The acoustics of phonation in Santiago Laxopa Zapotec[1]

Mykel Loren Brinkerhoff

1 Introduction

- Phonation is a process where the larynx is used to alter the way different sounds are produced.
- This use of the larynx produces sounds which vary from being more breathy or creaky.
  - Other types of phonation also exist (see Esling et al. 2019 for a detailed discussion on the different phonation types that exist and how the larynx produces them).
- In some languages these alterations are described as being pathological, with some speakers just being more breathy or creaky than others (e.g., Klatt & Klatt 1990).
- In other languages these phonation types are used phonemically like in Gujarati, where vowels can be breathy or modal (Fischer-Jørgensen 1968).
- This phonemic use of phonation is particularly common in the Otomanguean language family from southern Mexico (e.g., Suárez 1983, Campbell, Kaufman & Smith-Stark 1986, Silverman 1997a, Campbell 2017a,b).
- One of the common problems facing linguists who study phonation is determining what the acoustic correlates are for these different phonation types.
- This paper is a preliminary investigation into the acoustic correlates of the different phonation types in Santiago Laxopa Zapotec, an Otomanguean language spoken by approximately 1200 people.
- I show that that H1-H2 is not a very good indicator of phonation in Santiago Laxopa Zapotec, but like other Zapotecs H1-A3 is a good indicator of phonation.
- Additionally, Cepstral Peak Prominence is shown to be useful in distinguishing checked and laryngealized vowels apart for one of the speakers.

[1] 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 of 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.

2 Background

- It is commonly assumed that the vocal folds, which account for differences in pitch and voicing, are also responsible for differences in phonation.
- Ladefoged (1971) and Gordon & Ladefoged (2001) argued that phonation exists on a single dimension ranging from opened vocal folds to closed vocal folds.
- The variation in how open or closed the vocal folds are correspond to whether or not the sound produced is breathy or creaky.
- Ladefoged (1971) and Gordon & Ladefoged (2001) summarized this assumption in the diagram found in Figure 1.

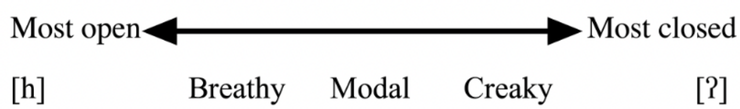


Figure 1: Simplified one-deminsional model for phonation. Based on Ladefoged (1971) and Gordon & Ladefoged (2001)

.

- In addition to this degree of openness, linguists have found that there are other acoustic measurements that correlate to different types of phonation.
- Commonly, these acoustic measurements are spectral-tilt measurements.
- Spectral-tilt measurements are when you take the relative amplitude of different harmonics and subtract them from each other.
- Fischer-Jørgensen 1968 showed that the difference between the amplitude of the first harmonic (H1) and second harmonic (H2) can account for the differences between breathy and modal vowels.
- Further research has continued to rely on H1-H2, both corrected and uncorrected, to distinguish different types of phonation (e.g., Huffman 1987, Klatt & Klatt 1990).
- Further research into spectral-tilt measurements began looking at other measurements besides H1-H2. These include looking at H1 minus the amplitude of the fourth harmonic (H4) or the amplitude of the harmonic closest to the different formants (A*n*).
- From these studies several patterns have emerged (see Garellek 2019 for an overview).

- Breathy phonation typically has a higher spectral-tilt measurement when compared to modal phonation.
- Creaky phonation typically has a lower spectral-tilt measurement than modals.

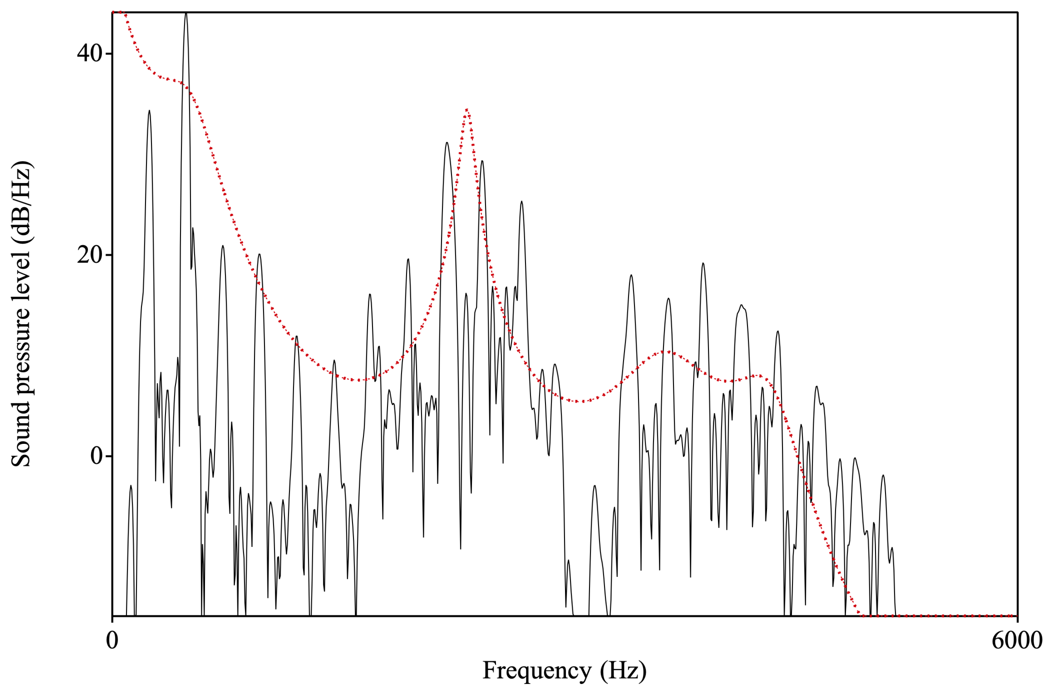


Figure 2: Spectral slice with LPC smoothed line overlaid for the vowel [e]. The harmonics in the spectral slice are represented by each of the dark peaks. The leftmost black solid line peak is the first harmonic (H1) and each subsequent peak represents the next highest harmonic (H2 through H*n*). The red dotted line represents an LPC smoothed line which identifies the formants by the peaks in the line. Each of the harmonics that are closest to the formant peak is identified as A1 through A*n*.

- The spectral-tilt measurement of H1-A3, the difference in amplitude between the first harmonic and the amplitude of the harmonic closest to the third formant, appears to be another of these measurements that robustly distinguishes different phonation types.
- Esposito (2010b) found that in Central Valley Zapotecs, phonation contrasts were best captured by different measurements depending on sex.
- H1-H2 was found to be effective only for female speakers, whereas male speakers were best characterized by H1-A3.
- This pilot study will focus on these two spectral-tilt measurements.

- In addition to spectral-tilt measurements, measurements of noise frequently accompany investigations into phonation (e.g., Garellek & Esposito 2021).
- One of the most commonly used measurements is Cepstral Peak Prominence (CPP; Hillenbrand, Cleveland & Erickson 1994, Hillenbrand & Houde 1996).
- CPP is similar to the harmonics-to-noise ratio measure of de Krom (1993) but differs in how the 'prominence' of the cepstral peak is calculated.
- Prominence is taken as the difference in amplitude of the cepstral peak and a regression line used to normalize for window size and overall energy.
- A more prominent cepstral peak indicates stronger harmonics above the floor of the spectrum (i.e., greater periodicity in the speech signal).
- CPP was originally used as a diagnostic for breathy or modal voice (Blankenship 2002, Esposito 2010b) In fact, Esposito showed that CPP was the best of the eight measures she considered for distinguishing modal from breathy phonation types.
- Further research has shown that CPP is also a good measurement for any type of non-modal phonation (e.g., Andruski & Ratliff 2000, Andruski 2006, Blankenship 2002, Wayland & Jongman 2003, Avelino 2010).
- This will be the third measurement considered for this pilot study.

3 Santiago Laxopa Zapotec

- Santiago Laxopa Zapotec (SLZ), endonym *Dille'xhunh Laxup*, is a a Northern Zapotec language spoken by approximately 1000 people in the municipality of Santiago Laxopa, Ixtlán, Oaxaca, Mexico and in diaspora communities in Mexico and the United States (Adler & Morimoto 2016, Adler et al. 2018, Foley, Kalivoda & Toosarvandani 2018, Foley & Toosarvandani 2020).
- Closely related to San Bartolomé Zoogocho Zapotec (Long & Cruz 2005, Sonnenschein 2005) and shares a high level of mutual intelligibility with it.
- SLZ is similar to other Zapotecan languages in distinguishing lenis and fortis consonants (e.g., Nellis & Hollenbach 1980, Jaeger 1983, Uchihara & Pérez Báez 2016).
- SLZ has a standard five vowel inventory.

Table 1: Consonant inventory for Santiago Laxopa Zapotec

		bilabial	alveolar	post-alveolar	retroflex	palatal	velar	labio-velar	uvular
stop	lenis	b	d				g	gʷ	
	fortis	p	t				k	kʷ	
fricative	lenis		z	ʒ	ʐ				ʁ
	fortis		s	ʃ	ʂ	ç			
affricate	lenis		d͡z						
	fortis		t͡s		t͡ʃ				
nasal	lenis		n						
	fortis	mː	nː						
lateral	lenis		l						
	fortis		lː						
trill			r						
approximate								w	

Table 2: Vowel qualities in Santiago Laxopa Zapotec.

	front	central	back
high	i		u
mid	e		o
low		a	

- These five vowels, additionally, appear with one of four different phonation types which will be discussed in greater detail in Section 3.2.

3.1 Tone in Santiago Laxopa Zapotec

- Similar to other Otomanguean languages, SLZ is tonal (Suárez 1983, Campbell, Kaufman & Smith-Stark 1986, Silverman 1997a, Campbell 2017a,b).

- SLZ has five distinct tonal patterns that appear on the syllables of nouns, see Table 3.

- Figure 3 shows the five tonal contrasts averaged for each tonal contrast from the onset to ending of the vowel.

- The first 20-25% of Figure 3 can be ignored due to the influence of consonantal transitions.

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

High	a^{H}	*xha*	[ʐa^{H}]	‘clothing.POSS’
Mid	a^{M}	*lhill*	[liʒ^{M}]	‘house.POSS’
Low	a^{L}	*yu’*	[çuˀ^{L}]	‘earth’
Rising	a^{MH}	*yu’u*	[çuˀu^{MH}]	‘quicklime (Sp. cal)’
Falling	a^{HL}	*yu’u*	[çuˀu^{HL}]	‘house’

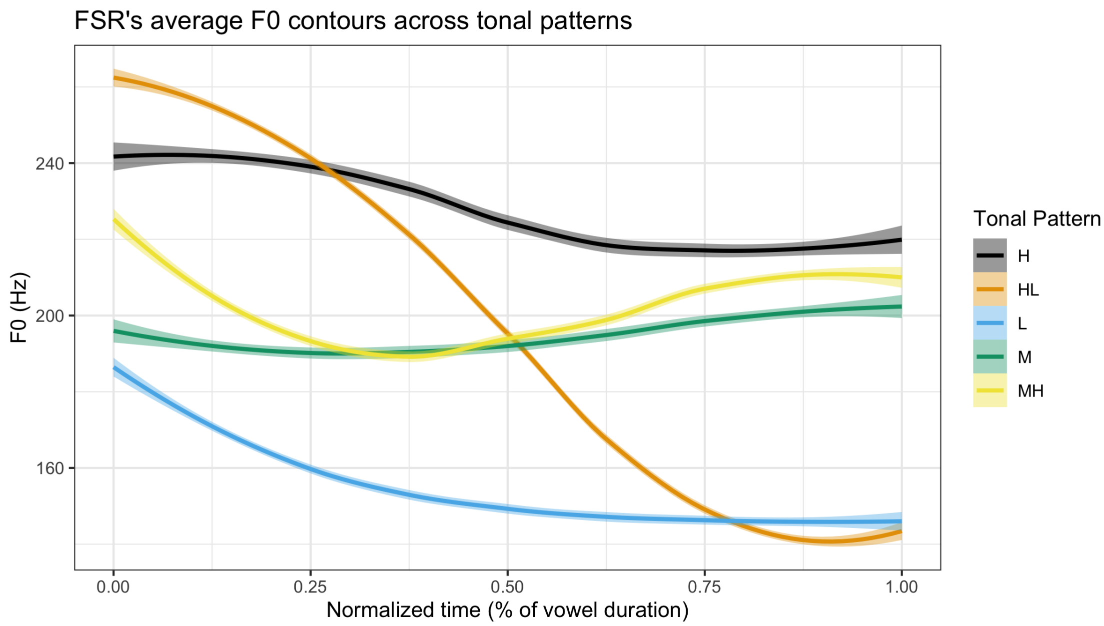


Figure 3: Tonal contrasts for FSR averaged and time normalized. Each line in this graph represents the average of approximately 10 syllables for each tonal pattern.

3.2 Phonation in Santiago Laxopa Zapotec

- Zapotecan languages commonly make use of contrastive phonation on vowels (e.g., Avelino 2004, 2010, Long & Cruz 2005, López Nicolás 2016, Chávez-Peón 2010, Ariza-García 2018).

- SLZ has four contrastive phonation types, see (1)

(1) Four-way near minimal phonation contrast

a. *yag* [çag^{L}] ‘tree; wood; almúd (unit of measurement approximately 4kg)’

b. *yah* [ça̤^{L}] ‘metal; rifle; bell’

c. *yu’* [çuˀ^{L}] ‘earth’

d. *ya’a* [çaˀa^{L}] ‘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 4.

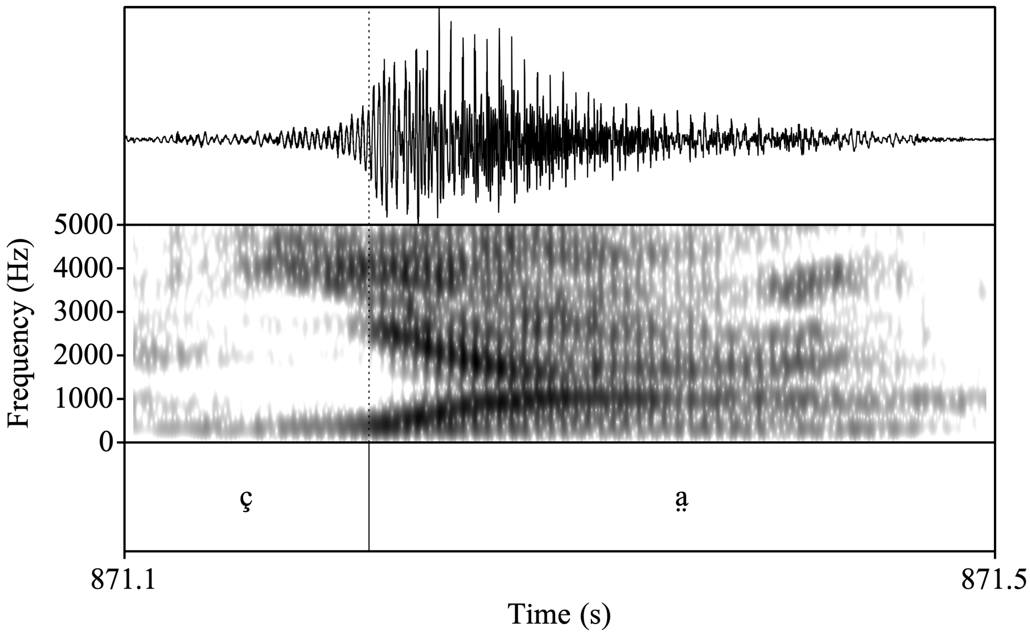


Figure 4: FSR's breathy vowel in the word *yah* 'metal; rifle'

- 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 5.
- Laryngealized vowels are quite common in Zapotecan languages and have received a wide 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, Ariza-García 2018).
- In addition to a wide number of different names these vowels also exhibit a wide range of allophones.
- Avelino (2010) found in the closely realted Yalálag Zapotec that among his consultants there were at least four different pronunciations as seen in Table 4.
- In SLZ, this vowel is also highly variable.
- One consultant does rearticulation, where there is a full glottal stop in the middle of the vowel, or creaky voice, see Figure 6. This alternation seems to be in free variation but there was a greater tendency to creak in words with a L tone, such as *xa'ag* [ʂa̰ːx] 'topil'[2].

[2] A topil is a type of government office in traditional Oaxacan communities somewhat akin to a sheriff.

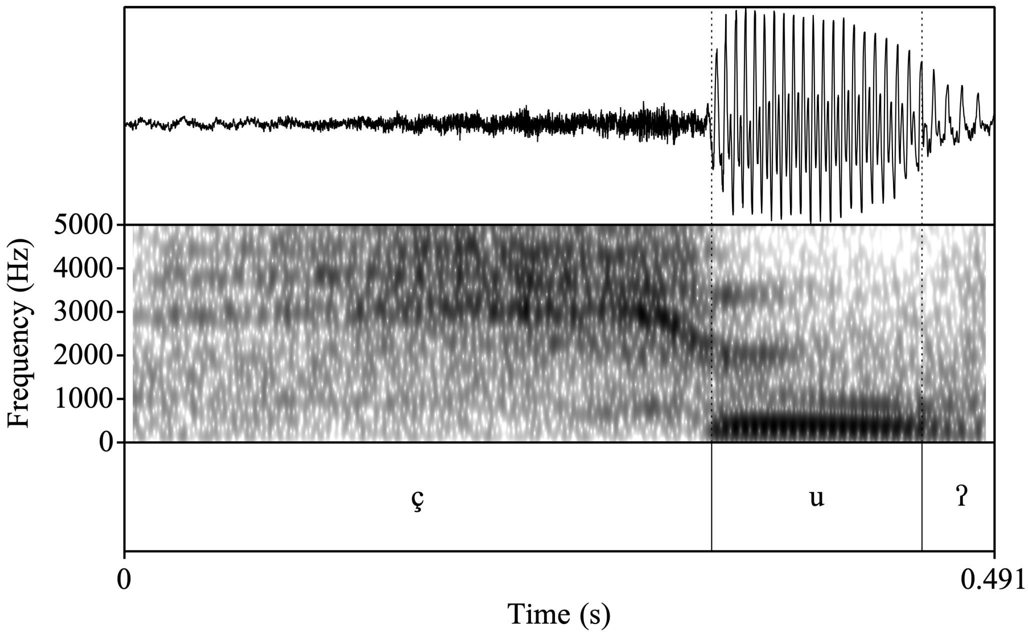


Figure 5: RD's checked vowel in the word *yu'* 'earth'

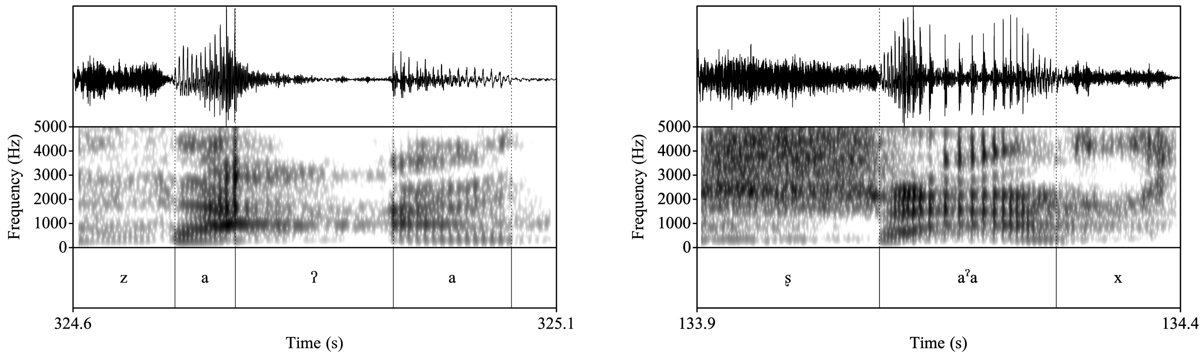
Table 4: Layngealized Vowels in Yalálag Zapotec

/Vˀ V/	[VʔV]
	[VV̰V]
	[VV̰ːV̆]
	[VV̰V̰]

- Another consultant only ever produced creaky voice for these vowels regardless of the tone with the word, see Figure 7.
- Further research will show whether these differences are sociolinguistic in nature or whether these differences are individualistic.

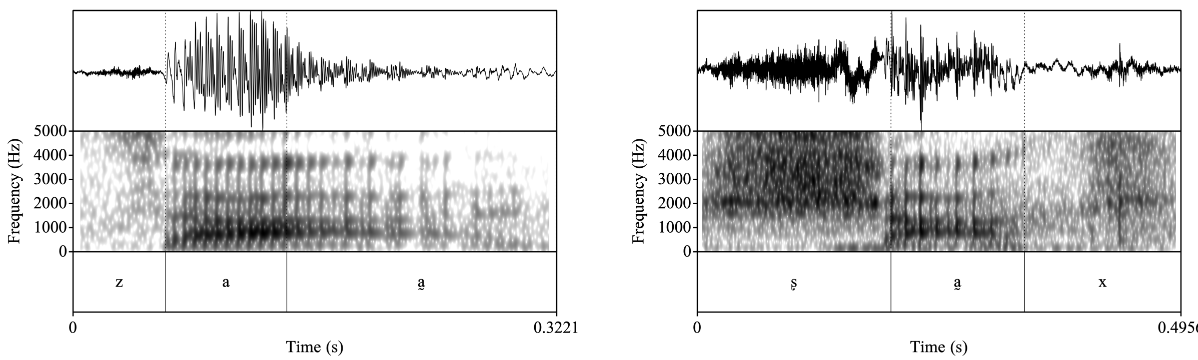
3.3 Interaction of tone and phonation in Santiago Laxopa Zapotec

- 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).
- Smalley (1976) and Ratliff (1992) both describe White Hmong's (Hmong-Mien) -*g* tone as being a mid-low tone with breathy phonation



(a) *za'a* 'corncob' (b) *xa'ag* 'topil'

Figure 6: Comparison of FSR's laryngealized vowels in *za'a* 'corncob' and *xa'ag* 'topil'



(a) *za'a* 'corncob' (b) *xa'ag* 'topil'

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

- Mandarin's (Sino-Tibetan) 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 (Sino-Tibetan) has found that two tones are only differentiated by the presence of some form of non-modal phonation (Boehm p.c.).
- In Oto-Manguean languages, there is strong evidence that tone and phonation are independent from each other (e.g., Silverman 1997a, Avelino 2004, 2010, Chávez-Peón 2010, Campbell 2011, Garellek & Keating 2011, Villard 2015, López Nicolás 2016, Ariza-García 2018).
- In San Lucas Quiaviní Zapotec (Valley Zapotec) almost all of the the different tones and phonation types can be combined with each other.

- The rising tone in SLQZ is restricted to only appearing with modal voice.

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

	Modal	**Breathy**	**Creaky**	**Interrupted**
High	✓	–	✓	✓
Low	✓	✓	✓	✓
Falling	✓	✓	✓	✓
Rising	✓	–	–	–

- SLZ is similar to SLQZ in allowing almost all combinations of tone and phonation.
- There are two gaps in the distribution of tone and phonation: Breathy-High and Checked-Falling.

Table 6: Observed combinations of tone and phonation in SLZ.

	Modal	**Breathy**	**Checked**	**Laryngealized**
High	✓	–	✓	✓
Mid	✓	✓	✓	✓
Low	✓	✓	✓	✓
Rising	✓	✓	✓	✓
Falling	✓	✓	–	✓

- 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.

4 Methodology

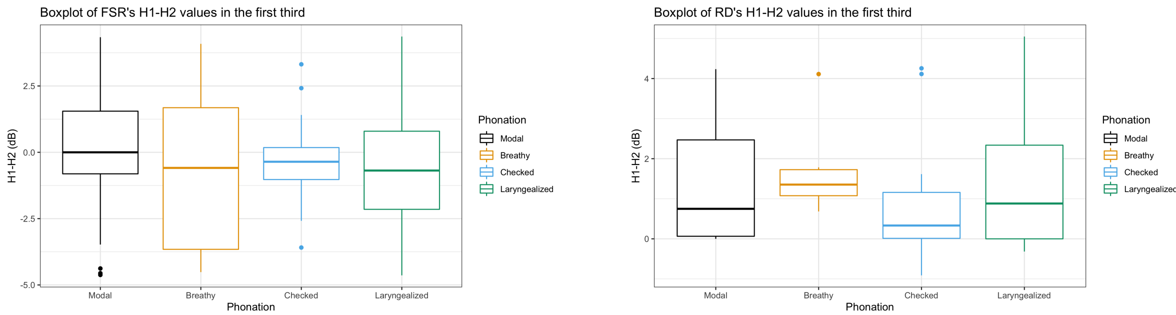
- Two native language speakers of SLZ who live in Santa Cruz, CA took part in this study (one male).

- Because of the COVID-19 pandemic data collection was conducted remotely using Zencastr, 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).
- The SLZ speakers were recorded saying approximately 100 words in isolation three times and those same words three times in the carrier phrase *shnia' X chonhe lhas* 'I say X three times'.
- The resulting files were first segmented in ELAN (Wittenburg et al. 2006) then each individual vowel from the target word was isolated from the carrier sentences using Praat (Boersma & Weenink 2021).
- After isolating each vowel token the files were fed into VoiceSauce (Shue, Keating & Vicenik 2009) for calculating the different spectral-tilt and noise measurements.
- Each vowel was normalized for time and then the various acoustic measurements were averaged for each third of the vowel following Garellek & Keating (2011).

5 Results

5.1 H1-H2 spectral-tilt results

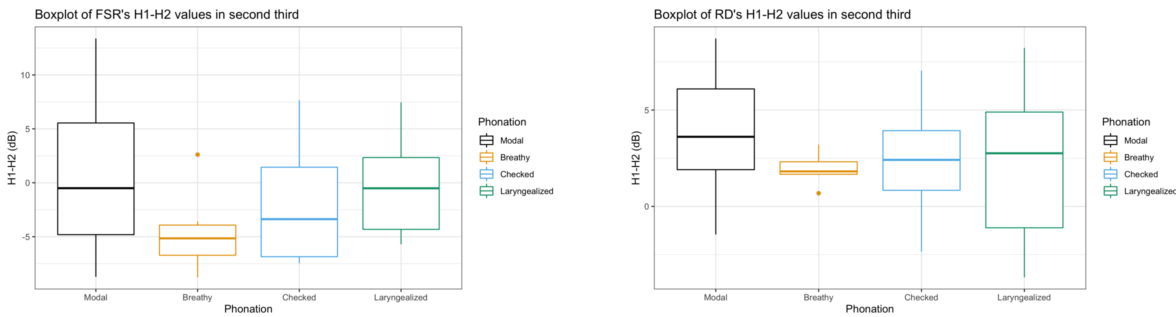
- In all three portions of the vowels, H1-H2 was not a very good measure for spectral-tilt in this variety of Zapotec.
- In all three thirds of the vowel, we see very little differences in the values for H1-H2 between modal, checked, and laryngealized.
- In the first third of the vowel there are virtually no differences between the four phonation types, see Figure 8.
- Beginning in the second third of the vowel and continuing throughout the vowel, the H1-H2 values for breathy voice is lower than the other three phonation types, which is the opposite of what we expect for breathy voice, see Figure 9 for the second third and Figure 10 for the final third.



(a) FSR's H1-H2 values.

(b) RD's H1-H2 values.

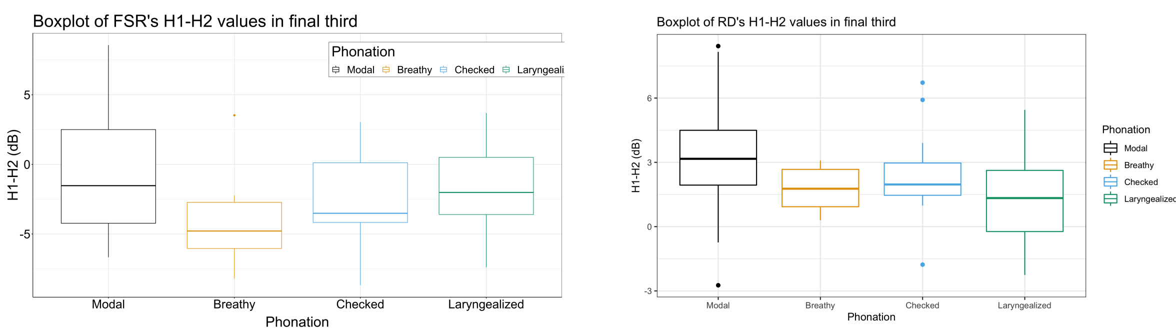
Figure 8: Mean H1-H2 values for the first third of the vowel according to each phonation type.



(a) FSR's H1-H2 values.

(b) RD's H1-H2 values.

Figure 9: Mean H1-H2 values for the second third of the vowel according to each phonation type.



(a) FSR's H1-H2 values.

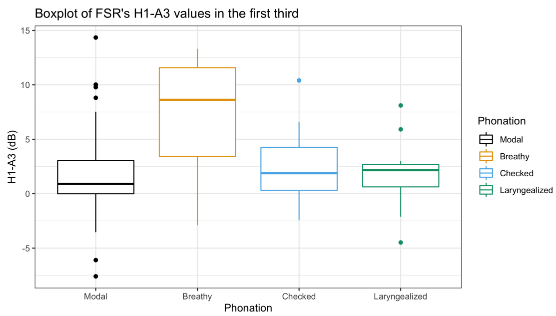
(b) RD's H1-H2 values.

Figure 10: Mean H1-H2 values for the final third of the vowel according to each phonation type.

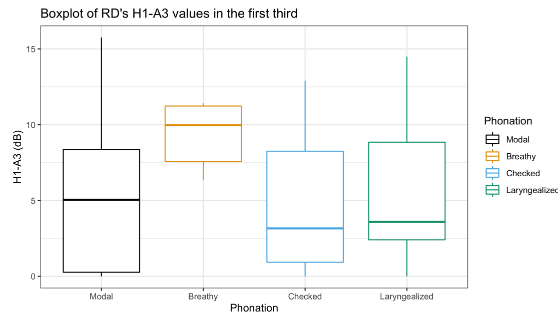
5.2 H1-A3 spectral-tilt results

- When we consider H1-A3, the measurement Esposito (2010b) found useful for male Zapotec speakers, several observations become apparent.
- In all portions of the vowel, breathy voice shows a much higher H1-A3 measurement than modals as observed in Figures 11, 12, and 13.

- This high H1-A3 value in both speakers is what we expect for breathy vowels (Fischer-Jørgensen 1968).



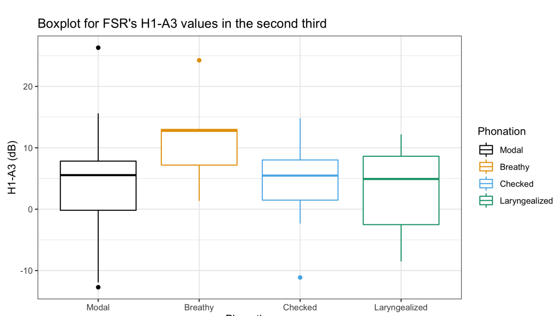
(a) FSR's H1-A3 values.



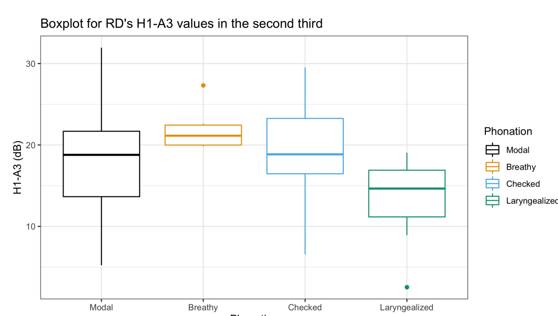
(b) RD's H1-A3 values.

Figure 11: H1-A3 values for FSR (a) and RD (b) for the first third of the vowel.

- In Figure 11, we see that there is a large degree of overlap between modals, checked, and laryngealized H1-A3 values. This behavior indicates that there is little to no difference in the phonation types for this portion of the vowel.
- In the second third of the vowel (Figure 12), we see that the checked and laryngealized vowels H1-A3 values for FSR are uninformative because of the large degree of overlap.
- Beginning in the middle of the vowel and continuing throughout the rest of the vowel, RD shows a lower H1-A3 value in the laryngealized vowels than the modals which is consistent with creakier productions of vowels, see Figure 12b and Figure 13b.
- In both speakers there is no difference between modal and creaky for the middle third of the vowel.



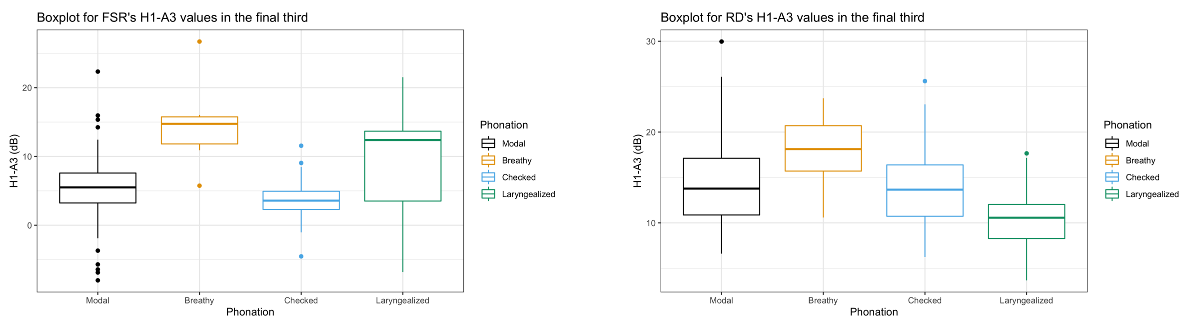
(a) FSR's H1-A3 values.



(b) RD's H1-A3' values.

Figure 12: H1-A3 values for FSR (a) and RD (b) for the second third of the vowel.

- In the final third of the vowel, FSR's production of checked vowels show a lower H1-A3 measurement, which suggests FSR produces a period of creakiness in the last third of these vowels.
- For RD, however, there are no differences between checked and modal vowels in this last third.



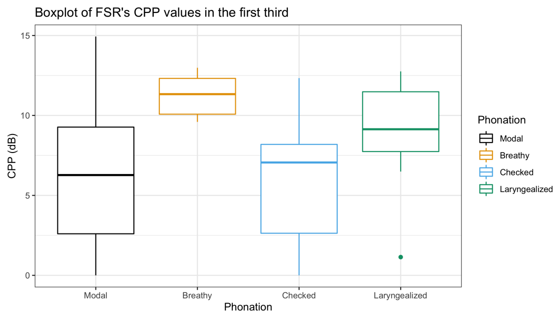
(a) FSR's H1-A3 values. (b) RD's H1-A3 values.

Figure 13: H1-A3 values for FSR (a) and RD (b) for the final third of the vowel.

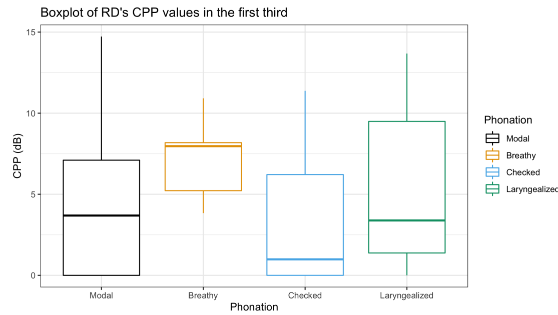
5.3 Cepstral Peak Prominence results

- In determining if there is any amount of aperiodicity in the signal, we turn to CPP (Hillenbrand, Cleveland & Erickson 1994, Hillenbrand & Houde 1996).
- A CPP measurement for any vowel with non-modal phonation should be lower than the CPP measurement for modal vowels.
- Throughout the whole vowel, CPP measures for both FSR and RD are consistently higher for breathy vowels than modal vowels.
- This observation seems to suggest that breathy vowels are in some way more periodic than modal vowels.

- When we turn to the middle third of the vowel, Figure 15, we observe that FSR's laryngealized vowels show a lower CPP value, indicating that there is aperiodicity in this portion of the vowel.
- For RD, there are no differences between modal, breathy, or checked in this portion of the vowel.

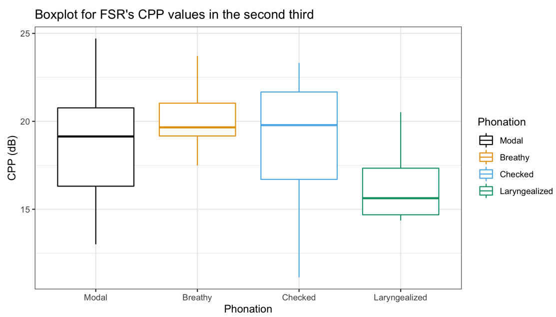


(a) FSR's CPP values.

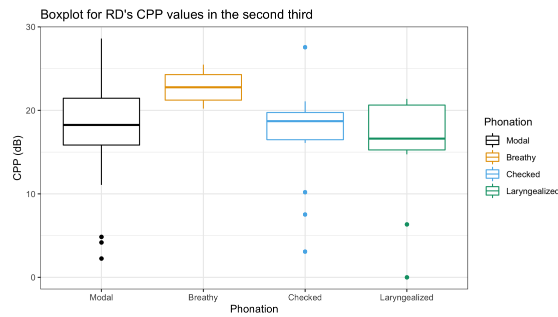


(b) RD's CPP values.

Figure 14: CPP values for FSR (a) and RD (b) for the first third of the vowel.



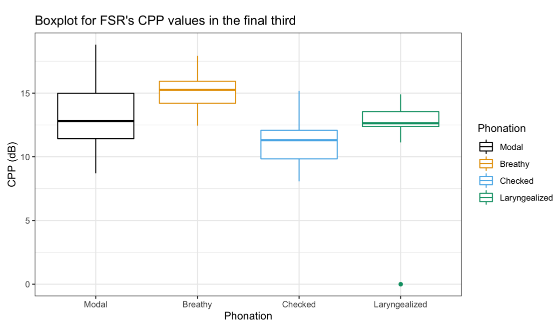
(a) FSR's CPP values.



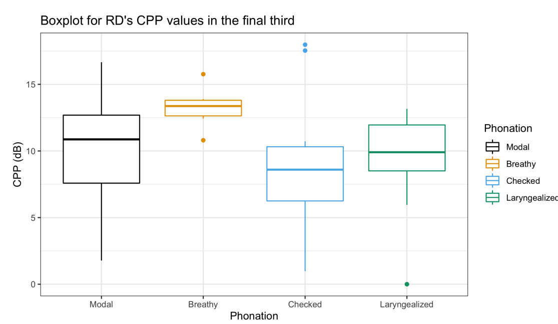
(b) RD's CPP values.

Figure 15: CPP values for FSR (a) and RD (b) for the second third of the vowel.

- In both FSR and RD, we observe a lower CPP value for checked vowels in the final third of the vowel.
- This is consistent with a period of aperiodicity at the end of the vowel.



(a) FSR's CPP values.



(b) RD's CPP values.

Figure 16: CPP values for FSR (a) and RD (b) for the final third of the vowel.

5.4 Statistical results

- In order to determine whether or not the different acoustic measurements were effective at capturing the different phonation types, a mixed-effects linear regression analysis was performed with a Satterthwaite's method for t-test analysis used to derive the p-values.
- Each of different measurements were treated as the dependent variable with word and speaker treated as random.

5.4.1 H1-H2 statistical results

- As can be recalled from Section 5.1, both FSR and RD show a lower value for H1-H2 for breathy vowels in the last two-thirds of the vowels when compared to the model vowel's H1-H2 values.
- The results of the statistical analysis for the last two-thirds of the vowel, presented in Table 8 and Table 9, show that this behavior is significant.
- At no point do the other phonation types reach significance with respect to H1-H2.

Table 7: Results of the mixed-effects linear regression analysis on the first third of the vowel for H1-H2.

	Estimate	Std. Error	df	t value	p-value
Breathy	-0.3049	0.5795	116.4255	-0.526	0.600
Checked	-0.4731	0.4029	145.6184	-1.174	0.242
Laryngealized	-0.2938	0.4675	101.9977	-0.629	0.531

Table 8: Results of the mixed-effects linear regression analysis on the second third of the vowel for H1-H2.

	Estimate	Std. Error	df	t value	p-value
Breathy	-3.4158	1.3400	126.9590	-2.549	0.0120
Checked	-1.7466	0.9197	146.9995	-1.899	0.0595
Laryngealized	-0.7405	1.0852	117.0669	-0.682	0.4963

Table 9: Results of the mixed-effects linear regression analysis on the final third of the vowel for H1-H2.

	Estimate	Std. Error	df	t value	p-value
Breathy	-2.2131	1.0291	125.2092	-2.151	0.0334
Checked	-1.3487	0.6952	146.2461	-1.940	0.0543
Laryngealized	-0.8676	0.8202	116.8963	-1.058	0.2923

5.4.2 H1-A3 statistical results

- The second statistical analysis for the H1-A3 measurement shows some very clear behavior for breathy voice.
- In Section 5.2 we observe that breathy voice is clearly identified in all portions of the vowel with an elevated H1-A3 value when compared to modals.
- We see that for all portions of the vowel, as seen in Tables 10, 11, 12 breathy voice reaches significance.

Table 10: Results of the mixed-effects linear regression analysis on the first third of the vowel for H1-A3.

	Estimate	Std. Error	df	t value	p-value
Breathy	4.29036	1.31288	137.12067	3.268	0.00137
Checked	0.15001	0.84476	134.62215	0.178	0.85932
Laryngealized	-0.05694	1.07130	137.16596	-0.053	0.95769

- When corrected for word and speaker, laryngealized vowels reach significance in the second third of the vowel.

Table 11: Results of the mixed-effects linear regression analysis on the second third of the vowel for H1-A3.

	Estimate	Std. Error	df	t value	p-value
Breathy	4.398	2.027	147.003	2.169	0.0317
Checked	1.195	1.426	147.000	0.838	0.4033
Laryngealized	-2.694	1.627	147.001	-1.656	0.0998

- In the final third of the vowel, H1-A3 was only significant for breathy voice.

Table 12: Results of the mixed-effects linear regression analysis on the final third of the vowel for H1-A3.

	Estimate	Std. Error	df	t value	p-value
Breathy	5.2928	1.7966	127.0400	2.946	0.00383
Checked	-0.7825	1.2487	147.1011	-0.627	0.53189
Laryngealized	0.2669	1.4224	117.4917	0.188	0.85151

5.4.3 CPP statistical results

- When we consider CPP, we see that this measurement was significant for all three of the phonation types but in different portions of the vowel.
- In Table 13, we see that the high CPP values observed in Section 5.3 for the first third of the vowel was significant.
- For checked and laryngealized vowels CPP fails to reach significance in this first third.

Table 13: Results of the mixed-effects linear regression analysis on the first third of the vowel for CPP.

	Estimate	Std. Error	df	t value	p-value
Breathy	3.4470	1.1743	145.9760	2.935	0.00387
Checked	-1.4190	0.7088	120.5020	-2.002	0.04754
Laryngealized	0.8240	0.9530	147.0083	0.865	0.38868

- In the middle of the vowel, only laryngealized voice was significant, see Table 14.

Table 14: Results of the mixed-effects linear regression analysis on the second third of the vowel for CPP.

	Estimate	Std. Error	df	t value	p-value
Breathy	1.5044	1.4073	128.0083	1.069	0.287094
Checked	-0.5804	0.9789	146.0792	-0.593	0.554154
Laryngealized	-2.2898	1.1354	117.6213	-2.017	0.046006

- In the final portion of the vowel, Table 15, only the CPP values for checked phonation reached significance.

Table 15: Results of the mixed-effects linear regression analysis on the final third of the vowel for CPP.

	Estimate	Std. Error	df	t value	p-value
Breathy	1.6391	1.0173	139.8563	1.611	0.10939
Checked	-2.1386	0.6449	129.1392	-3.316	0.00119
Laryngealized	-1.1158	0.8142	140.5570	-1.370	0.17274

6 Discussion

6.1 Acoustics discussion

6.1.1 H1-H2 discussion

- Overall, H1-H2 was uninformative for phonation in SLZ.
- This is in contrast to Adler & Morimoto (2016) which found for two male speakers of SLZ H1-H2 was informative.
- This difference could be attributed to a hypothesis put forward by Esposito (2010b) that maybe the speakers are producing the phonation types differently than the speakers in this study.
- Additionally, Chai & Garellek (2022) remark that there is a growing body of evidence that H1-H2 is not a reliable measure of phonation.
- Sundberg (2022) found that this unreliability is because H2 can be highly influenced by differences in subglottal pressure.
- More specifically, H2 is more sensitive to the pulse amplitude increase than H1.
- This sensitivity results in an inverse relationship between H1-H2 and pulse amplitude.
- Furthermore, the results also show that H1-H2 is not a reliable measure in SLZ.
- Another question that this raises is whether more speakers were sampled if the results would be different.

6.1.2 H1-A3 discussion

- H1-A3, however, appears to be a good measure for breathy voice in SLZ for both speakers.

- This shows that, similar to Esposito (2010b), H1-A3 is a good measure for capturing some of the phonation contrasts in SLZ.
- However, unlike Esposito, this measure was not informative for creaky or laryngealized voice as indicated by the linear regression.
- Similar to H1-A3 it is possible that the results might be different if there was a larger sample pool.

6.1.3 CPP discussion

- The results of CPP were both informative and concerning.
- Breathy voice, as seen in Section 5.3 and shown to be significant in Section 5.4.3, the CPP value was much higher for both speakers.
- This is the opposite of what we expect to see for periods of aperiodicity (Hillenbrand, Cleveland & Erickson 1994, Hillenbrand & Houde 1996).
- This suggests that breathy vowels were more periodic that modal vowels.
- It is possible that several of the vowels were mislabeled as being modal or breathy when they were actually a different phonation type.
  - This is because it is rather difficult to hear breathy voice when listening to the audio produced by Zencastr.
  - It is much easier to hear breathy voice when in person.
  - Additionally, when in Laxopa this summer, Maya and I realized that many words had the wrong phonation type associated with them.
- With checked and laryngealized voice, we see the drop in CPP we expect to see with aperiodic noise.
- These drops were significant even when correcting for word and speaker
- This suggests that aperiodic noise might be the main cue to laryngealized and checked voice for these speakers.
- Additionally, we notice that there is an apparent difference in timing between when this aperiodic noise begins.
  - Laryngealized vowels have this aperiodic noise centered in the middle of the vowel.

  - Checked vowel have this aperiodic noise begin in the final third of the vowel.

- This difference in timing is important because both phonation types are associated with creakiness in the vowel.

- This is consistent with the observation that somewhere in the middle of the vowel there is either a full glottal stop or a period of creakiness in the two speakers that were evaluated for this study.

- This observation also bears witness to observations made by Avelino (2010) about the how each of the different manners that laryngealized vowels are produced in Yalálag Zapotec.

- This difference in timing is similar to observations made by Silverman (1997a,b) and Blankenship (1997, 2002) for nonmodal phonation in the Laryngeal Complexity Hypothesis.

6.2 Laryngeal Complexity Hypothesis

- The Laryngeal Complexity Hypothesis (LCH) has its origin in work from Silverman (1997a) and Blankenship (1997, 2002).

- The LCH claims that languages that have both tone and phonation need them to be phased/ordered with respect to each other.

- This is required because it is assumed that the same mechanism for tone is also responsible for phonation.

  - Tone is the rate of vocal fold vibration.
  - Ladefoged (1971) and Gordon & Ladefoged (2001) argued that phonation exists on a single dimension ranging from opened vocal folds to closed vocal folds.

- The variation in how open or closed the vocal folds are correspond to whether or not the sound produced is breathy or creaky.

- Ladefoged (1971) and Gordon & Ladefoged (2001) summarized this assumption in the diagram found in Figure 1, repeated in Figure 17.

- Because the same organ is responsible for these two different phenomena there is a mismatch in trying to produce both at the same time.

- The LCH assumes that there needs to be a strict ordering in the glottal gestures.

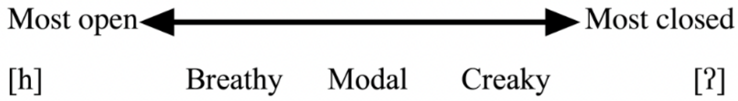


Figure 17: Simplified one-deminsional model for phonation. Based on Ladefoged (1971) and Gordon & Ladefoged (2001)

.

- 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.
- In Figure 18, 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.

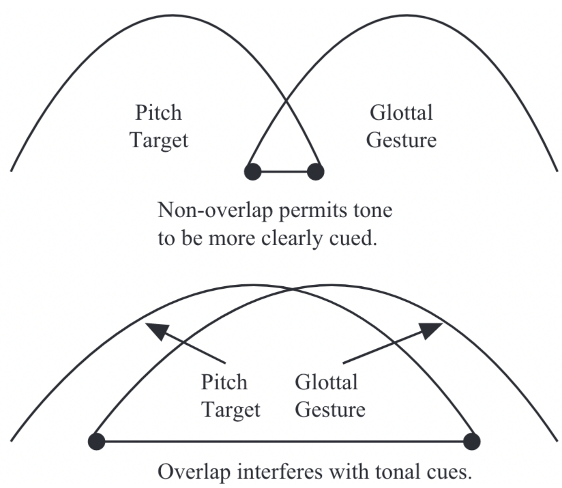


Figure 18: Representation taken from DiCanio (2012).

- DiCanio (2012) 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.

- 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.
- SLZ could potentially speak to these differences in timing for nonmodal phonation because of the four way phonation contrast that exists for vowels.
- It already appears from the results of CPP that there is a difference in timing between laryngealized and checked vowels.
- Additionally, SLZ could speak to the phasing of tone and phonation in the vowels, because tone and phonation is orthogonal to one another.
- In order to test the LCH in SLZ, a more controlled token list is needed as well as a larger participant pool to determine if the differences we observe between laryngealized and checked vowels are factual.

7 Conclusion

- This paper has provided a brief introduction to SLZ's phonation and the other aspects of its sound system.
- This work is important because SLZ is an understudied variety of Northern Zapotec and there is very little about the phonetics or phonology of this language available.
- Additionally, this paper has shown that H1-H2 was not a very reliable measurement for FSR's or RD's phonation type.
- In line with Esposito (2010b), H1-A3 seemed to be the more reliable phonation type measurement, at least for breathy voice.
- Additionally, CPP was shown to be a very good indicator for checked and laryngealized vowels in SLZ.
- Additionally, CPP seemed to be primarily different in terms of timing for checked and laryngealized.
- This study will benefit from further analysis and data collection.
- Now that the world is in a safer state in regards to COVID-19 it is important to collect data from more speakers to confirm the data and analysis from FSR and RD.

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. 2006. Tone clarity in mixed pitch/phonation-type tones. *Journal of Phonetics* 34(3). 388–404. https://doi.org/10.1016/j.wocn.2005.07.001.

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.

Ariza-García, Andrea. 2018. Phonation types and tones in zapotec languages: a synchronic comparison. *Acta Linguistica Petropolitana* XIV(2). 485–516. https://doi.org/10.30842/alp2306573714220.

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 Electroglottographic Analyses of Nonpathological, Nonmodal 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).

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 W. 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.

Chai, Yuan & Marc Garellek. 2022. On H1–H2 as an acoustic measure of linguistic phonation type. *The Journal of the Acoustical Society of America* 152(3). 1856–1870. https://doi.org/10.1121/10.0014175.

Chávez-Peón, Mario E. 2010. *The interaction of metrical structure, tone, and phonation types in Quiaviní 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. 2010a. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38(2). 306–316. https://doi.org/10.1016/j.wocn.2010.02.002.

Esposito, Christina M. 2010b. 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.

Fischer-Jørgensen, Eli. 1968. Phonetic Analysis of Breathy (Murmured) Vowels in Gujarati. *Annual Report of the Institute of Phonetics University of Copenhagen* 2. 35–85. https://doi.org/10.7146/aripuc.v2i.130674.

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.

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 & Christina M. Esposito. 2021. Phonetics of White Hmong vowel and tonal contrasts. *Journal of the International Phonetic Association*. 1–20. https://doi.org/10.1017/S0025100321000104.

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.

Hillenbrand, James & Robert A. Houde. 1996. Acoustic Correlates of Breathy Vocal Quality: Dysphonic Voices and Continuous Speech. *Journal of Speech, Language, and Hearing Research* 39(2). 311–321. https://doi.org/10.1044/jshr.3902.311.

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.

Klatt, Dennis H. & Laura C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87(2). 820–857. https://doi.org/10.1121/1.398894.

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 Zoochina*. 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.

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.

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.

Silverman, Daniel. 1997a. Laryngeal complexity in Otomanguean vowels. *Phonology* 14(2). 235–261. https://doi.org/10.1017/S0952675797003412.

Silverman, Daniel. 1997b. *Phasing and recoverability* (Outstanding Dissertations in Linguistics). New York: Garland Pub. 242 pp.

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.

Sundberg, Johan. 2022. Objective Characterization of Phonation Type Using Amplitude of Flow Glottogram Pulse and of Voice Source Fundamental. *Journal of Voice* 36(1). 4–14. https://doi.org/10.1016/j.jvoice.2020.03.018.

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.

Wayland, Ratree & Allard Jongman. 2003. Acoustic correlates of breathy and clear vowels: the case of Khmer. *Journal of Phonetics* 31(2). 181–201. https://doi.org/10.1016/S0095-4470(02)00086-4.

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.