has been based on descriptions of southeast and far east Asian languages, which lead to strong claims being made about what is possible for the interaction between tone and phonation (Masica 1976, Thurgood 2002, Yip 2002, Enfield 2005, Michaud 2012, Brunelle & Kirby 2016). The main claim from many of these authors is that tone and phonation are codependent (i.e., certain tones only appear with certain phonations).

An example for this type of co-occurrence between tone and phonation is Mandarin's Tone 3 which is frequently associated with creaky voice (Hockett 1947).

## 2 The Laryngeal Complexity Hypothesis

The LARYNGEAL COMPLEXITY HYPOTHESIS (LCH) has its origins in the work from Silverman (1997) and Blankenship (1997, 2002). The basic premise of the LCH is that in languages that have both tone and phonation there needs to be an ordering between the two laryngeal gestures for tone and phonation for optimal perceptibility. This premise comes from the understanding that the same mechanism that is responsible for tone is also responsible for the production of phonation: the vocal folds and glottis. The rate that vocal folds vibrate is responsible for changing the fundamental frequency, perceived as pitch, which is grammaticalized as tone in lexical tone languages. The faster the vocal folds vibrate the higher the pitch and the slower the vocal folds vibrate the lower the pitch.

The vocal folds are also the primary articulator for phonation. Ladefoged (1971) and Gordon & Ladefoged (2001) treated phonation as a by-product of how open or closed the glottis is during vocal fold vibration. This is schematized by Figure 1. The more open the glottis is the breathier the sound to the point where the sound becomes the sound [h]. The more closed the glottis is the creakier the sound is to the point where the sound becomes [?].

---

<sup>1</sup>I am grateful to Fe Silva-Robles and Raúl Díaz Robles for sharing their time and language expertise. I am also grateful to Grant McGuire, Jaye Padgett, Rachel Walker, Maziar Toosarvandani, Ben Eischens, Kim Tan, and Zach Horton for their help and discussions during all stages of this project. This project branched off from a collaboration with Jack Duff and Maya Wax Cavallaro. Various parts of this project were previously shared in joint presentations with Jack Duff and Maya Wax Cavallaro.

This work was supported in part by the National Science Foundation under Grant No. 2019804, the Humanities Institute at UC Santa Cruz, and the Jacobs Research Funds.

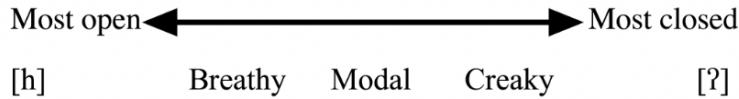


Figure 1: Simplified one-dimensional model for phonation. Based on Ladefoged (1971) and Gordon & Ladefoged (2001)

The LCH assumes that phonation and tone are both produced at the vocal folds and glottis. Because these same organs are responsible for these two different phenomena there is a mismatch in trying to produce both at the same time. It is assumed by the LCH that because of this issue there needs to be a strict ordering in the glottal gestures. This means that the tonal gesture needs to be produced either before or after the phonation gesture. The reason for this is that if the gestures were overlapped there will be a perturbation of the tone and the listeners will not be able to reliably differentiate what the tone is. The LCH assumes that there is a close link between production and perception. This assumption places the responsibility on making sure the acoustic cues are the most perceptually salient on the speaker and the speaker is responsible to produce both tone and phonation in such a way that the listener can clearly differentiate the different acoustic cues for both tone and phonation. These assumptions can best be represented by Figure 2.

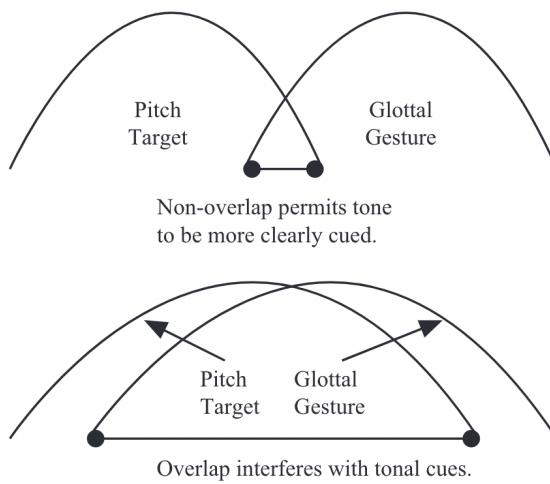


Figure 2: Representation taken from DiCanio (2012).

In Figure 2, which is taken from DiCanio (2012), the cue for tone is represented by the Pitch Target and the Glottal Gesture represents the gestures needed to produce phonation. When the Pitch Target is not co-articulated with the Glottal Gestures, then there is the greatest perceptual recoverability for the listener. This is argued for by the LCH as the tones are the most recoverable on modal vowels. This modal portion is then ordered or phased relative to the non-modal

portion. This ordering is then the responsibility of the speaker to accommodate for the listeners perceptibility. If, however, the Pitch Target and the Glottal Gesture were overlapping, as represented by lower half in Figure 2, then the cues for pitch and phonation would be overlapping and making it perceptually more difficult for

Some work has been done into investigating this in other languages, most notably DiCanio's (2012) investigation into glottals in Itunyoso Trique, which is also an Oto-Manguean language. DiCanio in his study found that when the magnitude of coarticulation for glottal consonants occurs on the vowels there is a strong correlation between the magnitude of overlap and the amount of perturbation in the f0 signal. If, however, the degree of overlap was minor then the acoustic signal had little to no perturbation. These results were found by consulting the spectral tilt of the vowels with the f0 measures and performing a generalized linear mixed effects model with speaker treated as a random effect.

Another study on Jalapa Mazatec (Garellek & Keating 2011) also investigated the interaction of tone and phonation. Jalapa Mazatec is a language with both contrastive tone and phonation and Garellek & Keating (2011) validated the claims made by the LCH, in that tone and phonation seemed to be ordered with each other when it comes to at least one of the phonation types.

In testing the claims made by the LCH, this paper investigates the interaction of tone and phonation in the Northern Zapotec language of Santiago Laxopa Zapotec. This language is ideal for testing the viability of the LCH because of its use of both contrastive tone and phonation. The rest of the paper will provide a description of the language in Section 3 with special emphasis on the tone and phonation systems. Following this discussion on Santiago Laxopa Zapotec, this paper discusses an acoustic analysis of phonation modelled on Esposito (2010), Garellek & Keating (2011), and DiCanio (2012) in Section 5.

### 3 Santiago Laxopa Zapotec

Santiago Laxopa Zapotec (SLZ; *Dilla'xhonh Laxup*) is a Northern Zapotec language spoken by approximately 1000 people in the municipality of Santiago Laxopa, Ixtlán District in the Sierra Norte of Oaxaca, Mexico (Adler & Morimoto 2016, Adler et al. 2018, Foley, Kalivoda & Toosarvandani 2018, Foley & Toosarvandani 2020). It is closely related to San Bartolomé Zoogocho Zapotec (Long & Cruz 2005, Sonnenschein 2005) and is mutually intelligible with this variety according to native speakers. As is common among Zapotecan languages, SLZ distinguishes between lenis and fortis consonants (e.g., Nellis & Hollenbach 1980, Jaeger 1983, Uchihara & Pérez Báez 2016) and has a fairly standard five-vowel inventory. These contrasts and inventories can be seen in Table 1 and Table 2.

In addition to the contrasts in both consonants and vowels, SLZ additionally has a phonation

Table 1: Consonant inventory for Santiago Laxopa Zapotec

		bilabial	alveolar	retroflex	alveo-palatal	palatal	velar	labio-velar	uvular
nasal	lenis		n						
	fortis	m:	n:						
stop	lenis	b	d				g	g <sup>w</sup>	
	fortis	p	t				k	k <sup>w</sup>	
fricative	lenis		z	z~t̪	ʒ	ç			χ~χ̪
	fortis		s	ʂ	ʃ				
affricate	lenis		ðz						
	fortis		ts		ʈʃ				
lateral	lenis		l~r̪						
	fortis		l:						
trill			r						
approximate					j			w	

Table 2: Vowels inventory in Santiago Laxopa Zapotec.

	front	central	back
high	i		u
mid	e		o
low		a	

and tonal contrast. I will first talk about the tonal contrasts in Section 3.1 followed by the tonal phonation in Section 3.2.

### 3.1 Tone in SLZ

As is common among Oto-Manguean languages, SLZ is a tonal language (Suárez 1983, Campbell, Kaufman & Smith-Stark 1986, Silverman 1997, E. W. Campbell 2017a,b). In SLZ there are five surface tones which appear on a syllable. Following other papers on SLZ, I am representing the tone using Pike's numbers (e.g. Sichel & Toosarvandani 2020). Pike's numbers are commonly used in describing Mesoamerican tone systems and has its origins in Pike's (1948) work with Mixtec. A good summary of this system and its comparisons to other tone marking systems is given in Wee (2019). Pike's numbers represent the highest tone with the number 1 and then each subsequent tonal level is represented with the next number until all levels have been assigned. In the case of a language with three distinct levels the number 1 represents the highest tone, the number 2 represents the mid tone, and finally the number 3 represents the lowest tone. Additionally,

these different numbers can be combined to represent the different contour tones that exist in the language, for example 21 representing a tone that starts at the mid level and then increases its pitch to high.

Table 3: Examples of the five tonal patterns observed in the Santiago Laxopa Zapotec words.

High	$a^1$	<i>xha</i>	[ za <sup>1</sup> ]	'clothing.POSS'
Mid	$a^2$	<i>lhill</i>	[ ri <sup>2</sup> ]	'house.POSS'
Low	$a^3$	<i>yu'</i>	[ ju <sup>3</sup> ]	'earth'
Rising	$a^{21}$	<i>yu'u</i>	[ ju <sup>2</sup> u <sup>21</sup> ]	'quicklime (Sp. cal)'
Falling	$a^{13}$	<i>yu'u</i>	[ ju <sup>2</sup> u <sup>13</sup> ]	'house'

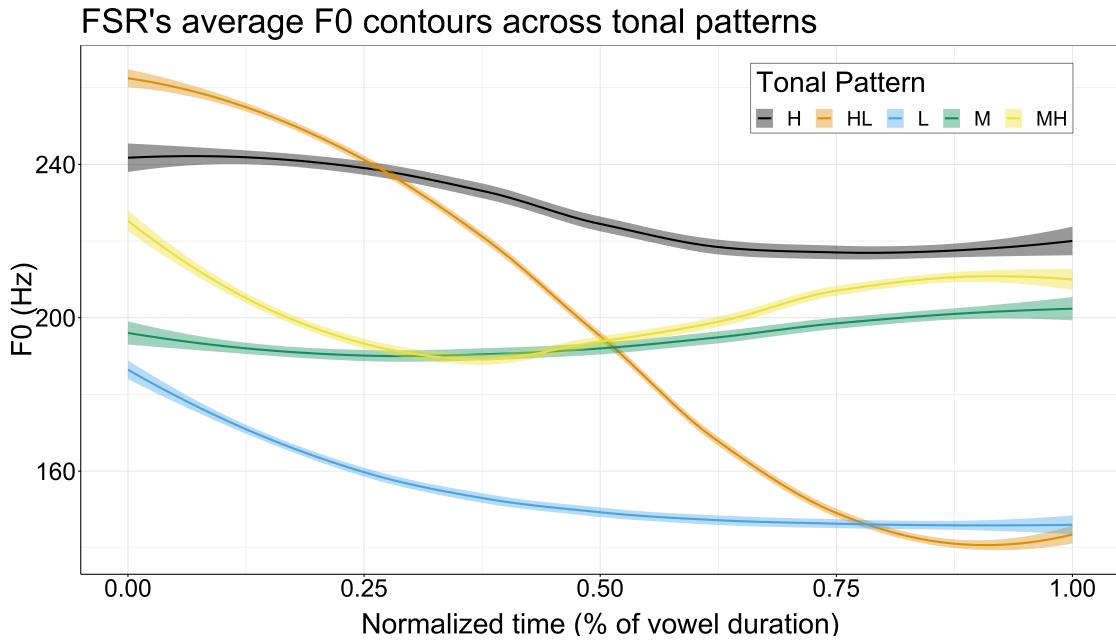


Figure 3: Tonal contrasts for FSR averaged and time normalized. Each line in this graph represents the average of approximately 10 syllables for each tonal pattern.

Following discussion from Brinkerhoff, Duff & Wax Cavallaro (2022) these tones appear to be limited in their distribution. It is true that all five patterns can surface on a syllable but there is a restriction in what tonal patterns are allowed to surface on words that are larger than bimoraic. The patterns that we observe on bimoraic nominals are: HL, MH, and LL.

### 3.2 Phonation in SLZ

Among Zapotecan languages it is quite common for languages to make use of contrastive phonation (e.g., Avelino 2004, Long & Cruz 2005, Avelino 2010, López Nicolás 2016, Chávez-Péon 2010).

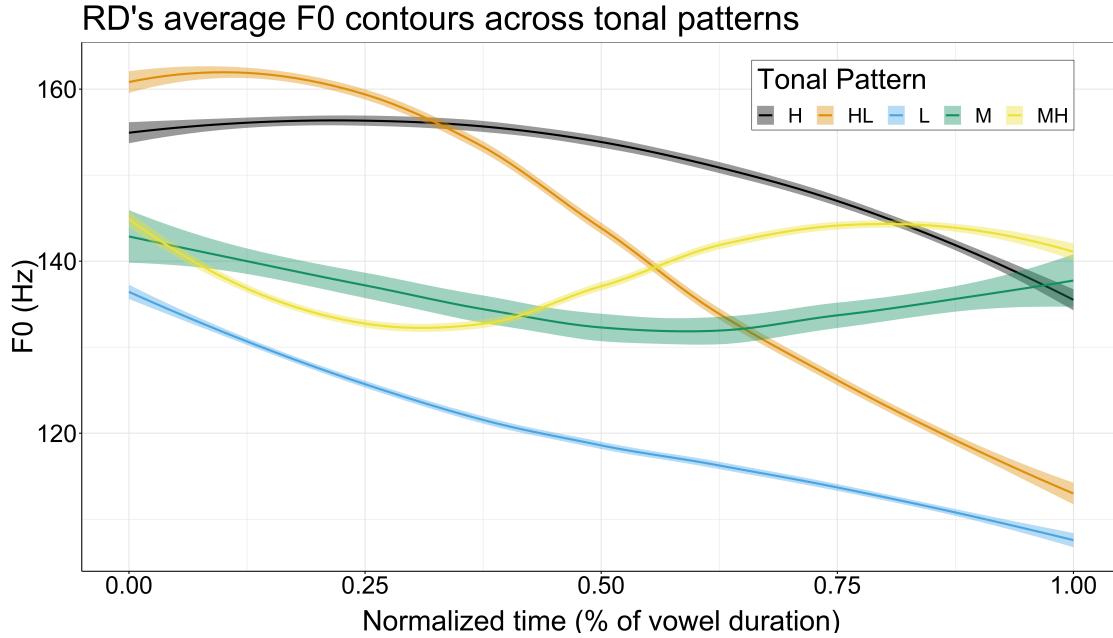


Figure 4: Tonal contrasts for RD averaged and time normalized. Each line in this graph represents the average of approximately 10 syllables for each tonal pattern.

SLZ, in addition to the five vowel qualities, has four contrastive phonation types which are: modal, breathy, checked, and laryngealized. These contrasts are exemplified in the minimal quadruple in (1). In representing the checked and laryngealized vowels, I follow the same procedure as other authors (e.g., Avelino (2010) and Uchihara (2016)) in representing the ‘glottal stop’ element as a superscripted glottal stop in the IPA transcription (i.e., [a']). This is primarily done as a way of standardizing the variable pronunciation of the glottal element in Zapotec ranging from a full glottal stop (i.e., [a?]) to creaky voice on a portion of the vowel (i.e., [aa]).

(1) Four-way minimal phonation contrast

- a. *ya* [ja<sup>3</sup>] ‘bell’
- b. *yah* [ja<sup>3</sup>] ‘metal/rifle’
- c. *ya'* [ja'<sup>3</sup>] ‘pound’
- d. *ya'a* [ja'a<sup>3</sup>] ‘market’

Breathy phonation on vowels is characterized by a raspy quality throughout the whole vowel or a portion toward the end of the vowel, see Figure 5.

Checked vowels on the other hand are characterized by an abrupt glottal closure which cuts the vowel short. This phonation is sometimes only realized as a very short period of creakiness at the end of the vowel, see Figure 6.

Laryngealized vowels are quite common in Zapotecan languages and have received a wide

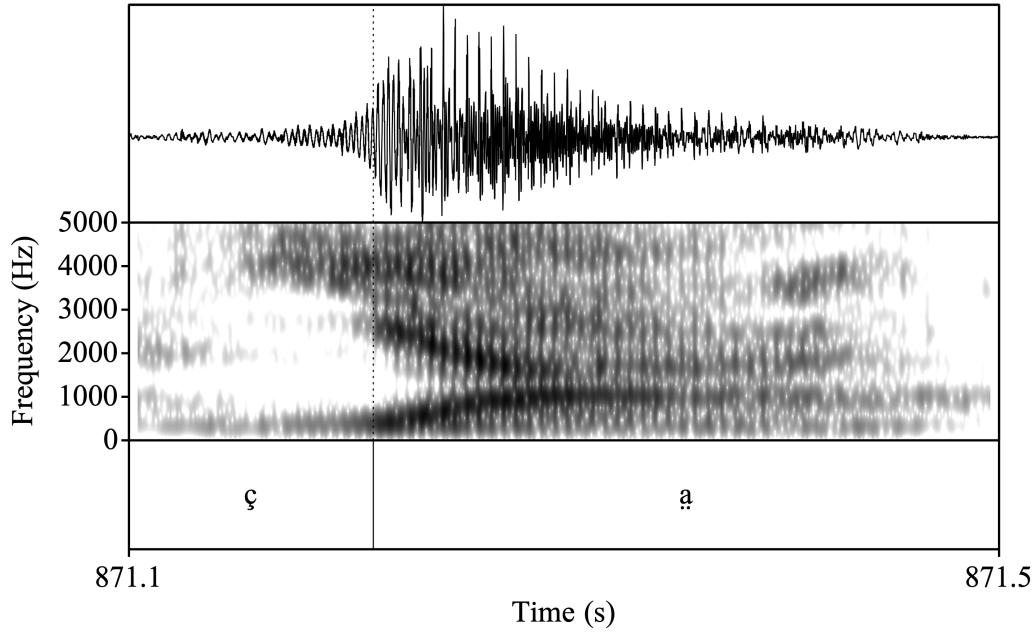


Figure 5: Breathy vowel in the word *yah* ‘metal/rifle’

[INSERT YA’ SPECTROGRAM AND WAVEFORM]

Figure 6: Breathy vowel in the word *ya'* ‘pound’

number of different names. Previous descriptions have used terms such as broken, rearticulated, interrupted, and creaky (Long & Cruz 2005, Avelino 2004, 2010, Sonnenschein 2005, Adler & Morimoto 2016). In order to avoid confusion, I will use the term laryngealized following Avelino (2010). In addition to a wide number of different names these vowels also exhibit a wide range of allophones.

Avelino (2010) found in the closely related Yalálag Zapotec that among his consultants there were at least four different pronunciations as seen in Table 4. In SLZ, each of the consulted

Table 4: Layngealized Vowels in Yalálag Zapotec

/V'V/	[V?V]
	[VV̚V]
	[VV̚:V̚]
	[VV̚V̚]

language experts would produce this vowel differently. One consultant would do rearticulation, where there is a full glottal stop in the middle of the vowel, or creaky voice. This alternation

seemed to be in free variation but there was a greater tendency to creak in low toned words, such as *xa'ag* [ʂə:g] ‘topil’<sup>2</sup>, see Figure 7 for a comparison between this consultant’s pronunciation of the laryngealized vowels.

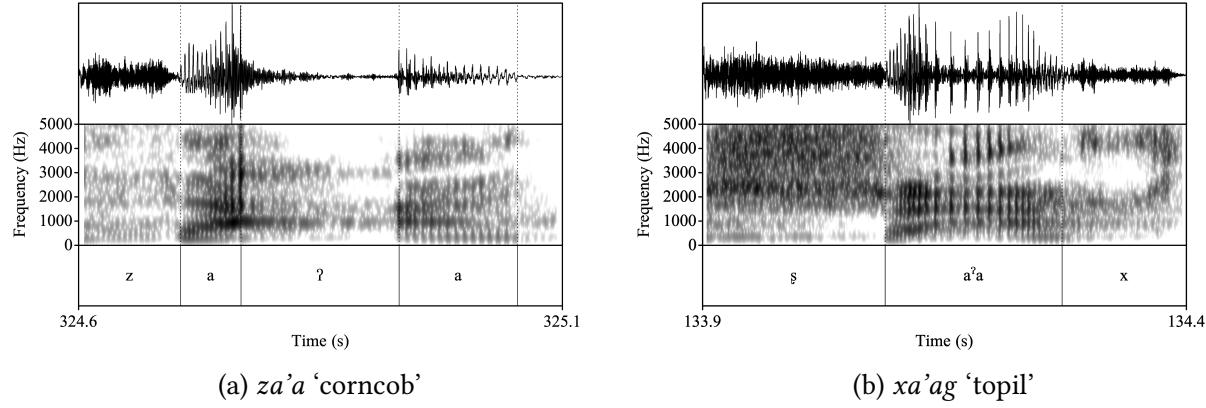


Figure 7: Comparison of FSR’s laryngealized vowels in *za'a* ‘corncob’ and *xa'ag* ‘topil’

The other consultant only ever produces creaky voice for these vowels regardless of the tone with the word. During one of the elicitation sessions, we conducted a perceptual check that these were in fact the same vowels and both consultants reliably identified the words and produced laryngealized vowels according to their own idiosyncrasies. However, a more detailed perception study is beyond the scope of this paper.

[INSERT SIDE BY SIDE SPECTROGRAMS]

Figure 8: Comparison of RD’s laryngealized vowels in *za'a* ‘corncob’ and *xa'ag* ‘topil’

## 4 Interaction of Tone and Phonation

Most previous work on the interaction of tone has been focused on the languages of East and Southeast Asia (e.g., Masica 1976, Thurgood 2002, Yip 2002, Enfield 2005, Michaud 2012, Brunelle & Kirby 2016). What has been found in these descriptions is that certain tones and phonations are codependent (i.e., only occur with each other). For example Smalley (1976) and Ratliff (1992) both describe White Hmong’s -g tone as being a mid-low tone with breathy phonation and Mandarin’s tone 3 is often associated with creaky phonation (Hockett 1947). Brunelle (2009) found that creaky phonation plays an important role in the production of certain tones. Additionally, work on S’gaw Karen has found that two tones are only differentiated by the presence of some form of amodal phonation (Boehm p.c.).

<sup>2</sup>A topil is a type of government office in traditional Oaxacan communities somewhat akin to a sheriff.

However, there has been some observations—especially in Mesoamerica—that tone and phonation can co-vary (e.g., Silverman 1997, Garellek & Keating 2011). This means that tone can independently occur with any phonation type. This has also been extensively described in multiple Zapotecan languages (e.g., Avelino 2004, 2010, Chávez-Péón 2010, E. Campbell 2011, Villard 2015, López Nicolás 2016)

Chávez-Péón (2010) has a detailed description of the tone and phonation interactions in San Lucas Quiaviní Zapotec (SLQZ), a central valley variety of Zapotec. The distribution of tone and phonation is found in Table 5. We see that in SLQZ, that both low and falling tone have the full range of possible combinations. However, we see gaps in the high tone for breathy and rising tone can only occur with modal phonation.

Table 5: SLQZ tone and phonation interactions (Chávez-Péón 2010).

	<b>Modal</b>	<b>Breathy</b>	<b>Creaky</b>	<b>Interrupted</b>
High	✓	–	✓	✓
Low	✓	✓	✓	✓
Falling	✓	✓	✓	✓
Rising	✓	–	–	–

Based on elicitation data collected from 2020-2022, SLZ has a more expansive distribution of tone and phonation when compared to SLQZ but seems to be very similar to other Northern Zapotec varieties (e.g., Avelino 2004). The distribution of SLZ tonal and phonation interactions are given in Table 6.

Table 6: Distribution of tone and phonation in SLZ

	<b>Modal</b>	<b>Breathy</b>	<b>Checked</b>	<b>Laryngealized</b>
High	✓	–	✓	✓
Mid	✓	✓	✓	✓
Low	✓	✓	✓	✓
High-Low	✓	✓	✓	✓
Mid-High	✓	✓	–	✓

One of the striking things in this is the lack of high tone with breathy phonation. This gap is interesting because of the long time association of high pitch with breathiness (a good overview of this association and other phonation types—is found in Esling et al. 2019). This gap of breathy phonation and high tone is quite common across the Zapotecan languages (Campbell p.c.). In the case of breathy phonation in SLQZ, Uchihara (2016) offers some convincing evidence that the phonation originated in syllables with low tone and then spread to other tones via analogy.

## 5 Acoustic investigation into the phonation

One of the primary ways to measure and investigate phonation is using spectral measures. The most frequent type of measurement is spectral-tilt. Spectral-tilt measurements are conducted by comparing the relative amplitude of different harmonics in the acoustic signals, which has primarily been the difference in the relative amplitude of the first and second harmonics. Other measurements make use of higher harmonics which are closest to the different formants, see Figure 9. These spectral-tilt measurements have been found to be particularly useful in languages such as Green Hmong (Huffman 1987, Andruski & Ratliff 2000) and Jalapa Mazatec (Silverman et al. 1995, Blankenship 1997).

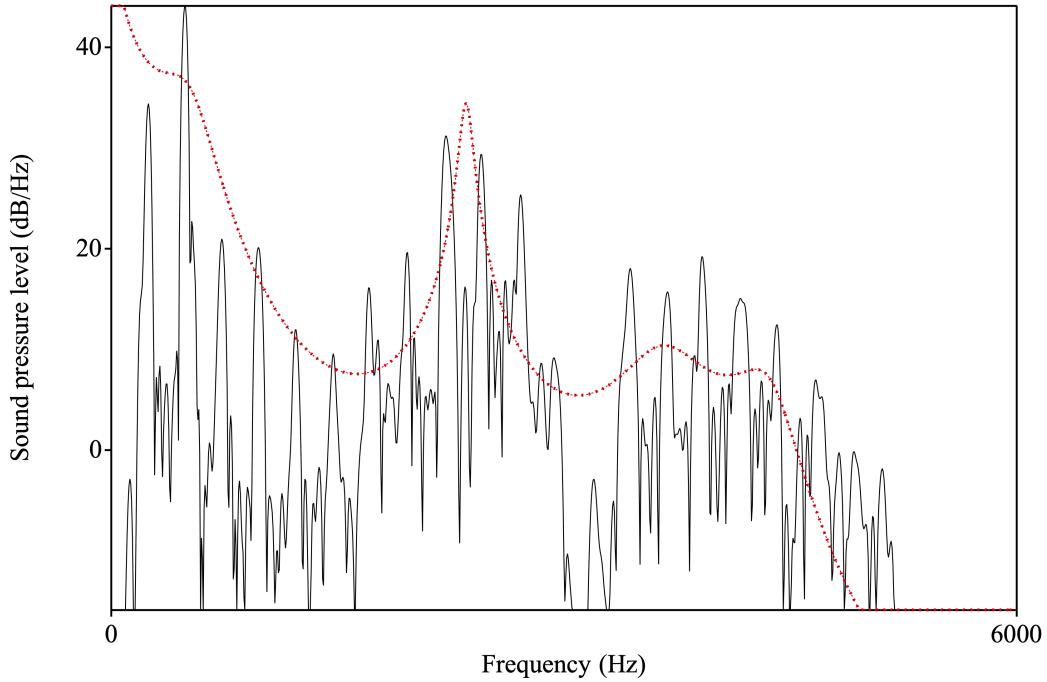


Figure 9: Spectral slice with LPC smoothed line overlaid for the vowel [e]. The harmonics in the spectral slice are represented by each of the dark peaks. The leftmost black solid line peak is the first harmonic ( $H_1$ ) and each subsequent peak represents the next highest harmonic ( $H_2$  through  $H_n$ ). The red dotted line represents an LPC smoothed line which identifies the formants by the peaks. Each of the harmonics that are closest to the formant peak is identified as  $A_1$  through  $A_n$ .

When conducting spectral-tilt measurements, there are two types of measurements that can be made: corrected and uncorrected. The status of corrected or uncorrected refers to whether or not the influence of formants are taken into account (Garellek 2019 provides a good overview of the differences). Most previous studies have not used corrected measure but have been focused on a single vowel /a/ because it minimizes the effects of the first formant (Esposito 2010).

Corrected measures take the formants into account during calculation and minimizes their influences. Fortunately, most software that is used for calculating the acoustic measurements such as VoiceSauce (Shue, Keating & Vicenik 2009) and PraatSauce (Kirby 2022) produces both corrected and uncorrected values.

A second type of spectral measure that one often encounters are harmonic-to-noise ratios. Harmonic-to-noise ratios refer to the difference in amplitude between the harmonic and inharmonic components of the source spectrum, as measured in the cepstral domain (de Krom 1993). Frequently these harmonics-to-noise ratios take the form of a Cepstral Peak Prominence (CPP) measurement (Hillenbrand, Cleveland & Erickson 1994).

In determining whether or not a given measure corresponds to a given phonation several patterns have been described and validated by many authors. A summary of these findings are given in Garellek (2019) where it is noted that when the values of the spectral-tilt measurement are higher than that of the modal's spectral-tilt measurement and the CPP value is lower than the modal's CPP value this more than likely indicates a vowel with breathy phonation. If, however, the spectral-tilt measurements are lower than the modal's spectral-tilt measurements and the CPP value is lower than the modal's CPP value this more than likely indicates a vowel with creaky phonation. This means that if the acoustic measurements for the SLZ phonation types match these findings we can be fairly confident that we are dealing with breathy and creaky phonation.

The rest of this section will describe the methods, results, and discussion of an acoustic analysis into SLZ's phonation.

## 5.1 Methods

Due to the impact of the COVID-19 pandemic, only two native language speakers of SLZ were able to take part in this study (one male; one female). Both speakers live in Santa Cruz, CA and data collection was done remotely using Zencastr<sup>3</sup>, a professional podcasting website, (44.1kHz, 16-bit) or in-person outside in a well ventilated location, using a Zoom H4n handheld recorder (44.1kHz, 16-bit). Participants were recorded saying approximately 100 words in the carrier sentence *shnia' X chone las* 'I say X three times'. This phrase was repeated three times.

After the elicitations were completed the audio was uploaded into ELAN (Wittenburg et al. 2006) for initial segmentation into sentences. This was then followed by segmenting of the vowel portion in Praat (Boersma & Weenink 2021). These segments were then inputted to VoiceSauce (Shue, Keating & Vicenik 2009) where each vowel was resampled at 16kHz for acoustic measurement. The acoustic measurements of VoiceSauce were then analyzed in R (R Core Team 2021). In order to investigate the differences in timing, each vowel was normalized for time.

---

<sup>3</sup><https://zencastr.com/>

## 5.2 Spectral-tilt results

The results of the spectral-tilt measurements were for the most part comparable to the findings by Esposito (2010) for Central Valley Zapotec. As noted by Esposito (2010) there is a difference in which acoustic measurements were most informative for the speakers based on their sex. Similar to Esposito, the female speaker was best characterized with the H1-H2 measurements and the male speaker was best characterized with H1-A3. For the female speaker this was true for all phonation types except for breathy voice which measured similarly to the other non-modal vowels. This is not too unsurprising considering the fact that the female speaker would more often produce a full glottal fricative in these environments than a breathy vowel. This can be seen in Figure 10 where the modal vowel is represented by the solid black line. All other lines which represent breathy, checked, and laryngealized are all lower than that of the modal.

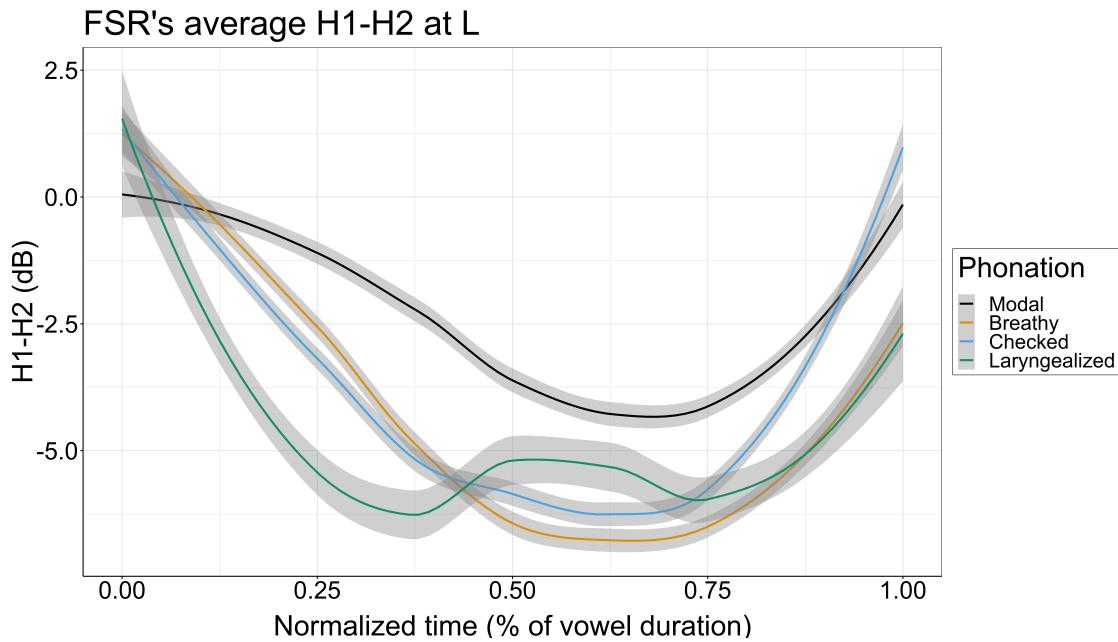


Figure 10: FSR's H1-h2 values.

As mentioned previously, one of our speakers would regularly produce the laryngealized vowels with rearticulation, meaning that they produced the vowel with glottalization, with either creakiness or a full glottal closure, in the middle of the vowel. This is in contrast with the checked vowels which have a short period of creakiness or a full glottal closure at the end of the vowel. This means that our spectral-tilt measurements should reflect a difference in time with creaky voice appearing earlier in the laryngealized vowels than checked vowels. In Figure 11, the spectral-tilt values for laryngealized and checked vowels are plotted along a time normalized scale. The important thing to be aware of is the difference in where the lowest part of the plot is located. For checked vowels, represented by the solid lines, the lowest part of the spectral-tilt

measurement is located in the last quarter of the vowel. This dip, however, occurs earlier in the vowel for the laryngealized vowels closer halfway to 75% of the vowel.

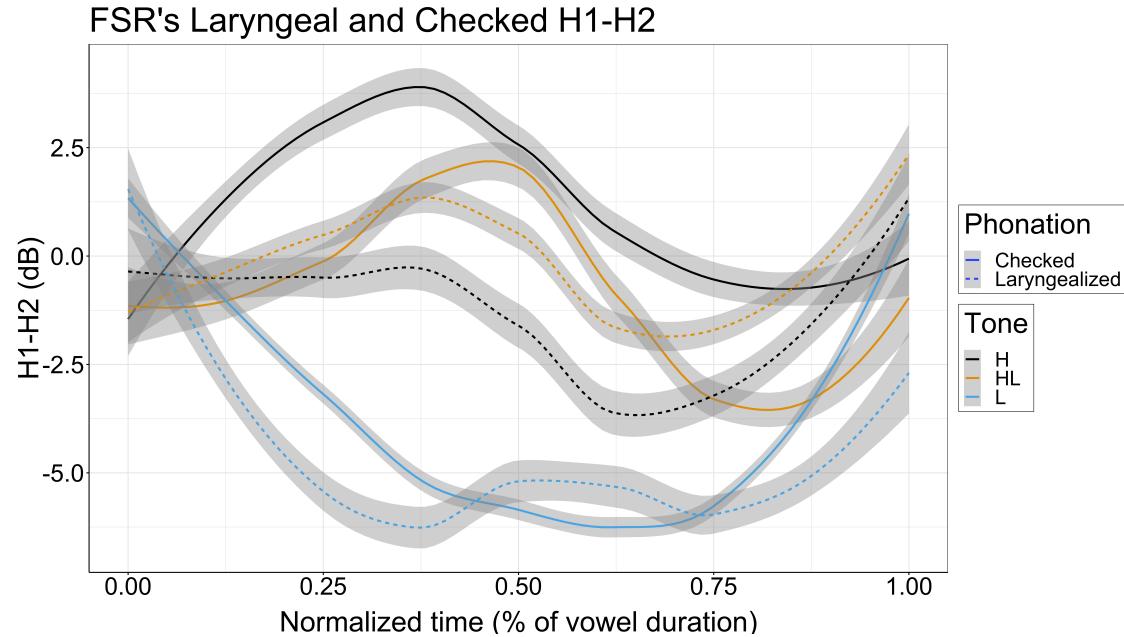


Figure 11: FSR's H1-h2 values for checked and laryngealized.

As mentioned previously, the results of the spectral analysis for the male SLZ speaker were inconclusive in terms of H1-H2. As seen in Figure 12 all of the different phonation types are essentially identical with less than 1 dB difference between the different phonation types. The H1-A3 spectral measures, however, show the expect patterns, as seen in Figure 13.

### 5.3 Statistical results

In order to determine whether or not the gestures for pitch and phonation are overlapping a mixed-effects linear regression analysis was performed with a Satterthwaite's method for t-test analysis used to derive the p-values. Following Garellek & Keating (2011), the results for both speaker were combined together and then divided into thirds based on the normalized times. This allows for analysis of each third of the vowel to determine if the phonation and pitch gestures are overlapping. To this effect, the f0 measurements where treated as the dependent variable with speaker treated as random. The results of the analysis showed that the phonation categories, spectral-tilt measurements, and CPP all had an effect on f0 in different portions of the vowel. Specifically the portions of the analysis showed that did not show any effects where the first third of the vowel for CPP and the last third for the laryngealized category. The spectral-tilt measurements of H1-H2 and H1-A3 had an effect on f0 in all portions of the vowel.

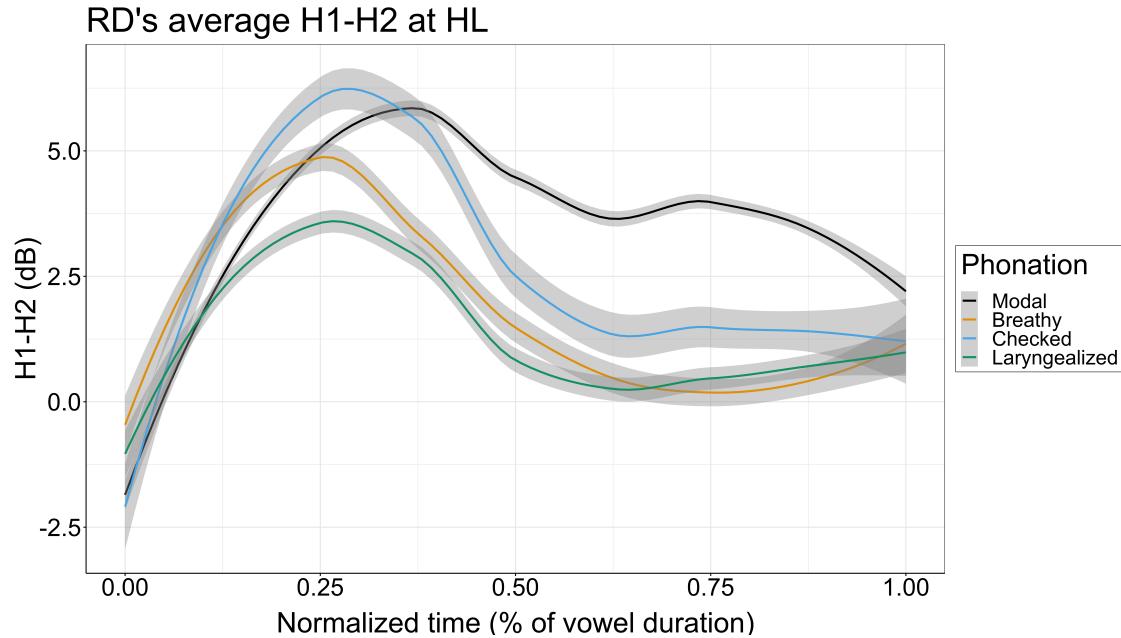


Figure 12: RD's H1-H2 values in HL toned syllables.

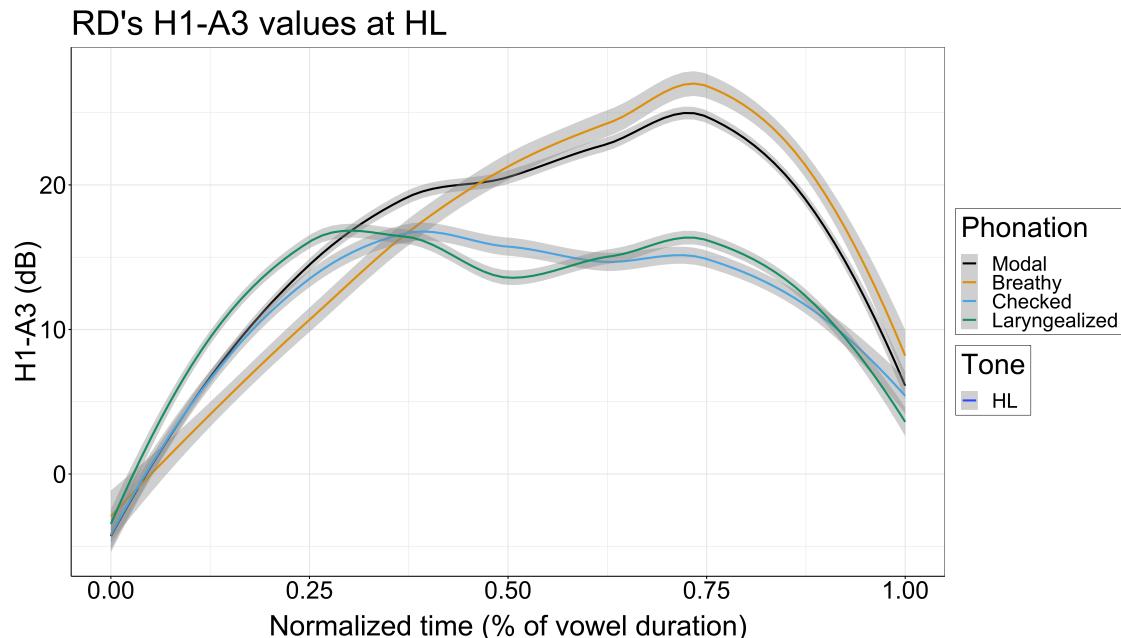


Figure 13: RD's total H1-A3 values in HL toned syllables.

## 6 Discussion of the results

Based on the results of the statistical analysis, SLZ appears to have a large degree of overlap in the tonal and phonation gestures in contrast to the strict ordering that is predicted to occur according to the LCH. The LCH predicts that languages that have contrastive tone and phonation should

Table 7: Results of the mixed-effects linear regression analysis on the first third of the vowel.

	Estimate	Std. Error	df	t value	p-value
Breathy	-23.41	0.8128	23710	-28.799	<0.001
Checked	-9.123	0.6933	23710	-13.159	<0.001
Laryngealized	9.712	0.6733	23710	14.424	<0.001
H1-H2	2.163	0.06640	23710	32.572	<0.001
H1-A3	-0.6678	0.03436	23710	-19.434	<0.001
CPP	0.04144	0.03289	23710	1.260	0.208

Table 8: Results of the mixed-effects linear regression analysis on the second third of the vowel.

	Estimate	Std. Error	df	t value	p-value
Breathy	-9.634	0.7874	23540	-12.235	<0.001
Checked	-3.509	0.6564	23540	-5.346	<0.001
Laryngealized	-2.261	0.6437	23540	-3.513	<0.001
H1-H2	3.371	0.04283	23540	78.707	<0.001
H1-A3	-0.4599	0.02592	23540	-17.741	<0.001
CPP	-0.9082	0.04806	23540	-18.897	<0.001

Table 9: Results of the mixed-effects linear regression analysis on the last third of the vowel.

	Estimate	Std. Error	df	t value	p-value
Breathy	-5.247	0.8224	24490	-6.380	<0.001
Checked	-4.494	0.6898	24490	-6.514	<0.001
Laryngealized	1.013	0.6670	24490	1.519	0.129
H1-H2	3.079	0.04900	24490	62.845	<0.001
H1-A3	-0.6611	0.02372	24490	-27.867	<0.001
CPP	0.5377	0.02958	24490	18.176	0.208

not have any overlap between tone and phonation. The rational for this comes from the one-dimensional nature of the larynx proposed by Ladefoged (1971). As previously mentioned, this model assumes that the vocal folds are the primary force in determining both tone and phonation. The assumption that tone and phonation are controlled by only the vocal folds is actually inaccurate. Recent work, summarized in Esling et al. (2019), actually shows that the whole larynx is involved with the production of the different phonation types. For example, based on laryngoscopy, creaky voice shows that in addition to the vocal folds, the thyroarytenoid muscle contract from the borders of the arytenoids to the front of the epilaryngeal tube. In addition, the ventricular folds are often also engaged in creaky voice which has the effect of increasing the vibrating mass in the larynx, adds damping, encourages irregular vibration, and perturbs the

mucosal wave from crossing the vocal folds (Moisik & Esling 2014).

Based on this and other work Esling et al. (2019) proposes a new model for accounting for the complex interactions in the larynx. This model termed the LARYNGEAL ARTICULATOR MODEL assumes that instead of a one-dimensional model you instead have a network of nodes that act in synergistic and anti-synergistic manners to each other. In a sense these nodes either work in a cooperative way strengthening or supplementing the workings of each other or the nodes are antagonistic and do not work together. These different nodes and their abbreviations are given in Table 10.

Table 10: A list of the different nodes and their abbreviations in the Laryngeal Articulator Model.

States/Nodes	Physiological description
vfo	vocal folds open (abducted)
vfc	vocal folds closed (adducted/prephonation)
epc	epilaryngeal constriction
epv	epilaryngeal vibration
tfr	tongue fronting
tre	tongue retraction
tra	tongue raising
tdb	tongue double bunching
↑lx	raised larynx
↓lx	lowered larynx
Hf0	increased vocal fold tension, less vibrating mass (high f0)
Lf0	decreased vocal fold tension, more vibrating mass (lower f0)

These twelve nodes not only represent the interactions of the larynx but also represents actual physiological representations. This means that any given node represents what is occurring with a given part of the larynx. For example, the node ‘epc’ represents any epilaryngeal constriction when activated. Now these nodes are not just independent entities but interact in complex ways with other nodes. These interactions are best captured as a network or web of nodes as seen in Figure 14.

Figure 14, shows which nodes interact in synergistic and anti-synergistic ways. This network is divided into three parts that represent the nodes which are responsible for vowel quality, phonatory quality, and tonal quality. For example, the node ‘epc’ which is one of the nodes responsible for phonatory quality (i.e., responsible for producing phonation) has synergistic relationships with the nodes representing larynx raising ( $\uparrow$ lx), vocal fold closure (vfc), tongue bunching(tdb), tongue retraction (tre), and for decreasing the vocal fold mass leading to a lowering of f0 (Lf0). In contrast this same node has anti-synergistic relationships with tongue fronting (tfr), larynx lowering ( $\downarrow$ lx), vocal fold opening (vfo), and increasing vocal fold tension which leads to less vi-

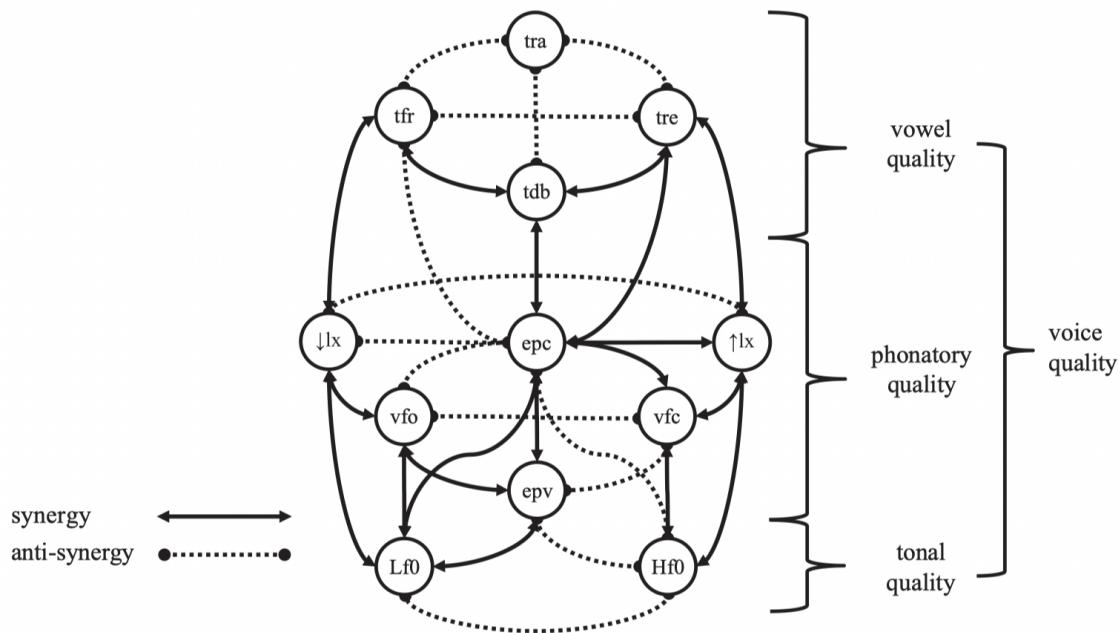


Figure 14: Network of synergistic and anti-synergistic nodes in the Laryngeal Articulator Model. Taken from Esling et al. (2019).

brating mass and an increase to f0 (Hf0). This means that when there is epilaryngeal constriction there is an increased chance for the synergistic relationships to also co-occur. However, this does not mean that the synergistic nodes must co-occur when the node is activated.

## 7 Conclusion

## References

- Adler, Jeff, Steven Foley, Jed Pizarro-Guevara, Kelsey Sasaki & Maziar Toosarvandani. 2018. The derivation of verb initiality in Santiago Laxopa Zapotec. In Jason Merchant, Line Mikkelsen, Deniz Rudin & Kelsey Sasaki (eds.), *A reasonable way to proceed: Essays in honor of Jim McCloskey*, 31–49. Santa Cruz, Berkeley, Chicago: University of California.
- Adler, Jeffrey M. & Maho Morimoto. 2016. Acoustics of phonation types and tones in Santiago Laxopa Zapotec. *The Journal of the Acoustical Society of America* 140(4). 3109–3109. <https://doi.org/10.1121/1.4969713>.
- Andruski, Jean E. & Martha Ratliff. 2000. Phonation types in production of phonological tone: the case of Green Mong. *Journal of the International Phonetic Association* 30(1-2). 37–61. <https://doi.org/10.1017/S0025100300006654>.

- Avelino, Heriberto. 2004. *Topics in Yalálag Zapotec, with particular reference to its phonetic structures*. Los Angeles, CA: University of California, Los Angeles dissertation. 315 pp.
- Avelino, Heriberto. 2010. Acoustic and Electroglossographic Analyses of Nonpathological, Non-modal Phonation. *Journal of Voice* 24(3). 270–280. <https://doi.org/10.1016/j.jvoice.2008.10.002>.
- Blankenship, Barbara. 1997. *The time course of breathiness and laryngealization in vowels*. Los Angeles, CA: University of California, Los Angeles dissertation.
- Blankenship, Barbara. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30(2). 163–191. <https://doi.org/10.1006/jpho.2001.0155>.
- Boersma, Paul & David Weenink. 2021. *Praat: Doing Phonetics by Computer [Computer Program]*. <http://www.praat.org/> (3 December, 2021).
- Brinkerhoff, Mykel Loren, John Duff & Maya Wax Cavallaro. 2022. Tonal patterns and their restrictions in Santiago Laxopa Zapotec.
- Brunelle, Marc. 2009. Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics* 37(1). 79–96. <https://doi.org/10.1016/j.wocn.2008.09.003>.
- Brunelle, Marc & James Kirby. 2016. Tone and Phonation in Southeast Asian Languages. *Language and Linguistics Compass* 10(4). 191–207. <https://doi.org/10.1111/lnc3.12182>.
- Campbell, Eric. 2011. Zenzontepec Chatino aspect morphology and Zapotecan verb classes. *International Journal of American Linguistics* 77(2). 219–246.
- Campbell, Eric W. 2017a. Otomanguean historical linguistics: Exploring the subgroups. *Language and Linguistics Compass* 11(7). e12244. <https://doi.org/10.1111/lnc3.12244>.
- Campbell, Eric W. 2017b. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11(4). e12240. <https://doi.org/10.1111/lnc3.12240>.
- Campbell, Lyle, Terrence Kaufman & Thomas Smith-Stark. 1986. Meso-America as a linguistic area. *Language* 62(3). 530–570.
- Chávez-Péón, Mario E. 2010. *The interaction of metrical structure, tone, and phonation types in Quiavini Zapotec*. University of British Columbia dissertation. <https://doi.org/10.14288/1.0071253>.
- DiCanio, Christian T. 2012. Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics* 40(1). 162–176. <https://doi.org/10.1016/j.wocn.2011.10.006>.
- Enfield, Nick J. 2005. Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology* 34. 181–206.

- Esling, John H., Scott R. Moisik, Allison Benner & Lise Crevier-Buchman. 2019. *Voice Quality: The Laryngeal Articulator Model*. 1st edn. (Cambridge Studies in Linguistics 162). Cambridge University Press. <https://doi.org/10.1017/9781108696555>.
- Esposito, Christina M. 2010. Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association* 40(2). 181–198. <https://doi.org/10.1017/S0025100310000046>.
- Foley, Steven, Nick Kalivoda & Maziar Toosarvandani. 2018. Forbidden clitic clusters in Zapotec. In Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit & Michael Tabatowski (eds.), *Proceedings of the Fifty-third Annual Meeting of the Chicago Linguistic Society*, 87–102.
- Foley, Steven & Maziar Toosarvandani. 2020. Extending the Person-Case Constraint to Gender: Agreement, Locality, and the Syntax of Pronouns. *Linguistic Inquiry*. 1–40. [https://doi.org/10.1162/ling\\_a\\_00395](https://doi.org/10.1162/ling_a_00395).
- Garellek, Marc. 2019. The phonetics of voice. In William F. Katz & Peter F. Assmann (eds.), *The Routledge handbook of phonetics* (Routledge Handbooks in Linguistics), 75–106. Abingdon, Oxon ; New York, NY: Routledge.
- Garellek, Marc & Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41(2). 185–205. <https://doi.org/10.1017/S0025100311000193>.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29(4). 383–406. <https://doi.org/10.1006/jpho.2001.0147>.
- Hillenbrand, James, Ronald A. Cleveland & Robert L. Erickson. 1994. Acoustic Correlates of Breathy Vocal Quality. *Journal of Speech, Language, and Hearing Research* 37(4). 769–778. <https://doi.org/10.1044/jshr.3704.769>.
- Hockett, Charles F. 1947. Peiping Phonology. *Journal of the American Oriental Society* 67(4). 253–267. <https://doi.org/10.2307/596062>.
- Huffman, Marie K. 1987. Measures of phonation type in Hmong. *The Journal of the Acoustical Society of America* 81(2). 495–504. <https://doi.org/10.1121/1.394915>.
- Jaeger, Jeri J. 1983. The fortis/lenis question: evidence from Zapotec and Jawoñ. *Journal of Phonetics* 11(2). 177–189. [https://doi.org/10.1016/S0095-4470\(19\)30814-9](https://doi.org/10.1016/S0095-4470(19)30814-9).
- Kirby, James. 2022. *PraatSauce*.
- de Krom, Guus. 1993. A Cepstrum-Based Technique for Determining a Harmonics-to-Noise Ratio in Speech Signals. *Journal of Speech, Language, and Hearing Research* 36(2). 254–266. <https://doi.org/10.1044/jshr.3602.254>.
- Ladefoged, Peter. 1971. *Preliminaries to linguistic phonetics*. Chicago: University of Chicago.

- Long, Rebecca & Sofronio Cruz. 2005. *Diccionario zapoteco de San Bartolomé Zoogocho, Oaxaca*. 2nd edition (Serie de vocabularios y diccionarios indígenas "Mariano Silva y Aceves" no. 38). Coyoacán, D.F. [Mexico]: Instituto Lingüístico de Verano. 531 pp.
- López Nicolás, Oscar. 2016. *Estudios de la fonología y gramática del zapoteco de Zochina*. Mexico City, Mexico: CIESAS Doctoral thesis.
- Masica, Colin P. 1976. *Defining a linguistic area: South Asia*. Chicago: The University of Chicago Press.
- Michaud, Alexis. 2012. The Complex Tones of East/Southeast Asian Languages: Current Challenges for Typology and Modelling. 8.
- Moisik, Scott R. & John H. Esling. 2014. Modeling the Biomechanical Influence of Epilaryngeal Stricture on the Vocal Folds: A Low-Dimensional Model of Vocal–Ventricular Fold Coupling. *Journal of Speech, Language, and Hearing Research* 57(2). S687–S704. [https://doi.org/10.1044/2014\\_JSLHR-S-12-0279](https://doi.org/10.1044/2014_JSLHR-S-12-0279).
- Nellis, Donald G. & Barbara E. Hollenbach. 1980. Fortis versus Lenis in Cajonos Zapotec Phonology. *International Journal of American Linguistics* 46(2). 92–105. <https://doi.org/10.1086/465639>.
- Pike, Kenneth L. 1948. *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion*. Ann Arbor, MI: University of Michigan Press.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. manual. R Foundation for Statistical Computing. Vienna, Austria.
- Ratliff, Martha Susan. 1992. *Meaningful tone: a study of tonal morphology in compounds, form classes and expressive phrases in White Hmong* (Special Report no. 27, 1992). DeKalb, Ill: Northern Illinois University, Center for Southeast Asian Studies. 269 pp.
- Shue, Yen-Liang, Patricia Keating & Chad Vicenik. 2009. VOICESAUCE: A program for voice analysis. *The Journal of the Acoustical Society of America* 126(4). 2221. <https://doi.org/10.1121/1.3248865>.
- Sichel, Ivy & Maziar Toosarvandani. 2020. The featural life of nominals. lingbuzz/005523.
- Silverman, Daniel. 1997. Laryngeal complexity in Otomanguean vowels. *Phonology* 14(2). 235–261. <https://doi.org/10.1017/S0952675797003412>.
- Silverman, Daniel, Barbara Blankenship, Paul Kirk & Peter Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* 37(1).
- Smalley, William A. 1976. The problems of consonants and tone: Hmong (Meo, Miao). In W. A. Smalley (ed.), *Phonemes and orthography: Language planning in ten minority languages of Thailand*, vol. 43 (Pacific Linguistics, Series C), 85–123. Canberra: Australian National University.

- Sonnenschein, Aaron H. 2005. *A descriptive grammar of San Bartolomé Zoogocho Zapotec*. Munich: Lincom Europa.
- Suárez, Jorge. 1983. *The Mesoamerican Indian languages*. Cambridge: Cambridge University Press.
- Thurgood, Graham. 2002. Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* 19(2). 333–363. <https://doi.org/10.1075/dia.19.2.04thu>.
- Uchihara, Hiroto. 2016. Tone and registrogenesis in Quiaviní Zapotec. *Diachronica* 33(2). 220–254. <https://doi.org/10.1075/dia.33.2.03uch>.
- Uchihara, Hiroto & Gabriela Pérez Báez. 2016. Fortis/lenis, glides and vowels in Quiaviní Zapotec. *Glossa: A Journal of General Linguistics* 1(1). 27. <https://doi.org/10.5334/gjgl.13>.
- Villard, Stéphanie. 2015. *The Phonology and Morphology of Zacatepec Eastern Chatino*. Austin, TX: University of Texas at Austin dissertation.
- Wee, Lian-Hee. 2019. *Phonological Tone*. 1st edn. Cambridge University Press. <https://doi.org/10.1017/9781316410912>.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. ELAN : a professional framework for multimodality research. In 1556–1559.
- Yip, Moira Jean Winsland. 2002. *Tone* (Cambridge Textbooks in Linguistics). Cambridge ; New York: Cambridge University Press. 341 pp.