

# Variation in contrastive phonation in Santa Ana Del Valle Zapotec

**Christina M. Esposito**

Department of Linguistics, Macalester College  
*esposito@macalester.edu*

The present study sets out to investigate variation due to gender, F0, and/or prosodic position in Santa Ana del Valle Zapotec (Oto-Manguéan), a language with phonemically breathy, modal and creaky vowels, each associated with a tone. Male and female speakers produced words in five prosodic positions: isolation (with focus, F0 higher than sentence-medial position), initial (focused, high F0), isolation (without focus, mid-range F0), medial (mid-range F0), final (lower F0). Two acoustic measures of phonation, H1-H2 and H1-A3, were made for each vowel. Results were inconclusive as to whether one gender was creakier or breathier than the other, though they did suggest that there was a difference in the PRODUCTION of phonation. In addition, there was also a strong effect of F0 on phonation, but not of position independently of F0. While the three-way phonation contrast was present in all five prosodic positions, it was not always well-defined. The contrast was minimized in isolation with focus (high F0) and initial position (high F0). The results obtained indicate that there is variation in phonation, even in a language with contrastive phonation.

## 1 Introduction

Studies have shown that there can be considerable variation in phonation based on gender, F0, and/or position.<sup>1</sup> For example, females were reported to be breathier than males in several

---

<sup>1</sup> Ladefoged (1971) proposed a continuum of phonation types characterized by their degree of glottal opening. The continuum in order from greatest glottal aperture to smallest is: voiceless, breathy, murmur, lax voice, voice (i.e. modal), tense voice, creaky voice, and glottal stop. For a full review of these and other phonation types see Ladefoged (1971), Laver (1980) and Gordon & Ladefoged (2001). This paper is concerned with three of the points on this continuum: breathy, modal and creaky. All three are produced with vibrating vocal folds. Differences between the phonations are due to the adductive and longitudinal tension of the vocal folds. During breathy phonation the vocal folds are minimally adducted with little longitudinal tension. For modal phonation, the vocal folds have normal adductive and longitudinal tension. For creaky phonation, the vocal folds are tightly adducted. However, more recently, Edmondson & Esling (2006) proposed that phonation is produced by manipulating six valves. The six valves are (i) glottal vocal fold adduction and abduction, (ii) ventricular incursion, (iii) sphincteric compression, (iv) epiglottal-pharyngeal constriction, (v) laryngeal raising, and (vi) pharyngeal narrowing. In this system, creaky voice is produced by vocal folds that vibrate slowly, with sphincteric compression, and little or no ventricular incursion. Breathier voice is produced by partial vocal fold adduction, which leaves a large opening between the arytenoid cartilages. Breathier phonation can be produced with vocal folds that oscillate anteriorly along the ligamental glottis, while the posterior portion of the glottis remains open.

languages/dialects such as Japanese (Todaka 1993), American English (Klatt & Klatt 1990, Hanson 1997, Hanson & Chuang 1999), and British English (Henton & Bladon 1985). Some of these gender-dependent variations could be due to physiological differences; for example, direct observations of the vocal folds via fiberoptic showed that females were more likely to have incomplete glottal closure than males, which can produce a breathier voice quality (Södersten & Lindestad 1990). Gender-based phonation differences, however, do not always vary in a systematic way. For example, while results of acoustic analyses showed that on average American-English-speaking females were breathier than males, there was a great deal of gender variation and, in fact, some males were breathier than the females (Klatt & Klatt 1990). Furthermore, Redi & Shattuck-Hufnagel (2001) examined glottalization (a phonation similar to creaky) and found that glottalization in American English could not be predicted from gender alone.

In addition to gender, studies have suggested that phonation can be sensitive to changes in fundamental frequency (F<sub>0</sub>), whether at the lexical or utterance level. For example, a lower F<sub>0</sub> was correlated with breathy-voiced stops in Hindi (Ohala 1973), while a higher F<sub>0</sub> was correlated with a tense voice in Jingpho, Lahu and Yi (Maddieson & Hess 1987). Furthermore, Hagen (1997) found that in English and German, glottalization occurred more often on words produced with a low F<sub>0</sub>.

Studies have also shown that phonation can vary due to position. For example, Epstein (2002) showed through inverse filtering that tense phonations were found utterance-initially regardless of F<sub>0</sub> in American English. Other studies have shown that the ends of sentences or paragraphs in English can be associated with either (i) creakiness (or glottalization) (Lehiste 1975, Kreiman 1982, Henton & Bladon 1985) or (ii) a 'breathy-laryngealized' voice, a type of voice quality characterized by a simultaneous increase in the posterior glottal chink creating breathy voice and a rotation of the anterior tips of the arytenoid cartilages creating laryngealization (Klatt & Klatt 1990). (This effect can likely be explained as the initiation of aryepiglottic trilling, Esling 2005.) Several studies have reported creak phrase-initially, phrase-finally and at the boundary of smaller prosodic units from direct measuring/observation of the voicing source (Pierrehumbert & Talkin 1992, Pierrehumbert 1995) or through qualitative assessment and labeling of waveforms (Dilley, Shattuck-Hufnagel & Ostendorf 1996, Hagen 1997). Additionally, the ends of sentences were associated with a lax voice quality in Swedish (Gobl 1988, using inverse filtering), and a creaky voice quality in Finnish (Ogden 2003, obtained through qualitative assessment and labeling of waveforms). With the exception of Epstein (2002), however, it is difficult to tell from these studies if it is phonological pitch or position that altered phonation.

Each of the aforementioned studies (with the exception of Hindi, Ohala 1973) examined variation in phonation in a language with either allophonic or suprasegmental non-modal phonation. Less attention has been given to this sort of variation in languages with phonemic phonation contrasts, though some gender-dependent differences have been observed. (No study has examined or observed changes in phonation due to sentence-level F<sub>0</sub> and/or position in languages with contrastive phonation.) In Jalapa Mazatec, a language that contrasts breathy, modal and creaky vowels, females were significantly breathier than males in their acoustic manifestation of breathiness (Blankenship 1997) and in San Lucas Quiavini Zapotec, a language that contrasts breathy, modal, creaky, and checked vowels (a vowel followed by a glottal stop with a phonation type distinct from modal and creaky, Munro & Lopez 1999), Gordon & Ladefoged (2001) observed spectrographic evidence that the female speaker was breathier than the male speaker. While there is some evidence that phonation will vary, at least as a function of gender, little is known about potential sources of variation in languages with contrastive phonation.

The present study sets out to investigate variation in phonation due to (i) gender, (ii) F<sub>0</sub>, and/or (iii) prosodic position in Santa Ana del Valle Zapotec, an Oto-Manguean language with phonemically breathy, modal, and creaky vowels.

Before discussing the present study in detail, there will first be a presentation of background information on Santa Ana del Valle Zapotec, then a review of the relevant acoustic measures of phonation, followed by methodology, results and discussion of the data.

## 2 Background

### 2.1 About the language

Santa Ana del Valle Zapotec (hereafter SADVZ) is an Oto-Manguean language spoken in Santa Ana del Valle, Oaxaca, Mexico. The Ethnologue (R. Gordon 2005) classifies SADVZ as belonging to the San Juan Guelavía Zapotec subgroup, which contains the numerous and diverse languages spoken in the Valley of Oaxaca, such as San Juan Guelavía (for which the subgroup is named), Jalieza Zapotec, Teotitlán del Valle Zapotec, and San Martín Tilcajete Zapotec, to name a few. There are approximately 28,000 speakers (1990 census) for the entire San Juan Guelavía Zapotec subgroup; it is not known what portion of this is composed of SADVZ speakers.

SADVZ has six vowels [i e ɨ a u o], each of which can have one of three contrastive phonations: modal, breathy, or creaky. In addition, tone is contrastive on modal vowels, which can have either a high or a rising tone. There is no difference in phonation between the high and rising modal vowels. Breathy and creaky vowels both have a falling tone (Esposito 2003, 2004a). Thus, there is a strong relationship between lexical F0 and phonation in that modal phonations are only produced with high F0s. The question of whether sentence-level F0 will have the same relationship with phonation will be explored in the current study.

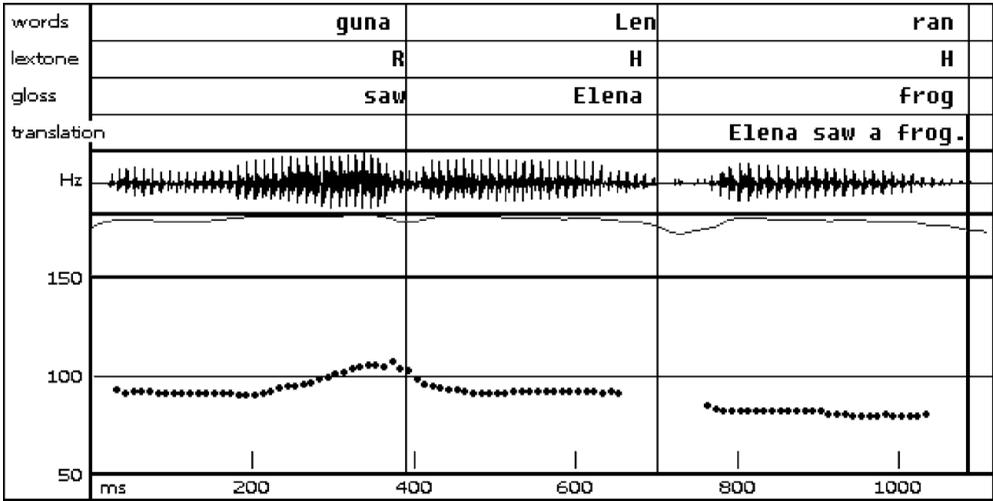
Santa Ana del Valle Zapotec, is a (V)erb–(S)ubject–(O)bject language. An example of a basic declarative sentence with VSO word order is given in (1).

- (1) gu-nǎ            lén    rán  
 PERFECTIVE-see Elena frog  
 ‘Elena saw a frog.’

However, SADVZ also allows SVO and OVS word order (in which case the preverbal material has a focused reading). The relatively free word order of SADVZ makes it an ideal language in which to examine the effects of position on phonation because words can be elicited in a variety of positions. Furthermore, there is no utterance-final drift in sentence-level F0, making it possible to study the effects of position without a strong influence of sentence-level F0. A study<sup>2</sup> of the intonation of SADVZ showed that in sentence-final position in a declarative sentence the original tonal pattern was preserved though the overall pitch value was lowered. For example, if a lexically high tone word was in sentence-final position in a basic declarative, there was a phonological lowering of lexical F0, but the level contour was preserved (i.e. it does not change to a falling F0). This is exemplified in Figure 1, a pitch track of the sentence in example (1); the end of the sentence ends with a low-level F0, and not a falling F0. In Figure 1 and throughout, the SADVZ words are written in IPA (with the exception of the ⟨r⟩ which represents [r]) on the word tier, followed by the lexical tone (R = rising tone, H = high tone), gloss and translation.

It is also possible to examine the potential influence of sentence-level F0 on phonation, independently of position, through the examination of a variety of intonational contours. For example, words in sentence-medial position have the same sentence-level F0 as words uttered in isolation without focus. Note the similarity in the F0 of ‘Elena’ [lén] as pronounced in

<sup>2</sup> The pitch tracks presented in this paper are from Esposito (2004b).



Note: The word 'frog' [rán] has a modal phonation and a phonemically high-level tone, which is lowered here due to position.

Figure 1 Pitch track of [gunã léñ rán] 'Elena saw a frog.'

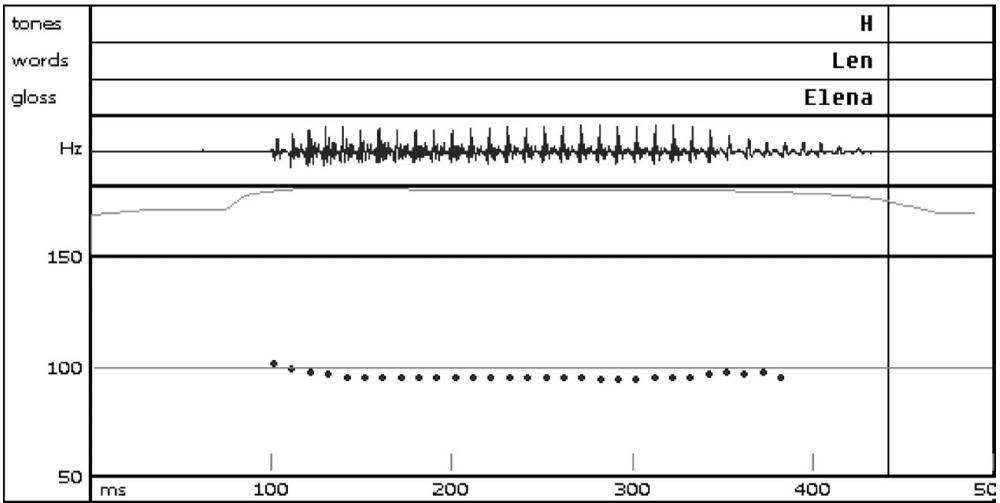


Figure 2 Pitch track of [léñ] 'Elena' produced in isolation without focus.

isolation (without focus) (Figure 2) to the way in which it is pronounced in sentence-medial position (Figure 1).

The independent relationship between sentence-level F0 and position, the relatively free word order, and the lack of an utterance-final drift in F0 make SADVZ an ideal language in which to study variation in phonation.

## 2.2 About measurements

There are numerous acoustic properties that can be useful measures of phonation, though spectral measures have been the most popular method for measuring phonation from an audio

signal. Spectral measures have been a reliable measure of phonation in Hmong (Huffman 1987, Andruski & Ratliff 2000), Mazatec (Kirk, Ladefoged & Ladefoged 1993, Silverman et al. 1995, Blankenship 1997), and !Xóõ (Bickley 1982, Ladefoged 1983, Ladefoged, Maddieson & Jackson 1988), to name a few. Primarily, the difference between the amplitudes of the first and second harmonics (H1-H2) has been used to distinguish phonation (e.g. Fischer-Jørgensen 1967; Bickley 1982). Other studies, however, have made use of the relationship between the amplitude of H1 compared to the amplitude of harmonics exciting higher formants (e.g. A1, A2, A3, and A4). These include: H1-A3 (Stevens & Hanson 1995, Blankenship 1997), H1-A1 or H1-A2 (Ladefoged 1983) and the average of H1-H2 compared to A1 (Stevens 1988).

The majority of studies on linguistically-relevant voice qualities (e.g. Bickley 1982, Ladefoged 1983, Huffman 1987, Ladefoged et al. 1988, Kirk et al. 1993, Silverman et al., 1995, Blankenship 1997) have not made use of corrected or normalized measures, but instead focused on /a/, because the high first formant minimizes the effects on the first and second harmonics. (Andruski & Ratliff 2000 also used uncorrected measures, but expanded their study to /i/ and /u/ in addition to /a/.)

The various spectral measures have been associated with physiological characteristics. Holmberg et al. (1995) showed that the difference between the first harmonic (H1) and the second harmonic (H2) correlated with the proportion of the glottal cycle during which the glottis is open (the open quotient). When the vibration of the vocal folds has a large open quotient, the amplitude of H1 is greater than the amplitude of H2. For example, breathy phonation has a large open quotient and a spectrum dominated by H1.

Furthermore, Stevens (1977) suggested that measures of spectral slope correlated with the abruptness of vocal fold closure. One way that vocal fold vibration can be achieved is with tightly adducted arytenoids, allowing vibration over the anterior portion of the vocal folds. As Stevens (1977: 274) notes, '[w]ithin this region, the mechanical properties of the folds are more uniform, and an abrupt closure along the length of the vibrating portion can be expected. A more rapid rate of closure is also expected, since the inward (adducting) force on the folds is greater when the arytenoids are tightly adducted'. This configuration produces waveforms with more high-frequency energy. For this reason, in creaky phonation, which can be characterized by vocal folds that close rapidly (but also open more slowly), the amplitude of the higher harmonics of the vowel is greater than that of the fundamental.

### 3 Variation in phonation

Gender variation will be examined first, followed by F0 and position.

#### 3.1 Gender variation

To determine if there is a possible gender difference in phonation in SADVZ, males' and females' phonations were compared and measured using two types of spectral measures, one reflecting the open quotient (H1-H2), the other reflecting the speed of vocal fold closure (H1-A3).<sup>3</sup>

##### 3.1.1 Methods

###### 3.1.1.1 Speakers

Five native speakers of SADVZ (three male and two female) were selected for this study. Speakers ranged from 40 to 70 years of age. Speaker 1 is a trilingual male, speaking SADVZ and Spanish natively, in addition to English, which he learned in his late twenties. Speakers 2 and 3 are bilingual males speaking SADVZ and Spanish natively. Speaker 4 is a bilingual

<sup>3</sup> A pilot study (Esposito 2003) compared six measures and determined that H1-A3 was the best spectral tilt measure in SADVZ.

**Table 1** Santa Ana del Valle Zapotec (SADVZ) wordlist. The fortis/lenis obstruents are represented with the symbols for voiceless and voiced consonants, respectively. This is their typical representation in Zapotec languages. The fortis/lenis sonorants are represented by their length contrast.

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[páp]	'potato'	[táp]	'four'	[gâts]	'hidden'
[pág]	'pay'	[bâd]	'Tlacolula'	[nâd]	'to stick'
[fáb]	'detergent'	[sâk]	'price'	[lâts]	'field'
[táp]	'lid'	[bâdɜ]	'tepache'	[dɜâʔp]	'girl'
[lát]	'(tin) can'	[râp]	'care'	[nâʔtʃ]	'rough'
[dâd]	'father'	[nâj]	'chocolate'	[bâd]	'scabies'
[bâd]	'duck'	[lâd]	'clothes'	[ɜân]	'butt'
[fâld]	'skirt'	[nâl]	'thick'	[bâʔt]	'skunk'
[mârkw]	'Mark'	[nân]	'cold (n.)'	[râp]	'have'
[lâʔs]	'skinny'	[dân]	'mountain'	[dâts]	'empty'

female speaking SADVZ and Spanish natively. Speaker 5 is a monolingual SADVZ-speaking female.

### 3.1.1.2 Speech materials

Speakers were asked to produce 10 monosyllabic [a]-vowel words per phonation type (breathy, modal, and creaky) for a total of 30 tokens per speaker (see Table 1 for the wordlist). (For modal phonation, this included five tokens with a high tone and five with a rising tone. Previous research showed that there was not a phonation difference between the high and high-rising tones; see Esposito 2003.) Only words with [a] were selected because the first formant of low vowels does not influence the amplitude of the first or second harmonics as much as in higher vowels. Each token was uttered in sentence-medial position in the frame [guni? \_\_\_ primer] 'Say \_\_\_ first.', and repeated ten times by each speaker. It was not possible to control the onset and coda consonants of the target words. However, pilot research suggested that the coda consonants did not have an effect on phonation (only coda consonants were tested because only the end of the vowels was measured). Paired t-tests indicated that the phonation of vowels produced before lenis consonants was not significantly different from the phonation before fortis consonants ( $df = 58$ ,  $t = -.14$ ,  $p = .88$ ).

### 3.1.1.3 Procedure

Speakers were recorded over multiple sessions in a soundproof booth at the UCLA Phonetics Lab. Tokens were digitized and analyzed in PCQuirer (Scicon R&D, Inc., Encino, CA) at a sampling rate of 22,050 Hz. Each vowel was divided into four equal parts. Two acoustic measures, H1-H2 and H1-A3, were made for each vowel over a 30 ms window in the last quarter of the vowel (approximately 50 ms). Previous research showed that the phonation contrasts were localized to the end of the vowel in SADVZ (Esposito 2003). Spectrograms were used to position the 30 ms window. Measurements were taken from a Fast Fourier Transform (FFT). A total of 1500 tokens were analyzed (5 speakers  $\times$  30 words  $\times$  10 repetitions each = 1500).

## 3.1.2 Results

Figures 3, 4, 5, 6, and 7 are graphs of the average H1-H2 and H1-A3 values per phonation type (breathy, modal, and creaky) for Speakers 1, 2, 3, 4, and 5, respectively. The difference between the amplitudes of the harmonics is given on the y-axis in dB. Greater dB values indicate more breathiness; smaller dB values indicate more creakiness. In the figures in this section and throughout, an arrow is pointing in the direction of increased breathiness and averages are given either above (for positive values) or below (for negative values) each column. Below each figure, the range of values and standard deviation (S.D.) is presented.

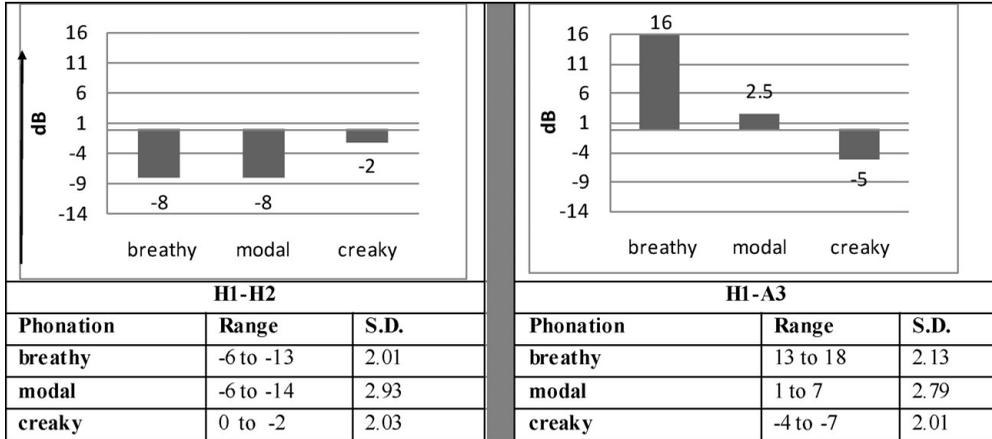


Figure 3 Average H1-H2 and H1-A3 values (in dB) for Speaker 1 (male). The range and standard deviations are presented below the graph.

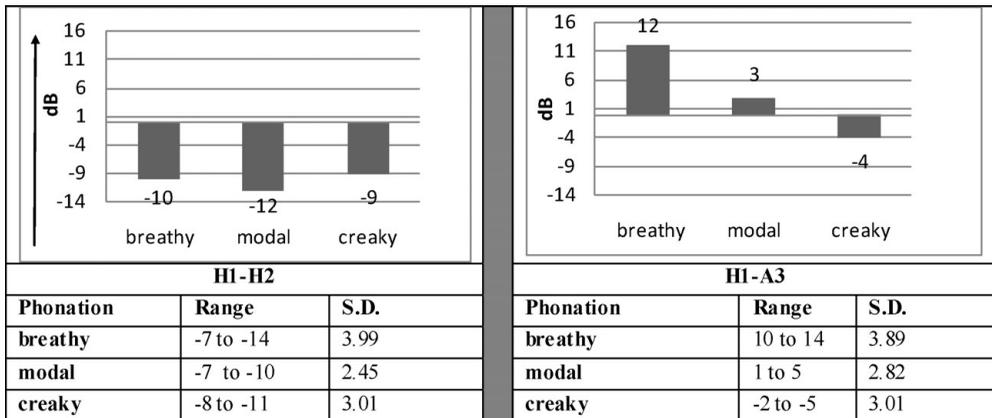


Figure 4 Average H1-H2 and H1-A3 values (in dB) for Speaker 2 (male). The range and standard deviations are presented below the graph.

For male speakers, H1-A3 distinguished the three phonemic phonations in the expected directions (i.e. with breathy phonation having the highest dB value, followed by modal and then creaky), but H1-H2 did not. For all three male speakers, the H1-A3 values for creaky phonation were consistently negative.<sup>4</sup> The creaky vowels for the male speakers were remeasured using H1\*-A3\* (where \* indicates that the measure was corrected) using VoiceSauce (Shue, Keating & Vicens 2009). Results were not significantly different ( $p > .05$ ) and were still negative.

<sup>4</sup> Previous studies using H1-A3 (or H1\*-A3\*, where the asterisk indicates that the measure was corrected) have not reported negative values for this measure. For example, Hanson & Chuang (1999) reported only positive values for H1\*-A3\* for male English speakers producing [ɛ], [æ] and [ʌ]. For [ɛ], the values ranged from 5.7 dB to 23.1 dB; for [æ], 6.2–24.1 dB; and for [ʌ], 4.8–22.8 dB. In addition, Stevens & Hanson (1995) also reported positive values for H1\*-A3\* for female English speakers producing [ɛ], [æ] and [ʌ]. For [ɛ], the values ranged from 17.9 dB to 35.2 dB; for [æ], 18.0–39.1 dB; and for [ʌ], 14.1–33.5 dB. While previous studies have not reported negative values for H1-A3 (or H1\*-A3\*), Blankenship (1997) reported negative values for a similar measure, H1-A2, for laryngealized vowels in Mazatec.

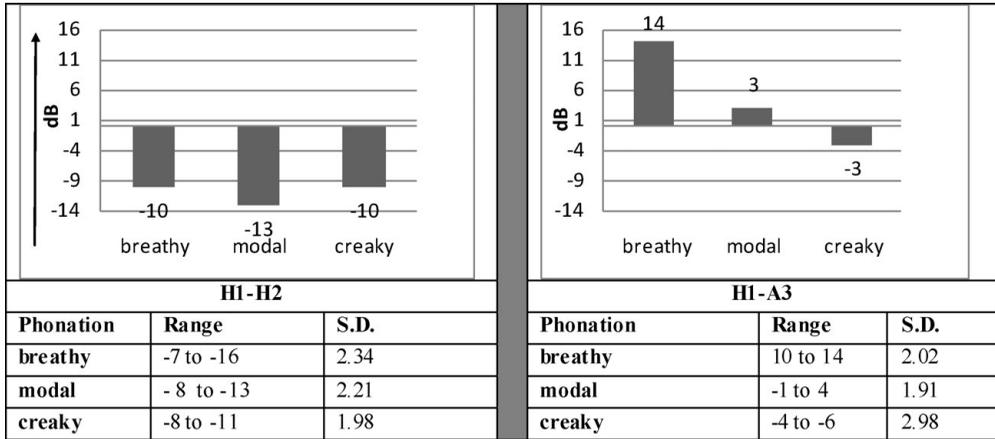


Figure 5 Average H1-H2 and H1-A3 values (in dB) for Speaker 3 (male). The range and standard deviations are presented below the graph.

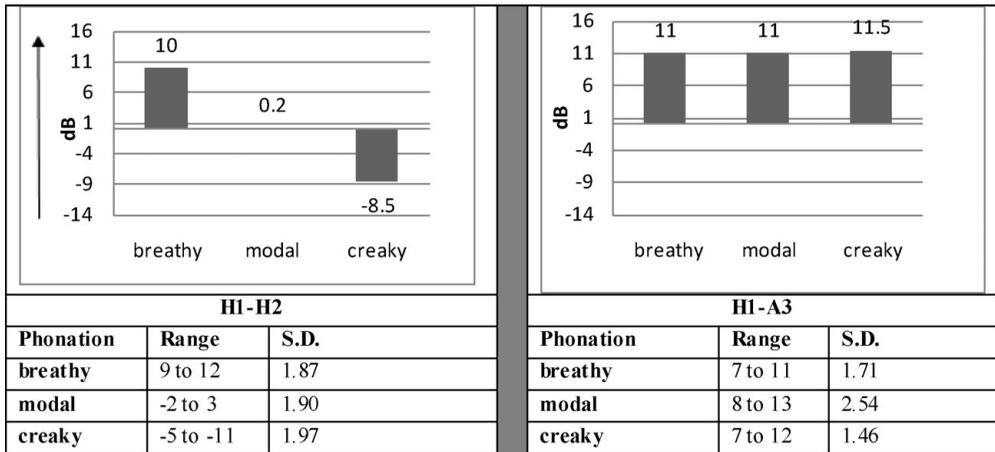


Figure 6 Average H1-H2 and H1-A3 values (in dB) for Speaker 4 (female). The range and standard deviations are presented below the graph.

Figure 8 presents a waveform, spectrogram, and FFT for [lâts] ‘field’ produced by Speaker 2. Note how the amplitude of A3 is higher than H1.

For female speakers, H1-H2 successfully distinguished the three phonemic phonation categories in the expected directions, while H1-A3 did not.

Unfortunately, it is difficult to determine if male SADVZ speakers are breathier/creakier than females (and vice versa) because of the success of the different measures. However, the success of H1-A3 for the male speakers and H1-H2 for the female speakers suggests that men and women are using different laryngeal settings to produce the same linguistic contrast, though direct observations of the vocal folds are needed to support this claim.

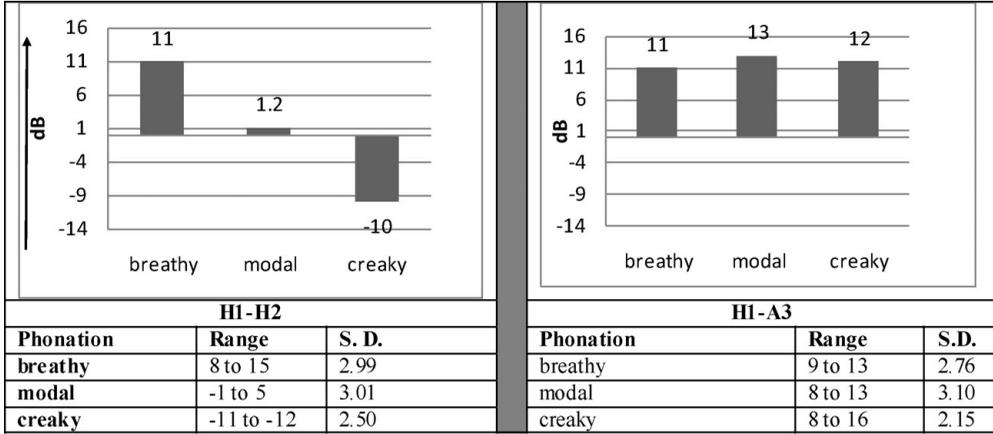


Figure 7 Average H1-H2 and H1-A3 values (in dB) for Speaker 5 (female). The range and standard deviations are presented below the graph.

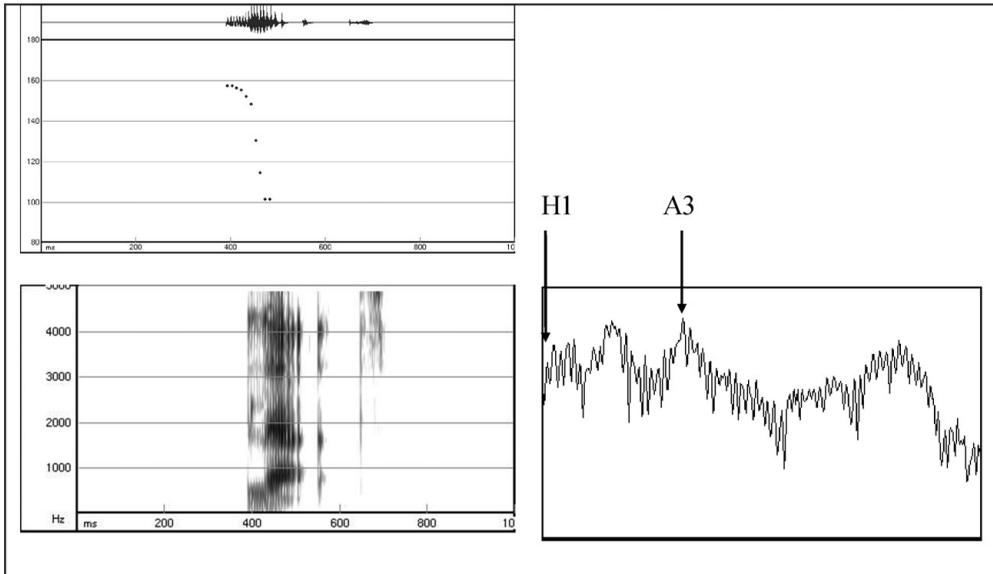


Figure 8 Waveform, spectrogram and FFT (calculated over a 30-ms window) of [lãts] 'field' produced by Speaker 3. Pitch track fails at the end.

### 3.2 Prosodic variability

To determine if SADVZ phonation is sensitive to changes in sentence-level F0 and/or position, acoustic measures (H1-A3 for the male speakers and H1-H2 for the female speakers) were taken of the same words in five prosodic positions, each associated with an F0. The effects of sentence-level F0 and position will be teased apart in the results section.

### 3.2.1 Methods

#### 3.2.1.1 Procedure

The same speakers as those used in the previous task were asked to produce ten [a]-vowel tokens per phonation (as with the previous experiment, this included five tokens with a high tone and five with a rising tone for modal phonation) in the following five prosodic positions:

- (i) isolation (with focus), which has a higher F0 than sentence-medial position<sup>5</sup>
- (ii) initial position (focused), which has a higher F0 than sentence-medial position
- (iii) isolation (without focus), which has a mid-range F0
- (iv) medial position, which has a mid-range F0
- (v) final position (end of a declarative sentence), which has a lower F0 than sentence-medial position

Isolation (with focus) was elicited by asking the speakers a question to which the response was the target word. Isolation (without focus) was elicited by asking the speakers to simply say the word in isolation. Words uttered in isolation and not as an answer to a question were always produced without focus. H1-A3 and H1-H2 were measured from an FFT following the same procedure established above. In addition, F0 was measured at three time-points (beginning, middle, and end).

### 3.2.2 Results

Figures 9, 11, and 13 are graphs of the average H1-A3 value for each phonation category in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) for Speakers 1, 2, and 3, respectively. Figures 15 and 17 are graphs of the average H1-H2 for each phonation category in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) for Speakers 4 and 5, respectively. The difference between the amplitudes of the harmonics is given on the y-axis in dB. The range and standard deviations are presented under the graphs. Figures 10, 12, 14, 16, and 18 are graphs of the average F0 (in Hz) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) for Speakers 1, 2, 3, 4, and 5, respectively. In the graphs of the F0, the high and rising tone modal vowels are separated.

In all five prosodic positions, there is a three-way contrast in phonation. This contrast, however, is not always well-defined. In isolation (mid-range F0), medial position (mid-range F0) and final position (lower F0), the three-way phonation contrast is clearest for both the male and the female speakers (but with some changes for the female speakers, whose modal vowels were creakier word-finally than in medial position). The contrast was minimized in isolation with focus (high F0) and initial position (high F0). In these two positions, breathy and creaky vowels had a much more modal phonation than when produced in isolation (mid-range F0), medial position (mid-range F0) or final position (lower F0).

There is evidence that it is F0, independent of position, that is influencing phonation in SADVZ. Tokens with the same position, but different F0s, have different phonations. More specifically, tokens in isolation can be produced with either a high F0 (when focused) or a mid-range F0 (when not focused). It was only when the tokens were produced with the high F0, however, that the three-way contrast in phonation was minimized. Furthermore, tokens

<sup>5</sup> Words in isolation can be elicited with two sentence-level F0s; one high and one mid-range. While technically not two separate positions, both were elicited because they were helpful in determining if F0 influences phonation independently of position.

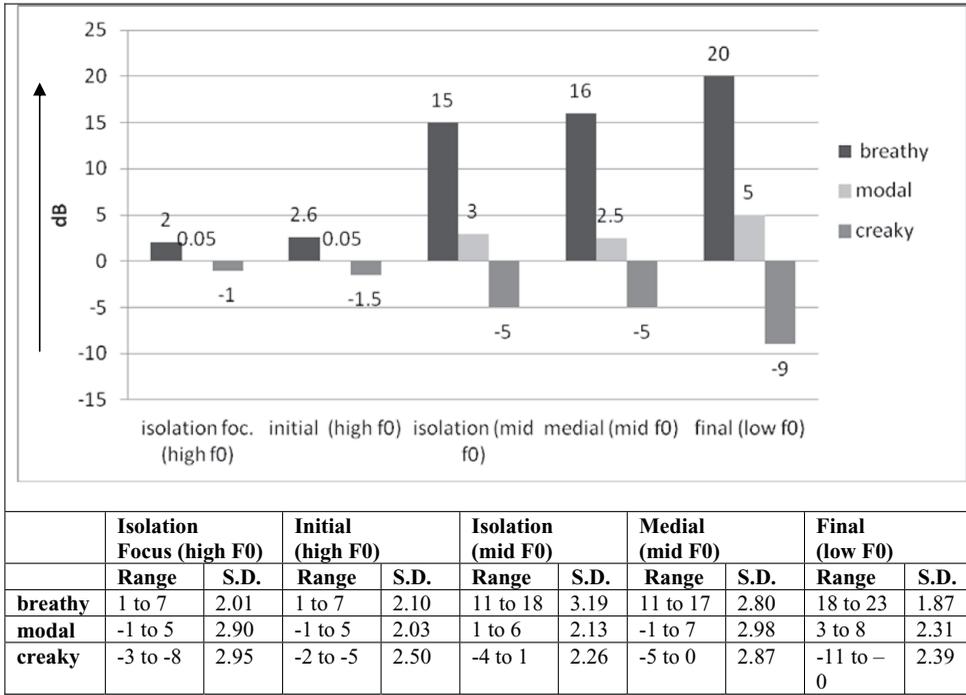


Figure 9 Average H1-A3 values for breathy, modal, and creaky phonation in five prosodic positions for Speaker 1 (male). The range and standard deviations are presented below the graph.

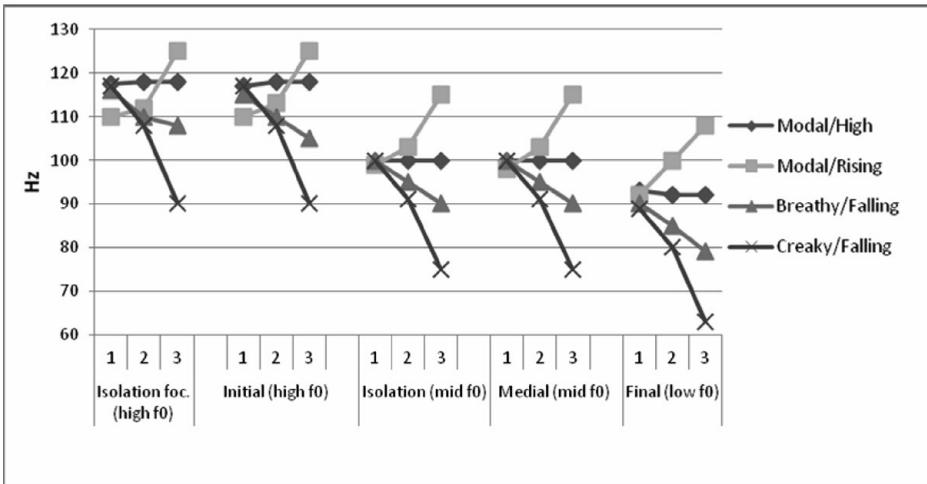
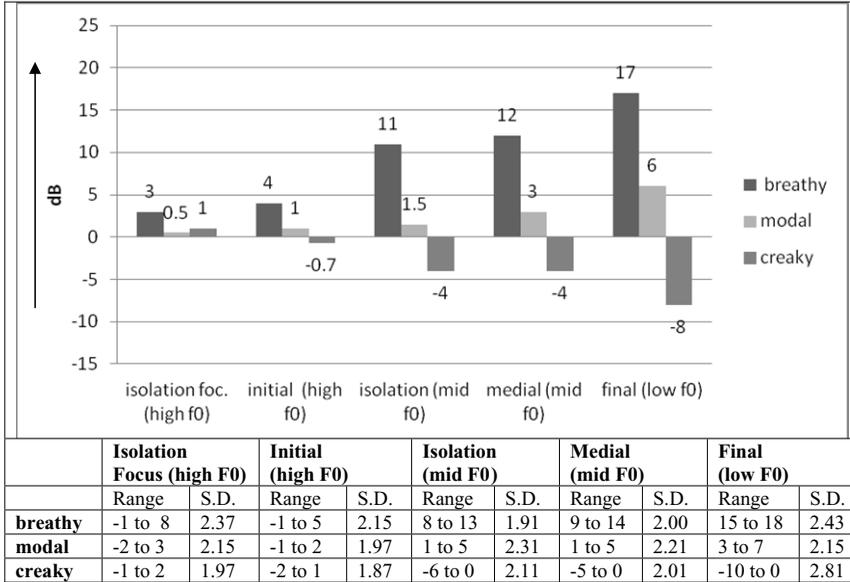
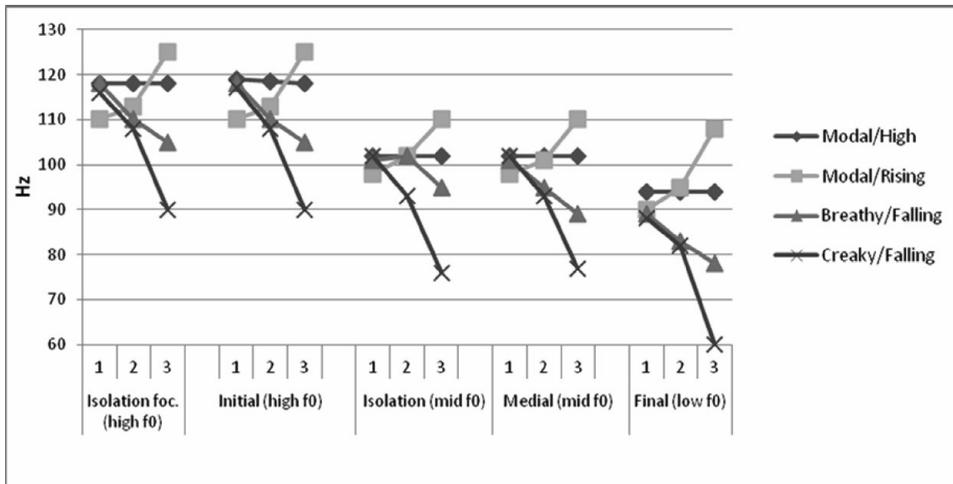


Figure 10 Average F0 for Speaker 1 (male) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) at three timepoints (1, 2, 3) per vowel.

with the same F0 but different positions have similar phonations: in positions with mid-range F0 (i.e. isolation (non-focused) and sentence-medial position), there was a clear three-way contrast in phonation.



**Figure 11** Average H1-A3 values for breathy, modal, and creaky phonation in five prosodic positions for Speaker 2 (male). The range and standard deviations are presented below the graph.



**Figure 12** Average F0 for Speaker 2 (male) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) at three timepoints (1, 2, 3) per vowel.

### 4 Discussion and conclusion

The current study set out to determine if phonation would vary as a function of (i) gender, (ii) F0, and/or (iii) position in Santa Ana del Valle Zapotec, a language where differences in phonation convey important linguistic information.

Results were inconclusive as to whether or not one gender was creakier or breathier than the other because of the success of different acoustic measures for the two genders. However,

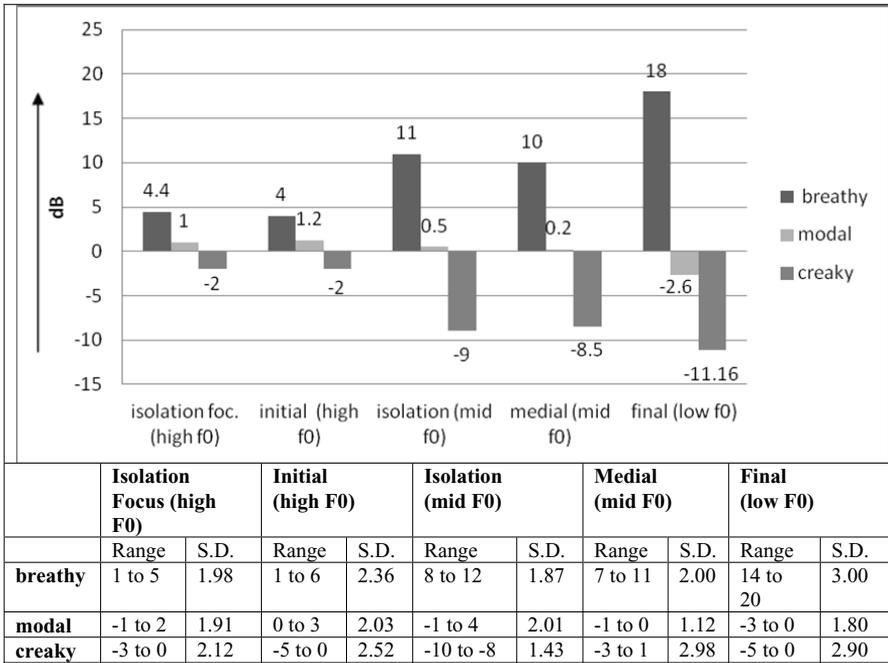


Figure 13 Average H1-A3 values for breathy, modal, and creaky phonations in five prosodic positions for Speaker 3 (male). The range and standard deviations are presented below the graph.

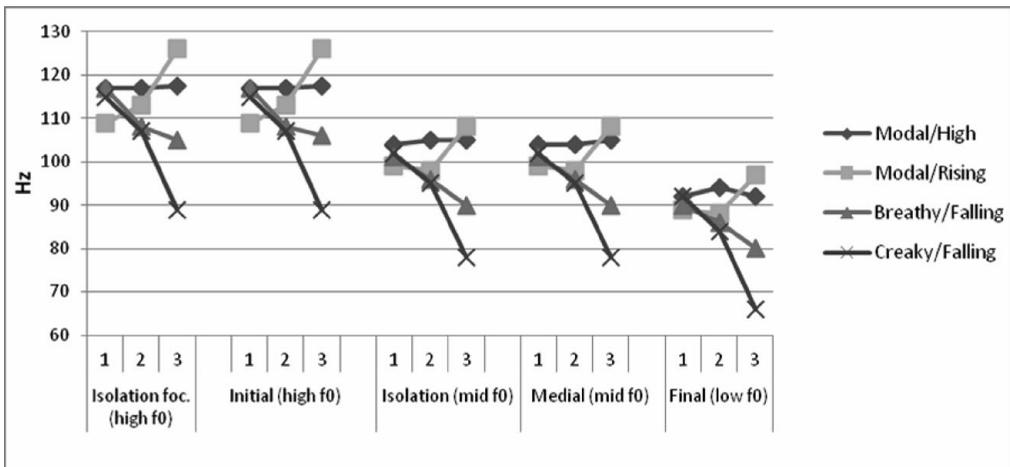


Figure 14 Average F0 for Speaker 3 (male) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) at three timepoints (1, 2, 3) per vowel.

the success of H1-H2 for females and H1-A3 for males suggests that there is a difference in the PRODUCTION of phonation between the genders, though the reason for this difference remains unknown. Presumably, all healthy humans are capable of producing all possible articulations; perhaps the gender-based phonation differences in SADVZ are determined by sociolinguistic factors. However, Södersten & Lindestad (1990) did show that female speakers (of English)

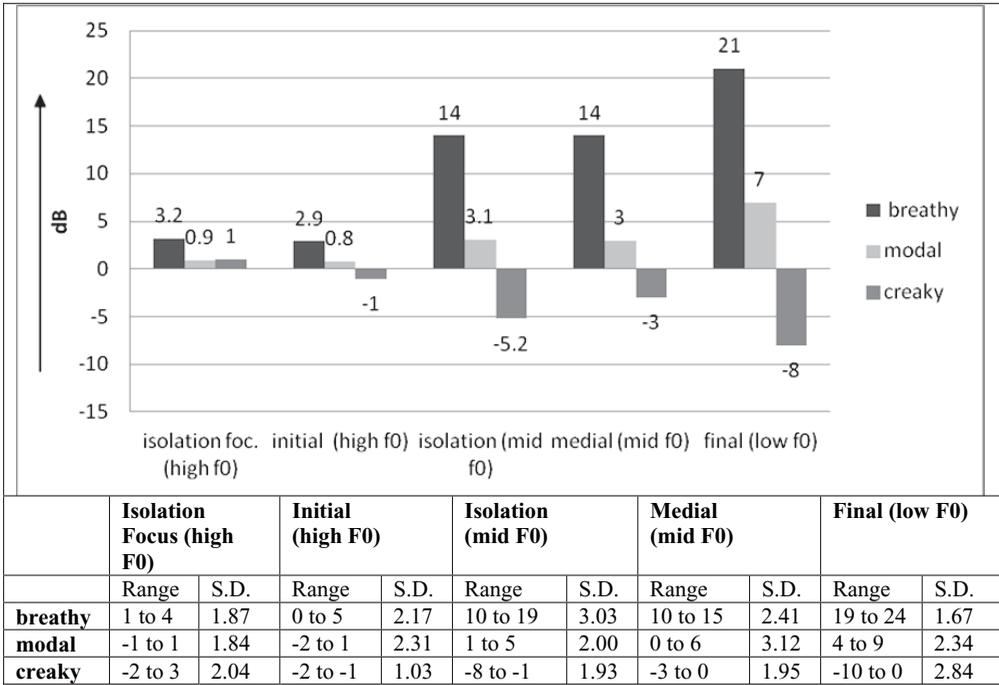


Figure 15 Average H1-H2 values for breathy, modal, and creaky phonations in five prosodic positions for Speaker 4 (female). The range and standard deviations are presented below the graph.

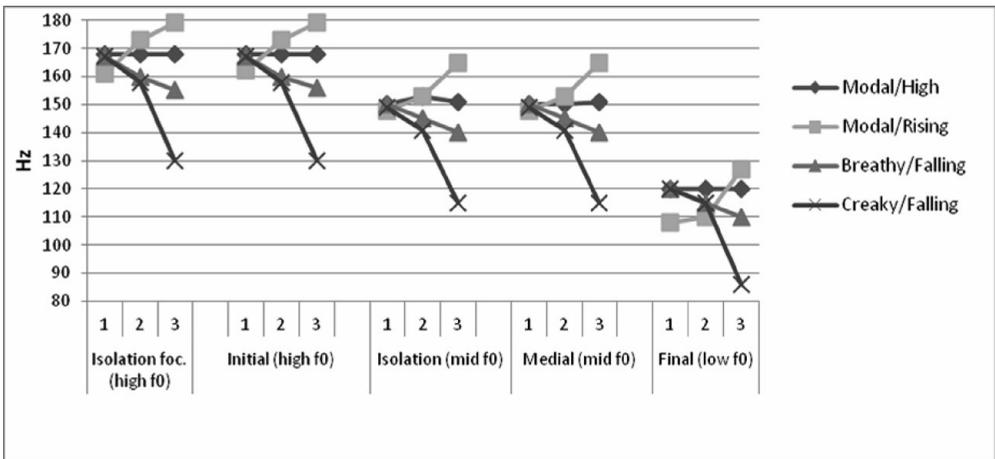


Figure 16 Average F0 for Speaker 4 (female) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) at three timepoints (1, 2, 3) per vowel.

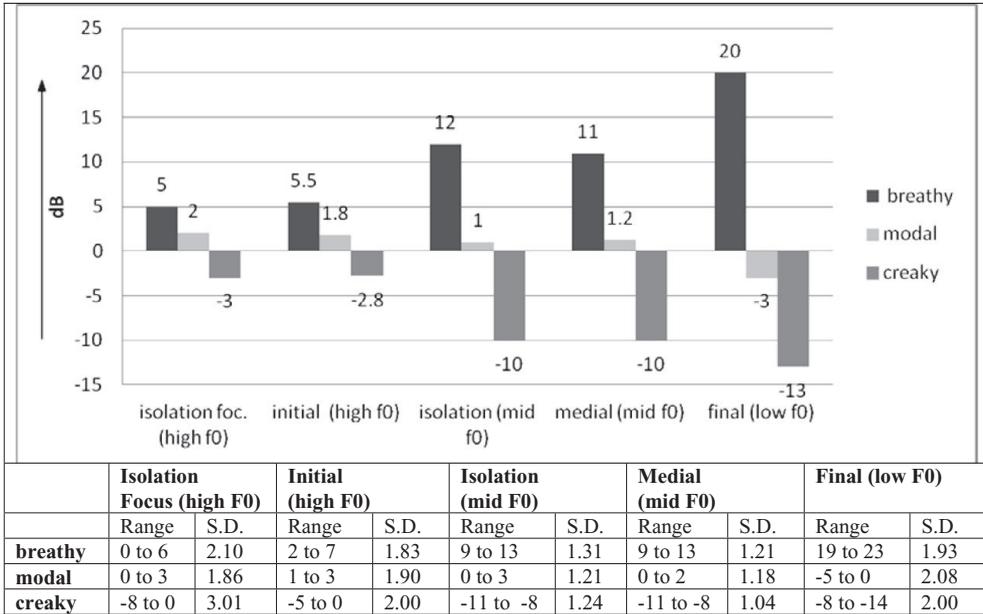


Figure 17 Average H1-H2 values for breathy, modal, and creaky phonations in five prosodic positions for Speaker 5 (female). The range and standard deviations are presented below the graph.

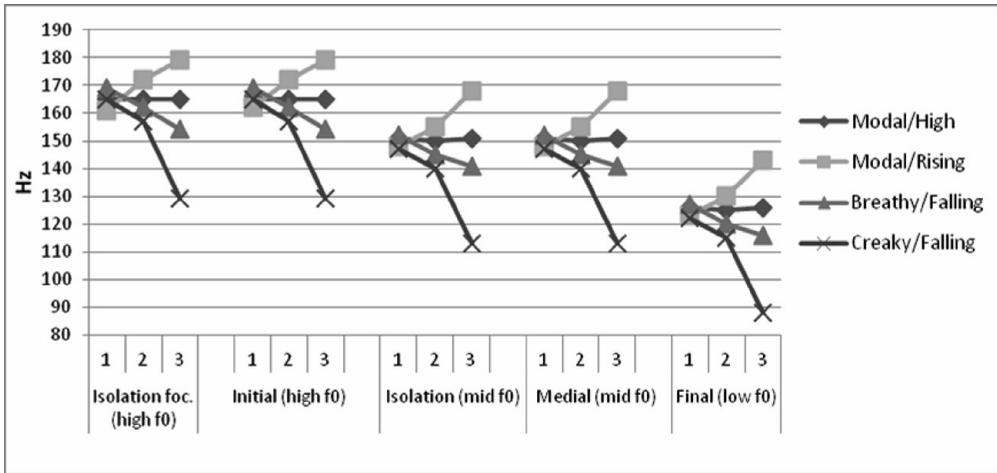


Figure 18 Average F0 for Speaker 5 (female) in five prosodic positions (isolation with focus, initial, isolation (not focused), medial, and final position) at three timepoints (1, 2, 3) per vowel.

were more likely to have incomplete glottal closure. Thus, another possible explanation for the gender-based differences could be a physiological one, though direct observations of the vocal folds, the larynx and the valves of the throat are needed to verify this hypothesis. Furthermore, these results raise the question of whether or not SADVZ listeners can perceive the difference between H1-H2 and H1-A3, and whether or not these differences are associated

with gender differences (i.e. do H1-H2 differences sound ‘feminine’ while H1-A3 differences sound ‘masculine’?). A perception study investigating these issues would be an important follow-up to the current study.

Additional findings showed that F0, independent of position, had a strong effect on phonation. When F0 was high (i.e. isolation and initial position), the three-way contrast in phonation was minimized, with non-modal phonation having an H1-H2/A3 value typically associated with modal phonation. (While the H1-H2/A3 values at high F0 indicate a modal-like phonation, it is possible that the creaky phonation is not actually more modal, but rather a type of harsh or pressed voice.) The same pattern is also seen with lexical F0; in SADVZ modal phonation is produced with high F0s and non-modal phonations are produced with lower F0s. Cross-linguistically, high F0, which can increase vocal fold tension and length, is associated with modal phonation (though F0 is not always correlated with phonation; see Ladefoged 1973, Laver 1980). This glottal configuration (i.e. increased vocal fold tension and length) is the opposite of the typical configuration for breathy phonation, which could explain why speakers are producing phonemically breathy phonations with H1-H2/A3 values typically associated with modal phonation in environments with high F0. Furthermore, creaky phonation can also lose its characteristic creak at high F0s. Creaky phonation (with fast closure and long and slow vocal fold opening) is produced when the aryepiglottic sphincter is engaged (Esling 2005). This shortens the length of the glottis and of the epilaryngeal structures over the glottis. When F0 is high, the larynx will lengthen longitudinally, stretching the vocal folds in the opposite dimension of the typical configuration for creaky phonation. (For more information on the relationship between phonation type and pitch, see Esling & Harris 2005.)

The results obtained in this study raise an interesting issue for fieldworkers working on Zapotec languages, and perhaps languages with phonation contrasts in general. In Santa Ana Del Valle Zapotec, when F0 is high, the phonemic three-way contrast in phonation is not well preserved. Thus, words elicited in isolation with focus, which have a higher F0 than sentence-medial position, will not show an obvious phonation contrast. In order to see a full range of phonation contrasts, it is important to elicit data that display a full range of F0s.

The current study shows that there is variation in phonation, even in languages with contrastive phonation. However, it remains to be seen how SADVZ compares with other languages with contrastive phonation. In order to truly understand the nature of the type of variation demonstrated in this study, it is necessary to replicate these results with other languages that contrast phonation and to elicit data in more natural discourse contexts.

## Acknowledgements

This article derives from a larger research project guided by Matthew Gordon, Patricia Keating, and Sun-Ah Jun. I thank them for their time and guidance. Thanks also to my Zapotec consultants for providing the data used in this paper.

## References

- Andruski, Jean & 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.
- Bickley, Corine. 1982. Acoustic analysis and perception of breathy vowels. *Speech Communication Group Working Papers* 1, 73–93. Cambridge, MA: MIT.
- Blankenship, Barbara. 1997. *The time course of breathiness and laryngealization in vowels*. Ph.D. dissertation, University of California, Los Angeles.
- Dilley, Laura, Stephanie Shattuck-Hufnagel & M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24, 423–444.

- Edmondson, Jerold A. & John H. Esling. 2006. The valves of the throat and their functioning in tone, vocal register, and stress: Laryngoscopic case studies. *Phonology* 23(2), 157–191.
- Epstein, Melissa. 2002. *Voice quality and prosody in English*. Ph.D. dissertation, University of California, Los Angeles.
- Esling, John H. 2005. There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics* 50, 13–44.
- Esling, John H. & Jimmy Harris. 2005. States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In William J. Hardcastle & Janet Beck (eds.), *A figure of speech: A Festschrift for John Laver*, 347–383. Mahwah, NJ: Lawrence Erlbaum.
- Esposito, Christina M. 2002. Pilot study: Santa Ana del Valle Zapotec Phonation. Ms., University of California, Los Angeles.
- Esposito, Christina M. 2003. Santa Ana del Valle Zapotec phonation. MA thesis, University of California, Los Angeles.
- Esposito, Christina M. 2004a. Santa Ana del Valle Zapotec phonation. *UCLA Working Papers in Phonetics* 103, 71–105.
- Esposito, Christina M. 2004b. Santa Ana del Valle Zapotec Intonation. Ms., Macalester College.
- Fischer-Jørgensen, Eli. 1967. Phonetic analysis of breathy (murmured) vowels. *Indian Linguistics* 28, 71–139.
- Fujimura, Osamu (ed.). 1988. *Vocal physiology: Voice production, mechanisms and functions*, 297–317. New York: Raven Press.
- Fujimura, Osamu & Minoru Hirano (eds.). 1995. *Vocal fold physiology: Voice quality control*. San Diego, CA: Singular Publishing Group.
- Gobl, Christopher. 1988. Voice source dynamics in connected speech. *STL-QPSR* 1, 123–159.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.
- Gordon, Raymond, Jr. (ed.). 2005. *Ethnologue: Languages of the world*, 15th edn. Dallas, TX: SIL International. <http://www.ethnologue.com/web.asp> (March 2009).
- Hagen, Astrid. 1997. *Linguistic functions of glottalizations and their language specific use in English and German*. Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg & MIT.
- Hanson, Helen. 1997. Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America* 101(1), 466–481.
- Hanson, Helen & Erika Chuang. 1999. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America* 106(2), 1064–1077.
- Henton, Caroline & R. Anthony Bladon. 1985. Breathiness in normal female speech: Inefficiency versus desirability. *Language and Communication* 5(3), 221–227.
- Holmberg, Eva B., Roger E. Hillman, Joseph Perkell, Peter Guiod & Susan L. Goldman. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research* 38, 1212–1223.
- Huffman, Marie. 1987. Measures of phonation type in Hmong. *Journal of the Acoustical Society of America* 81(2), 495–504.
- Kirk, Paul L., 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.
- Klatt, Dennis & Laura Klatt. 1990. Analysis, synthesis and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87(2), 820–857.
- Kreiman, Jody. 1982. Perception of sentences and paragraph boundaries in natural conversation. *Journal of Phonetics* 10, 163–175.
- Ladefoged, Peter. 1971 *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.
- Ladefoged, Peter. 1983. The linguistic use of different phonation types. In Diane Bless & James Abbs (eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 351–360. San Diego, CA: College-Hill Press.

- Ladefoged, Peter, Ian Maddieson & Michael Jackson. 1988. Investigating phonation types in different languages. In Fujimura (ed.), 297–317.
- Laver, John. 1981. *Phonetic description of voice quality*. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1975. The phonetic structure of paragraphs. In Antonie Cohen & Sibout Nooteboom (eds.), *Structure and process in speech perception*, 195–203. Heidelberg & New York: Springer.
- Maddieson, Ian & Susan Hess. 1987. The effect of F0 on linguistics use of phonation types. *UCLA Working Papers in Phonetics* 67, 112–118.
- Munro, Pamela & Felipe Lopez, with Olivia Mendez, Rodrigo Garcia & Michael Galant. 1999. *Di'syonaary X:tee'n Diizh Sah Sann Lu'uc (San Lucas Quiavini Dictionary/Diccionario Zapoteco de San Lucas Quiavini)*. Los Angeles, CA: Chicano Studies Research Center Publications.
- Ogden, Richard. 2003. Voice quality as a resource for the management of turn-taking in Finnish talk-in-interaction. *15th International Conference of Phonetic Sciences (ICPhS XV)*, Barcelona, 123–126.
- Ohala, John. 1973. The physiology of tone. *Southern California Occasional Papers in Linguistics* 1, 1–14.
- Pierrehumbert, Janet. 1995. Prosodic effects on glottal allophones. In Fujimura & Hirano (eds.), 39–60.
- Pierrehumbert, Janet & David Talkin. 1992. Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody*, 90–117. Cambridge: Cambridge University Press.
- Redi, Laura & Stephanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407–429.
- Shue, Yen, Patricia Keating & Chad Vicenik. 2009. VoiceSauce: A program for voice analysis. *Journal of the Acoustical Society America* 124(4), 2221.
- Silverman, Daniel, Barbara Blankenship, Paul Kirk & Peter Ladefoged. 1995. Phonetic structures of Jalapa Mazatec. *Anthropological Linguistics* 37, 70–88.
- Södersten Maria & Per-Åke Lindestad. 1990. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hearing Research* 33(3), 601–611.
- Stevens, Kenneth. 1977. Physics of laryngeal behavior and larynx modes. *Phonetica* 34, 264–279.
- Stevens, Kenneth. 1988. Modes of vocal fold vibration based on a two-section model. In Fujimura (ed.), 357–371.
- Stevens, Kenneth & Helen Hanson. 1995. Classification of glottal vibration from acoustic measurements. In Fujimura & Hirano (eds.), 147–170.
- Todaka, Yuichi. 1993. *A cross-language study of voice quality: Bilingual Japanese and American English speakers*. Ph.D. dissertation, University of California, Los Angeles.