

UNIVERSITY OF CALIFORNIA

Los Angeles

Santa Ana Del Valle Zapotec Phonation

A thesis submitted in partial satisfaction
of the requirements for the degree Master of Arts
in Linguistics

by

Christina Marie Esposito

2003

The thesis of Christina Marie Esposito is approved.

Sun-Ah Jun

Matthew Gordon

Patricia A. Keating, Committee Chair

University of California, Los Angeles

2003

Table of Contents

List of Figures	v
1 Introduction	1
1.1 Phonemic Inventory	2
1.1.1 Consonants	2
1.1.2 Vowels	4
1.2 Phonation and Tone	4
1.3 Syntax	4
1.3.1 Basic Word Order	4
1.3.2 Focus	5
1.3.3 Negation	5
1.4 Some Basics of SADVZ Intonation	5
1.4.1 Basic Declaratives	6
1.4.2 Focus	6
1.4.3 Isolation	8
1.4.4 Negative Yes/No Questions	8
1.5 Structure of SADVZ Words	9
1.6 Influence of Spanish	10
2 Phonation	11
2.1 Overview of Phonation and Tone in Zapotec Languages	12
2.2 Overview of Phonation across Languages	15
2.2.1 Localization of Non-Modal Phonation	16
2.2.2 Duration of Vowels with Non-Modal Phonation	17
3 Measuring Non-Modal Phonation	18
3.1 Periodicity	18
3.2 Acoustic Intensity	19
3.3 Spectral Noise	19
3.4 Spectral Tilt	19
4 Experimental Comparison of Phonation in SADVZ	20
4.1 Selecting a Measurement of Phonation	21
4.1.1 Spectral Measurements	21
4.2 Phonation in SADVZ	28
4.2.1 Methods	28
4.2.2 Changes in Phonation within a Syllable (Contour Phonations)	29
4.3 Three phonations (Modal, Breathy, Creaky)	31
4.3.1 Other Differences Among and Within Phonations	34
4.4 Summary of Results	39
4.4.1 Modal	40
4.4.2 Breathy	41
4.4.3 Creaky	43
5 Testing other Vowels	45
5.1 Method	46
5.1.1 Speech Materials	46

5.1.2	Procedure	46
5.2	Results.....	46
5.2.1	Modal	49
5.2.2	Breathy.....	51
5.2.3	Creaky.....	52
6	Polymorphemic, Disyllabic Words.....	53
6.1	Methods.....	53
6.2	Speech Materials.....	53
6.3	Procedure	54
6.4	Results.....	54
6.4.1	Other Differences.....	55
6.4.2	Modal	57
6.4.3	Breathy.....	59
6.4.4	Creaky.....	60
7	Summary.....	61
8	Effects of f_0 and Position on Phonation	62
8.1	Effects of F_0 versus Position	63
8.1.1	Method: Procedure.....	63
8.1.2	Results.....	64
9	The Effects of F_0 on Phonation	66
9.1	Methods.....	66
9.1.1	Speakers and Speech Materials.....	66
9.1.2	Procedure	67
9.2	Results for Isolation and Initial Position.....	67
9.2.1	Isolation.....	67
9.2.2	Initial (Focused) Position.....	69
9.2.3	Breathy.....	72
9.2.4	Creaky.....	74
9.3	Results for Final Position.....	75
9.3.1	Modal	78
9.3.2	Breathy.....	80
9.3.3	Creaky.....	82
10	Summary and Conclusion.....	83
	Appendix.....	87
	References.....	90

List of Figures

Figure 1: Pitchtrack of [ran elen bág] “Elena sees a cow.”	6
Figure 2: [bág guna elen] “Elena saw a COW./ A COW saw Elena.”	7
Figure 3: [elén gudau limoni] “ELENA eats lemons.”	7
Figure 4: [bág] ‘cow’	8
Figure 5: [teka nijau bed bág] “Didn’t Peter see the cow?”	9
Figure 6: Waveform, spectrogram and FFT of ‘can’ [lat] produced by Speaker 1. Pitch track fails at end.	23
Figure 7: Waveform, spectrogram and FFT of ‘place’ [lat] produced by Speaker 1	24
Figure 8: Waveform, spectrogram and FFT of ‘field’ [lat] produced by Speaker 1. Pitch track fails at end.	24
Figure 9: Waveform, spectrogram and FFT of ‘can’ [lat] produced by Speaker 3.	25
Figure 10: Waveform, spectrogram and FFT of ‘place’ [lat] produced by Speaker 3. Pitch track fails at end.	25
Figure 11: Waveform, spectrogram and FFT of ‘field’ [lat] produced by Speaker 3. Pitch track fails at end.	26
Figure 12: Frequency of H1-F3 values for Speaker 1 (initial portion of vowel).	30
Figure 13: Frequency of H1-F3 values for Speaker 1 (medial portion of vowel).	30
Figure 14: Frequency of H1-F3 values for Speaker 1 (final portion of vowel).	31
Figure 15: H1-F3 for Speaker 1 (sentence-medial position)	32
Figure 16: H1-F3 for Speaker 2 (sentence-medial position)	32
Figure 17: H1-H2 for Speaker 3 (sentence-medial position)	33
Figure 18: F0 for Speaker 1 (sentence-medial position)	35
Figure 19: F0 for Speaker 2 (sentence-medial position)	35
Figure 20: F0 for Speaker 3 (sentence-medial position)	36
Figure 21: H1-F3 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 1)	36
Figure 22: H1-F3 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 2)	37
Figure 23: H1-H2 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 3)	37
Figure 24: Spectrogram, waveform, pitchtrack and FFT of ‘mat’ [daʔa] Pitchtrack fails at end.	38
Figure 25: Duration of Modal and Non-Modal Phonation (Speaker 1)	39
Figure 26: Spectrogram, waveform and pitch track for a modal vowel with level tone ‘lid’ [táp] . Uttered in sentence-medial position by Speaker 1.	40
Figure 27: Spectrogram, waveform and pitch track for a modal vowel with high rising tone ‘father’ [däd] Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end.	41

Figure 28: Spectrogram, waveform and pitch track for a breathy vowel ‘four’ [t̪ap]. Uttered in sentence medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.	42
Figure 29: Spectrogram, waveform and pitch track for a breathy vowel ‘four’ [t̪ap]. Uttered in sentence medial position by Speaker 3. Pitchtrack fails at end.	43
Figure 30: Spectrogram, waveform and pitch track for a creaky vowel ‘field’ [l̪ats]. Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end.	44
Figure 31: Spectrogram, waveform and pitch track for a creaky vowel ‘field’ [l̪ats]. Uttered in sentence-medial position by Speaker 3. Pitchtrack fails at end. There is a marker at 500 ms, which is not significant for this study.	45
Figure 32: H1-F3 for [e] (Speaker 1, medial position)	47
Figure 33: F0 for [e] (Speaker 1, medial position)	47
Figure 34: H1-F3 for [u] (Speaker 1, medial position).....	47
Figure 35: F0 for [u] (Speaker 1, medial position)	48
Figure 36: H1-F3 for [i] (Speaker 1, medial position).....	48
Figure 37: F0 for [i] (Speaker 1, medial position)	48
Figure 38: H1-F3 for [o] (Speaker 1, medial position).....	49
Figure 39: F0 for [o] (Speaker 1, medial position)	49
Figure 40: Spectrogram, waveform and pitch track for a modal vowel with a level tone ‘Lucas’ [lúk] Uttered in sentence-medial position by Speaker 1. (The f0 is level during the vowel and low during the [l]).	50
Figure 41: Spectrogram, waveform and pitch track for a modal vowel with a high rising tone ‘gopher’ [lũʒ] Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.	51
Figure 42: Spectrogram, waveform and pitch track for a breathy vowel in ‘cough’ [r̪ù]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.	52
Figure 43: Spectrogram, waveform and pitch track for a creaky vowel ‘you’ [l̪ù]. Uttered in sentence medial position by Speaker 1. Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.	53
Figure 44: Graph of H1-H3 for the First and Second Syllables (Speaker 1, medial position)	54
Figure 45: Duration for the First and Second Syllable (in ms) (Speaker 1, medial position)	55
Figure 46: F0 for the First Syllable in Disyllabic Words. The phonation given in the legend corresponds to the phonation of the second syllable. (Speaker 1, medial position)	56
Figure 47: F0 for the Second Syllable in Disyllabic Words (Speaker 1, medial position).....	56
Figure 48: Spectrogram, waveform and pitch track for ‘Santa Ana’ [ʃandán]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.....	58

Figure 49: Spectrogram, waveform and pitch track for ‘cheap’ [ba.rǎt]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.....	59
Figure 50: Spectrogram, waveform and pitch track for ‘itches’ [rabàb]. Uttered in sentence-medial position by Speaker 1. Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.	60
Figure 51: Spectrogram, waveform and pitch track for ‘warm’ warm’ [nadʒâ]. Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end. There is a marker at 500 ms, which is not significant for this study.	61
Figure 52: Results of H1-F3. High f0, different position-in-utterance (Speaker 1)	65
Figure 53: Results of H1-H2. High f0, different position-in-utterance (Speaker 3).....	65
Figure 54: Results of H1-F3. Different f0, final position-in-utterance (Speaker 1)	65
Figure 55: Results of H1-F3. Different f0, final position-in-utterance (Speaker 3)	66
Figure 56: H1-F3 (isolation, Speaker 1).	68
Figure 57: F0 (isolation, Speaker 1).	68
Figure 58: H1-F3 (isolation, Speaker 2).	68
Figure 59: F0 (isolation, Speaker 2).	69
Figure 60: H1-H2 (isolation, Speaker 3).....	69
Figure 61: F0 (isolation, Speaker 3).	69
Figure 62: H1-F3 (initial (focused) position, Speaker 1).....	70
Figure 63: F0 (initial (focused) position, Speaker 1).....	70
Figure 64: H1-F3 (initial (focused) position, Speaker 2).....	70
Figure 65: F0 (initial (focused) position, Speaker 2).....	71
Figure 66: H1-H2 (initial (focused) position, Speaker 3).....	71
Figure 67: F0 (initial (focused) position, Speaker 3).....	71
Figure 68: Spectrogram, waveform and pitch track for a breathy token ‘four’ four’ [tâp] pronounced in isolation (Speaker 1). Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.	73
Figure 69: Spectrogram, waveform and pitch track for a breathy token ‘four’ [tâp] pronounced in isolation (Speaker 3).	74
Figure 70: Spectrogram, waveform and pitch track for a breathy token ‘field’ [lâts] pronounced in isolation (Speaker 1). There is a marker at 500 ms, which is not significant for this study.....	75
Figure 71: H1-F3 for [a] (final position, Speaker 1).....	76
Figure 72: F0 (final position, Speaker 1).....	76
Figure 73: H1-F3 for [a] (final position, Speaker 2).....	76
Figure 74: F0 (final position, Speaker 2).....	77
Figure 75: H1-H2 for [a] (final position, Speaker 3).....	77
Figure 76: F0 (final position, Speaker 3).....	77
Figure 77: Spectrogram, waveform and pitch track for a breathy token ‘lid’ [táp] pronounced in final position (Speaker 1). Pitch track fails at end.	79

Figure 78: Spectrogram, waveform and pitch track for a breathy token ‘lid’ [táp] pronounced in final position (Speaker 3). Pitch track fails completely.	80
Figure 79: Spectrogram, waveform and pitch track for a breathy token ‘four’ [təp] pronounced in final position (Speaker 1). Pitch track fails at end.	81
Figure 80: Spectrogram, waveform and pitch track for a breathy token ‘four’[təp] pronounced in final position (Speaker 3). Pitch track fails at end.	82
Figure 81: Spectrogram, waveform and pitch track for a breathy token ‘field’ [lâts] pronounced in final position (Speaker 1). Pitch track fails completely.	83

Acknowledgements

This thesis could not have been possible without the help of many people. I would like to thank my consultants, Fantino Aquino, and Fantino and Inés Gutierrez. I would also like to thank my committee: Patricia Keating, Sun-Ah Jun, and Matt Gordon for their time and support. In addition, I would like to thank Argelia Andrade, G. Aaron Broadwell, Olivia Martínez and Pamela Munro, who all have worked on SADVZ, at various times.

ABSTRACT OF THE THESIS

Santa Ana Del Valle Zapotec Phonation

by

Christina Marie Esposito

Master of Arts in Linguistics

University of California, Los Angeles, 2003

Professor Patricia A. Keating, Chair

Accounts of the phonations of Zapotec languages describe more interesting phonations than expected. The goal of this study is to investigate the phonation and tones in Santa Ana del Valle Zapotec, experimentally.

The first part of the study defines the phonations and tones of SADVZ. Speakers produced words that could exhibit a more complex phonation contrast than simply breathy, modal and creaky. H1-F3 was the measure of phonation for the male speakers; H1-H2, for female. Results showed that SADVZ's three-way contrast of modal, breathy and creaky vowels is distinguished acoustically in the directions expected. No change in phonation within a syllable was found. Each of the phonations is associated with one or two tonal patterns.

The second part of the study examines the effects of f_0 and position on phonation. First, it was determined that f_0 influenced phonation (not position-in-utterance). Speakers produced words with each phonation in four prosodic positions: in isolation (high f_0), sentence initially (high f_0), sentence medially (mid f_0), and sentence finally (lower f_0). In medial position, a contrast between breathy, modal and creaky voice was found. In

sentence-final position, when tokens had lower f_0 , the contrast between modal, breathy and creaky voice was preserved, but with the modal and breathy words both breathier. In both isolation and sentence-initial position, when tokens had higher f_0 , the contrast between phonations was reduced to modal versus creaky. The effect of f_0 is so strong that the phonemic three-way contrast in phonations is not preserved when the f_0 is high.

1 Introduction

The Zapotec language family provides an interesting case for the study of contrastive phonation.¹ The various Zapotec languages have been described as having complex phonation contrasts that vary from language to language. Accounts of Zapotec languages describe more interesting phonations than are typically expected, such as the possibility of a four-way phonation contrast and the possibility of complex phonation contours within a single syllable. In addition, many accounts report phonations that are coupled with a wide array of tone. However, little is known about the phonetic properties of these phonation and tones. This study investigates the phonation and tones in Santa Ana del Valle Zapotec, experimentally, in the hopes that this approach will shed light on the complex phonation and tones of these languages.

Santa Ana del Valle Zapotec (hereafter SADVZ) is an OtoManguean language spoken in Santa Ana del Valle, Oaxaca, Mexico. The Ethnologue (Grimes, 1990) classifies SADVZ into the San Juan Guelavía Zapotec subgroup. (The Ethnologue classifies the numerous Zapotec languages into 58 different sub-groupings based on a variety of complex criteria.) The San Juan Guelavía subgroup contains the numerous and diverse languages spoken in the Valley of Oaxaca such as San Juan Guelavía (for which the subgroup is named), San Lucas Quiavini, Tlacolula, Jalieza, Mitla and Teotitlán del Valle Zapotec. 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 (Grimes, 1990).

¹ Abbreviations: perf = perfective, hab = habitual, irr = irrealis.

The Ethnologue considers San Juan Guelavía Zapotec to be the ‘lingua franca’ of the group, even though it is quite different from the other Zapotec languages spoken in the Valley. (The languages within the San Juan Guelavía subgroup have various levels of mutual intelligibility (59-100%).) For example, San Juan Guelavía Zapotec does not contain breathy vowels (Jones and Knudson, 1997) while many of the other languages in the subgroup do (e.g. San Lucas Quiavini Zapotec (Munro and Lopez et al., 1999), Mitla Zapotec (Stubblefield and Stubblefield, 1991).

Perhaps the best studied languages in this subgroup are San Juan Guelavía Zapotec (Jones and Knudson, 1997) and San Lucas Quiavini Zapotec (Munro and Lopez et al., 1999). While both of these languages have a practical orthography, this orthography is not used for SADVZ, which remains unwritten.

There has been no previous research on SADVZ, except for an unpublished wordlist produced by G. Aaron Broadwell (1991). The data and description here come from fieldwork conducted at UCLA by Argelia Andrade, Olivia Martínez, Pamela Munro and myself. The analysis presented here and any errors are my own.

1.1 Phonemic Inventory

SADVZ has 30 consonants and 6 vowels, which are presented in the sections below.

1.1.1 Consonants

The consonant inventory of SADVZ is shown in Table 1. The Zapotec languages have been described as having a fortis/lenis contrast for both the obstruents and sonorants, rather than a difference in voicing (Jaeger, 1983; Avelino, 2001). The fortis/lenis distinction is characterized by greater duration and increased energy for the

fortis consonants compared to the lenis ones. In the chart below, the symbol to the left of each pair is fortis and the right, lenis. (I have represented the fortis/lenis obstruents with the symbols for voiceless and voiced consonants, respectively, since this is their typical representation in Zapotec languages. The fortis/lenis sonorants are represented by their length contrast.) Consonants in parenthesis only appear in loan words, which are generally from Spanish.

Table 1: Consonants of SADVZ

	Labial	Dental	palatal-alv.	retroflex	palatal	velar	glottal
Stop	(p) b	t d				k g	ʔ
fricative	(f)	s z	ʃ ʒ	ʂ ʐ			
affricate			tʃ dʒ				
nasals	m : m	n : n				ŋ : ŋ	
laterals		l : l					
approximants					j	(w)	
Trill		r : r					
Tap		ɾ : ɾ					

I have analyzed the glottal stop as a phoneme, even though the distribution is somewhat defective since it only occurs after a vowel. Analyses of other Zapotec languages include the glottal stop as a vocalic feature and vowels that bear this feature are referred to as a “checked vowel” (Munro and Lopez et al., 1999; Jones and Knudson, 1977; Nellis and Hollenbach, 1980). The possibility of checked phonation will be considered in this study.

In addition, SADVZ stops are optionally released word finally and the lenis stops appear to be in free variation with fricatives. The bilabial stop [b] freely alternates with [β], the velar stop [g] alternates with [h], and the alveolar stop [d] alternates with [ð].

1.1.2 Vowels

SADVZ has six vowel qualities /a,e,i,o,u,ɨ/ which are presented in Table 2. The vowel /ɨ/ is rare.

Table 2: Vowels of SADVZ

	Front	Mid	Back
high	I	ɨ	u
mid	E		o
low			a

1.2 Phonation and Tone

No research has been done on the phonation and tones of SADVZ. It is expected that this language shares some of the features of phonation and tone found in other Zapotec languages. Section 2.1 will be an overview of phonation and tone in Zapotec languages.

1.3 Syntax

While a full description of the syntax of SADVZ is out of the scope of this project, it is necessary to mention some basic syntactic constructions that are relevant to this study.

1.3.1 Basic Word Order

SADVZ, like most Zapotec languages, is a VSO language. This can be seen in the following example where the verb ‘see’ [guṇa] is followed by the subject ‘Elena’ [elen] and object ‘cow’ [bag] :

1. gu- ṇa elen bag
perf -see Elena cow
“Elena saw a cow.”

1.3.2 Focus

SADVZ also allows SVO and OVS word order, in which case the preverbal material has a focused reading. Since Zapotec does not mark subject or object, the preverbal material is ambiguous. For example in 2, ‘Elena’ [elen] is focused and can be interpreted as either the subject or the object.

2. elen gu- n̄a bag
Elena perf-see cow
“ELENA saw a cow./The cow saw ELENA.”

1.3.3 Negation

In a negative, [teka] "no" is in sentence initial position, followed by the normal VSO word order.

3. teka gu - dau elen lim
No perf-eat Elena limes
“Elena didn’t eat limes.”

1.3.3.1 Negative Yes/No Questions

Negative yes/no questions are identical to negative statements and only distinguished by a change in intonation. Thus, the sentence [teka gudau elen lim] can mean “Elena didn’t eat limes” or “Didn’t Elena eat limes?” depending on the intonation.

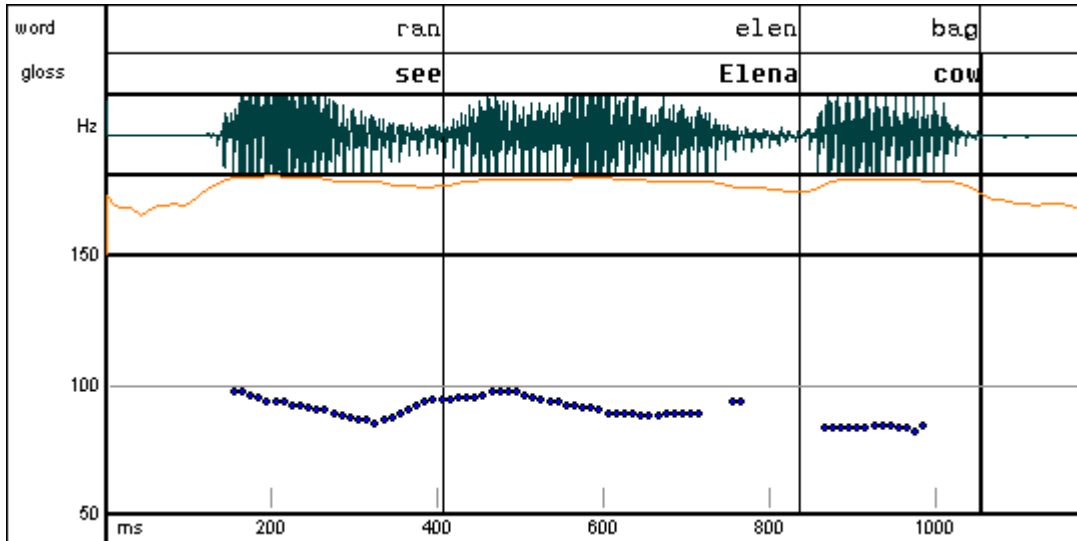
1.4 Some Basics of SADVZ Intonation

In this section, I will discuss the basic intonational contours associated with declaratives, focus, wh-questions and negative yes-no questions. Each sentence (except Figure 3) contains the target word ‘cow’ [bág], which has modal phonation and a high level tone. Each sentence was produced by the same speaker. The intonation of these sentence types is relevant for the experiment discussed in section 8.

1.4.1 Basic Declaratives

In basic declaratives, the end of the sentence has an overall low f0 and there is no sentence final fall of the f0, regardless of lexical f0. In Figure 1, the sentence ends with a low level f0.

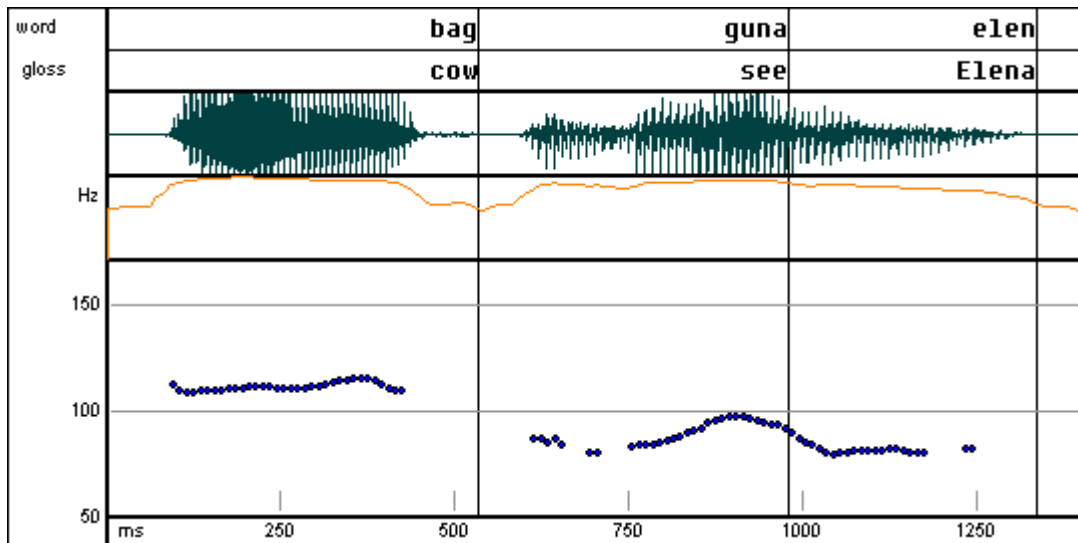
Figure 1: Pitchtrack of [ɾan elen bág] “Elena sees a cow.”



1.4.2 Focus

When a word is focused, it has an overall higher f0 and a greater duration than when it is uttered in a non-focus context. Compare the f0 of the word ‘cow’ [bág] in focused position in Figure 2, to sentence final position in Figure 1. The focused production has an overall higher, but still level, f0. There is no fall of the f0 on the focus item. The focus item is, by itself, a single intonational phrase.

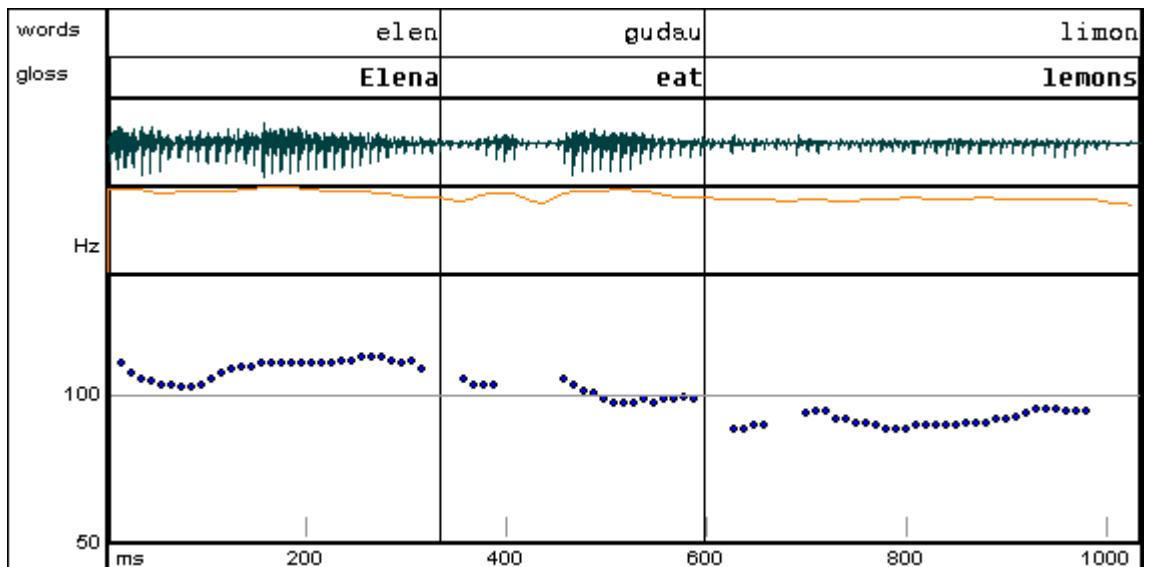
Figure 2: [bág guna elen] “Elena saw a COW./ A COW saw Elena.”



An additional example, with [elen] ‘Elena’ focused, is presented in Figure 3.

Again, there is no final fall of the f0 on the focused item.

Figure 3: [elén gudau limoni] “ELENA eats lemons.”

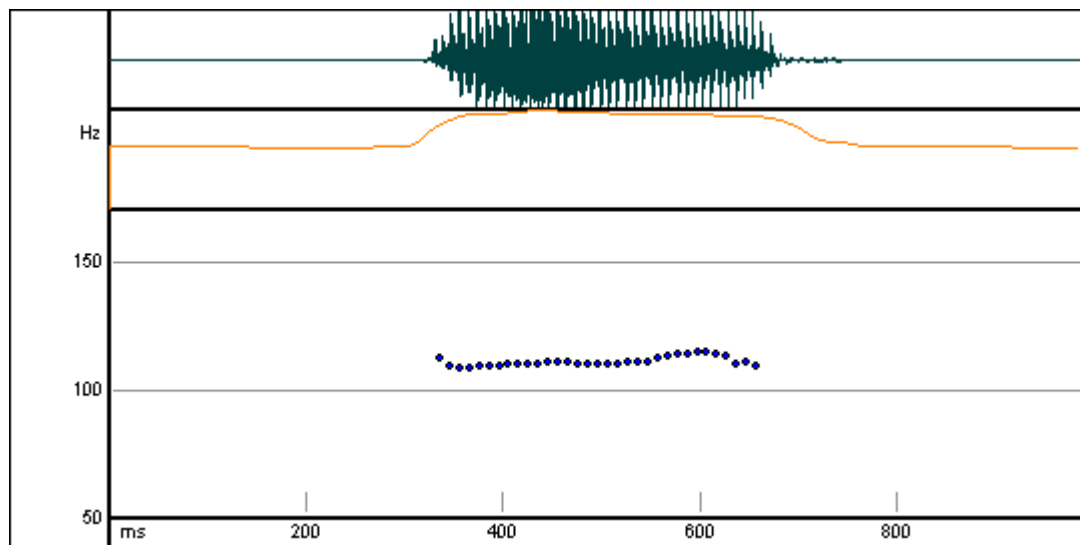


1.4.3 Isolation

Words uttered in isolation display the same f0 contour as words that are focused.

When a word is uttered in isolation, the f0 is still level, but overall higher than when produced in sentence final position in the declarative (Figure 1). (A direct comparison cannot be made between a word in isolation and a non-focused word in sentence initial position. Verbs are the only element that can be in non-focused sentence initial position. But, verbs are difficult to elicit in isolation.)

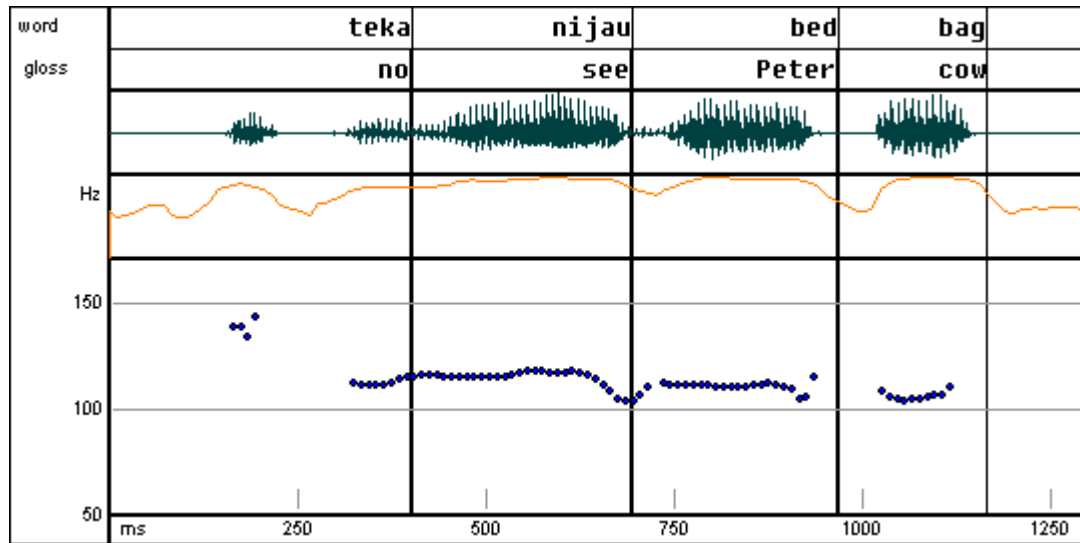
Figure 4: [bág] ‘cow’



1.4.4 Negative Yes/No Questions

Negative yes/no questions have a higher pitch range than declaratives. In Figure 5, ‘cow’ [bág] is higher (but still level) than when produced in sentence final position in a basic declarative (Figure 1).

Figure 5: [teka nijau bed bág] “Didn’t Peter see the cow?”



1.5 Structure of SADVZ Words

The vast majority of native Zapotec words are monosyllabic, while Spanish loan words are typically polysyllabic. Native polysyllabic words can be formed by the addition of morphemes; these often create changes in phonation. For example, the irrealis aspect can trigger a change from breathy to modal phonation. In the example below, the verb “to wash” [rig̃ibi] has a breathy vowel in the habitual form (4), but a modal verb in the irrealis² form (5):

4. ri-g̃ibi
hab-wash

5. k-ibi
irr.-wash

² The underlying form of the irrealis of ‘wash’ is /g- g̃ibi/. The /g-g/ sequence is realized phonetically as [k].

To avoid any changes in phonation, the majority of words examined in this study were monosyllabic. Section 6 of this study concentrates on polysyllabic words. These words were carefully chosen to avoid morphemes that would trigger changes in phonation.

1.6 Influence of Spanish

For over 500 years, Spanish words have been borrowed into SADVZ. The borrowing process is a complicated one, and will not be fully described here. However, I will mention some aspects that are relevant to this thesis.

When a word is borrowed from Spanish into SADVZ, several changes take place. The unstressed final vowel of the Spanish word is always dropped. In each of the examples in Table 3, the final vowel of the Spanish has been dropped in the SADVZ. For example, the Spanish [ˈlata] ‘(tin) can’ *lata* is [lát] in SADVZ.

In addition, any other unstressed vowel or syllable may be dropped (though the factors which determine whether an unstressed syllable will be deleted or will remain have yet to be determined). For example, [gaˈbino] ‘Gabino’ *Gabino*, is borrowed into SADVZ as [bín], losing both the final vowel [o] and the unstressed syllable [ga]. But, [baˈrato] ‘cheap’ *barato*, also a three syllable word, only loses the final vowel in the SADVZ [baɾăt].

Spanish vowels are always borrowed with modal phonation. Modal phonation, on both native and borrowed words, can have either a high or a high rising tone. Thus, borrowed words exhibit a tonal contrast not in the original language. Examples of borrowed words that display this tonal contrast are presented in Table 3. At this time, the

factor(s) that determine whether a loanword will be borrowed with a high or a high rising tone is not known. Cross-linguistically, a usual conditioning factor for a high versus rising tonal distinction is the voicing of the initial consonant; this does not seem to be the case for SADVZ. For example, [' falda] *falda* ‘skirt’ is borrowed with a high rising tone [fǎld] but [' fab] *faba* ‘brand of detergent (Fab)’ is borrowed with a high tone [fáb]. Nonetheless, since minimal pairs distinguishing the high and high rising tones have not been found, it is possible that tone is predictable in borrowed words. Further research is necessary to fully understand the borrowing process.

Table 3: Spanish Loan Words

Spanish	SADVZ
[' lata] <i>lata</i> ‘(tin) can’	[lát]
[' pato] <i>pato</i> ‘duck’	[bǎd]
[ba ' rato] <i>barato</i> ‘cheap’	[ba rǎt]
[' fab] <i>faba</i> ‘brand of detergent (Fab)’	[fáb]
[ga ' bino] <i>Gabino</i> ‘Gabino’	[bín]
[' falda] <i>falda</i> ‘skirt’	[fǎld]

In general, most of the words with modal phonation recorded for this study are loans from Spanish. In fact, it was difficult to find native words with modal phonation, though many neighboring languages do have native words with modal vowels and modal vowels are believed to have existed in Proto-Zapotec (Swadesh, 1947).

2 Phonation

The goal of this study is to understand phonation and tones, and the complex relation between the two, in a Zapotec language. But, before discussing the results

obtained for SADVZ, I will review the descriptions of phonation and tone in other Zapotec languages and then present the phonetic literature on phonation, in general.

2.1 Overview of Phonation and Tone in Zapotec Languages

There are numerous and diverse accounts of the phonation types and tones in Zapotec languages. Here, I briefly summarize some of the analyses of Zapotec phonation and tones.

Cajonos Zapotec (Nellis and Hollenbach, 1980) includes four vowels /i,o,e,a/ which can occur with one of three phonations: modal, checked, and laryngealized. Checked vowels, which only occur in open syllables, are articulated as a vowel followed by a glottal stop. Laryngealized vowels are defined as “a rapid sequence of two vowels with an intervening glottal stop” (Nellis and Hollenbach, 1980: 99) that is analyzed as a single-vowel segment. There is also a four-way tonal distinction between high, low and mid tones, and a “downglide from high to low tone” (Nellis and Hollenbach, 1980: 99). The tonal distribution is quite complex and determined by a number of factors such as the number of syllables in a word, if the vowel is followed by a lenis or fortis consonant, and the phonation of the vowel. For example, in “... monosyllabic roots with open syllables, simple vowels closed by either a fortis or lenis consonant, or checked vowels, there is a contrast only among the high...low...and glide” (Nellis and Hollenbach, 1980: 101). In monosyllabic roots “with laryngealized vowels, there is a four way contrast: both morae may have high tone, both low tone, the first high tone and the second mid, or the first high and the second low” (Nellis and Hollenbach, 1980: 101). In native disyllabic roots,

the tonal pattern is “high-high, low-high, and low-low” (Nellis and Hollenbach, 1980: 101).

Choapan Zapotec (Lyman and Lyman, 1977) includes six vowels /i,e,ɛ,u,o,a/ and three phonations: laryngealized, checked, modal (“unchecked”). A laryngealized vowel is a sequence of two vowels, which do not have to be of the same quality, with an intervening glottal stop; and a checked vowel is a vowel followed by a glottal stop. A typical Choapan syllable can have up to three phonation types over one vowel “segment”. The tones of Choapan Zapotec include: a high tone (which is slightly rising), a low tone (which is perceptually longer in duration than the other tones) and a mid tone. It is not clear if there is any relationship between tone and phonation in this language.

San Juan Guelavía (Jones and Knudsen, 1977) includes six vowels /i,e,ɨ, a, u, o/ each of which can occur with one of three phonations: modal, laryngealized or checked. Modal and checked vowels contrast phonemically. Laryngealized and checked vowels occur in complementary distribution; laryngealized vowels occur preceding lenis consonants or glides and checked vowels occur elsewhere. In addition, there are three level phonemic tones: low, mid, and high, which can occur on any of the three phonations.

San Lucas Quiaviní Zapotec (SLQZ) (Munro and Lopez et al., 1999) includes six vowels /a,e,i,o,u, ɨ/ which can occur with four phonations: modal, breathy, checked and creaky. A checked vowel is defined as a vowel followed by a glottal stop; creaky vowels co-occur in syllables with checked vowels. An SLQZ syllable can have up to three different phonations. “The fullest SLQZ syllable template is CCGVVVCG, where C represents a true consonant, G a glide and V is a vowel” (Munro and Lopez et al.,

1999:3). Most often, “only the final syllable of a non-compound uninflected native word ... is as elaborated as this template” (Munro and Lopez et al., 1999:3). Within the syllable the VVV sequence, referred to as the vowel complex, “may contain up to three individual vowels, each with its own phonation” (Munro and Lopez et al., 1999:3). SLQZ is also a tonal language, but the tones are dependent on the number and type of phonations found in a word. A characterization of the tonal system of SLQZ is as follows: “plain vowels (especially two plain vowels together) have high tone, creaky vowels have low tone ... and final phrasal breathy vowels have an extra-low tone. The tone of checked and other breathy vowels is derived from the vowel complex environment in which they occur” (Munro and Lopez et al., 1999:3).

In sum, the phonations of Zapotec languages are complex and vary across languages. Some studies describe either a four-way contrast between modal, creaky, breathy and glottalized (“checked”) vowels (e.g. San Lucas Quiaviní Zapotec) or a three-way contrast between modal, laryngealized or checked vowels (e.g. Cajonos Zapotec; Choapan Zapotec; San Juan Guelavía). For many analyses, it is not clear to what extent the checked vowel is thought to be a distinct phonation. It is possible that the checked vowel is a combination of a modal vowel followed by a glottal stop, but this is not explicit. On the same note, some analyses use the term laryngealized, but it is not clear if the term laryngealized is used to refer to creaky phonation or the presence of a glottal stop.

Tonal analyses also vary from language to language. Languages such as San Lucas Quiaviní Zapotec (Munro and Lopez et al., 1999) describe a strong correlation between

tone and phonation, while in languages such as San Juan Guelavía (Jones and Knudsen, 1977) there is no link between tone and phonation type.

In addition, some accounts describe the possibility of having up to three phonation types in one syllable (San Lucas Quiavini Zapotec (Munro and Lopez et al., 1999) and Choapan Zapotec (Lyman and Lyman, 1977)).

Based on the accounts of other Zapotec languages, it is possible that SADVZ could have up to four phonations, which may or may not be correlated with tone, and could have changes in phonation within a single syllable.

2.2 Overview of Phonation across Languages

In this section, I will present an overview of phonation across languages, to see how Zapotec phonation fits in cross-linguistically. I will begin by discussing how the different phonations are characterized in the phonetic literature, limiting the discussion to languages that use phonation types contrastively, and then discuss properties commonly associated with non-modal phonation, such as an increase duration and localization of non-modal phonation.

There are three main types of phonations that occur cross-linguistically: modal, breathy, and creaky voice. Modal voice is the neutral phonation to which other phonations are compared. Modal phonation is characterized by vocal folds with normal adductive and longitudinal tension. During modal phonation, the vocal folds are open during approximately half of the glottal cycle and closed for the other half, which results in regularly spaced glottal pulses.

Breathy phonation is characterized by vocal folds that vibrate, but without much contact. For breathy voice, there is minimal adductive tension and little longitudinal tension in the vocal folds. Breathy voice is contrasted with modal voice on vowels in languages such as Gujarati and Tamang (as reviewed by Gordon and Ladefoged, 2002).

Creaky voice is distinguished from breathy and modal voice by little longitudinal tension and high adductive tension in the vocal folds. During creaky voice, the vocal folds are opened just enough to allow for voicing, which often results in irregularly spaced glottal pulses. Creaky voice is contrasted with modal voice on vowels in Sedang and Southern Nambiquara. In addition, certain languages use a voice quality that is similar to creaky voice, but does not fall precisely under the category of creaky.

Examples of this are Bruu, which has a contrast involving stiff and slack vocal folds and !Xóõ, which uses a strident voice quality to distinguish vowels (as reviewed by Gordon and Ladefoged, 2002).

Some languages, such as Jalapa Mazatec and Chong, have a three way contrast in phonation between modal, breathy and creaky vowels (Blankenship, 1997). In the phonetic literature, no languages have been described as contrasting four phonation types or as having a sequence of contrastive non-modal phonations within a syllable. (Laver (1980) presents a more elaborate set of phonetic phonations, but only breathy, modal and creaky phonations are known to contrast.)

2.2.1 Localization of Non-Modal Phonation

Non-modal phonation on vowels is often confined to a portion of the vowel, especially in languages with contrastive phonation and contrastive tone. This is the case

for Jalapa Mazatec, a language with both contrastive tone and phonation. In this language, creaky and breathy voice are localized to the first portion of the vowel (Silverman, 1997; Blankenship, 1997). “The second portion of the vowel usually possesses severely weakened breathiness or creakiness, verging on modal phonation” (Silverman, 1997: 238). Silverman (1997) suggests that the localization of non-modal phonation is linked to the use of tone. Since non-modal phonation influences the saliency of the fundamental frequency, it would be difficult for creaky or breathy vowels to carry a full range of tonal contrasts. But, by limiting creaky or breathy voice to the beginning of the vowel, a portion of the vowel remains modal, and is thus able to support a full range of tonal contrasts (Silverman, 1997).

2.2.2 Duration of Vowels with Non-Modal Phonation

In some languages, non-modal phonation is associated with an increase in duration. This is true of Jalapa Mazatec, where creaky and breathy vowels are longer than modal vowels (Silverman et al. 1995). However, this is not true of all languages that have contrastive phonation. In San Lucas Quiavini Zapotec, a language that also contrasts creaky and breathy vowels, there is not a significant duration difference between modal and non-modal phonation (Gordon and Ladefoged, 2002).

In summary, the Zapotec languages presented in section 2.1 are typologically unusual when compared cross-linguistically. For example, the four-way contrast in phonation (e.g. San Lucas Quiavini Zapotec (Munro and Lopez et al., 1999) is typologically rare (since most languages have a maximum of three phonations). Another unusual feature is the sequence of phonations within a syllable that is described for some

Zapotec languages (e.g. Cajonos Zapotec (Nellis and Hollenbach, 1980) and San Lucas Quiavini Zapotec (Munro and Lopez et al., 1999)). In the next section, I will review some of the common measurements of non-nodal phonation that can assist with identifying and defining the phonations of SADVZ.

3 Measuring Non-Modal Phonation

There are numerous acoustic and auditory properties that can be useful measures of non-modal phonation. In this section, I will briefly discuss some of the major methods for measuring non-modal phonation from an audio signal. (Measures based on the glottal source will not be discussed here.)

3.1 Periodicity

Creaky phonation is generally characterized by aperiodic glottal pulses. The aperiodicity of creaky vowels can be quantified by jitter, “the variation in duration of successive fundamental frequency cycles” (Gordon and Ladefoged, 2002:15). Jitter values are greater for creaky vowels than for modal or breathy vowels, where the glottal pulses are more regularly spaced.

Another measure of aperiodicity is cepstral peak prominence. A cepstrum is an “inverse spectrum generated by taking the FFT ... of the log magnitude values of a power spectrum” (Blankenship, 1997:8). For modal phonation, which is a periodic signal, the spectrum has well-defined harmonics and therefore a higher (in amplitude) cepstral peak. Non-modal phonation has less distinct harmonics and therefore a lower cepstral peak. (Blankenship, 1997).

3.2 Acoustic Intensity

In some languages, breathy and creaky voice are associated with a decrease in acoustic intensity compared to modal voice. This is true of breathy vowels in Gujarati and Chong, and of creaky vowels in Chong and Hupa (as reviewed by Gordon and Ladefoged, 2002).

3.3 Spectral Noise

Breathy voice can be characterized by an increase in spectral noise, particularly at higher frequencies. This noise is due to the continuous leakage of air through the glottis that occurs during breathy phonation. Languages that express breathiness through noise include Newar, Jalapa Mazatec and San Lucas Quiaviní Zapotec (as reviewed by Gordon and Ladefoged, 2002).

3.4 Spectral Tilt

One of the major ways to measure phonation is spectral tilt. Spectral tilt “is the degree to which intensity drops off as frequency increases” (Gordon and Ladefoged, 2002:15). Subtracting the amplitude of a higher frequency harmonic from the amplitude of the fundamental frequency (also called the first harmonic) yield a largely positive value for breathy vowels, a smaller positive value for modal vowels, and a negative value for creaky vowels. Spectral tilt has been a reliable measure of phonation in numerous languages such as Jalapa Mazatec, Gujarati, Kedang and Hmong (as reviewed by Gordon and Ladefoged, 2002).

There are different ways to characterize spectral tilt. Primarily, the difference between the amplitudes of the first and second harmonics, which correlates with the portion of the glottal cycles in which the glottis is open (the open quotient), has been used

to distinguish between modal and breathy phonation. However, other studies have made use of the relationship between H1 and harmonics exciting higher formants, which correlates with the abruptness of the closure of the vocal folds. These measurements include: H1-F3 (Stevens and Hanson, 1995; Blankenship, 1997), H1-F1 or H1-F2 (Ladefoged, 1983) and the average of H1-H2 compared to F1 (Stevens, 1988). Other studies have used the relationship of higher formants to lower ones such as F2-F3 (Klatt and Klatt, 1990).

Because they reflect different aspects of phonation, the different measures of spectral tilt do not always distinguish phonation types, even within a single language. For example, Blankenship (1997) found that in Mpi the measurement of H1 – H2 was a more reliable measurement of phonation on vowels with high tone than with mid or low tone. This suggests that in Mpi, the phonations of high tone vowels differ in their open quotient, while the phonations of low tone vowels do not.

4 Experimental Comparison of Phonation in SADVZ

Some of the Zapotec languages discussed in section 2.1 have a three-way contrast in phonation between breathy, creaky and modal vowels. However, other accounts propose a possible fourth phonation, a “checked vowel.” Analyses also differ in the number and types of phonation that can be produced on a single syllable and the possible tonal contrast on different phonations. Thus, the numerous and diverse accounts of phonation in Zapotec languages cast doubt on a simple three way contrast as an accurate description for the phonation of SADVZ.

As a starting point, however, I propose a three-way contrast in phonation for SADVZ, because a near minimal triplet (distinguished by the phonation of the vowel) can be found. In the next section, I will determine the most appropriate measure of phonation to distinguish between the modal, breathy and creaky vowels of the near minimal triplet. Once a measure of phonation has been selected to distinguish the minimal triplet, it will be possible to see if the proposed three-way contrast in phonation should be expanded to include a possible fourth “checked” phonation or any other additional phonation types. In addition, once a measure of phonation has been selected, it will be possible to determine if phonation ever changes over the course of a syllable.

4.1 Selecting a Measurement of Phonation

Spectral measurements were used to measure phonation in SADVZ. Spectral noise, periodicity, and acoustic intensity were not measured.

4.1.1 Spectral Measurements

Different spectral measures yield different results depending on the nature of the phonation. In this section, I will concentrate on selecting the most appropriate measure of phonation in SADVZ. To determine which spectral measurement would best distinguish phonations for this study, 6 measures ($H1-H2$, $H1+H2/2-F1$, $H1-F1$, $H1-F2$, $H1-F3$, and $F2-F3$) were tested on the near minimal triplet consisting of the three words shown in Table 4. The near minimal triplet was selected based on cognates in a related language, San Lucas Quiaviní Zapotec. In Table 4, the San Lucas Quiaviní data is presented in the orthography for this language; I have written the phonation that

corresponds to the San Lucas Quiaviní orthographic conventions in parenthesis next to the word. Perception confirmed that these words were cognates in SADVZ.

Each word was uttered in sentence medial position, and repeated ten times by three native speakers ranging from 40-50 years of age (two male, Speaker 1 and Speaker 2, and one female, Speaker 3). Tokens were digitized and analyzed in PCQuirer at a sampling rate of 22050 Hz. Figures 6-8 and Figure 9-11 are spectrograms, waveforms and FFTs of the sample data as pronounced by Speaker 1 and Speaker 3, respectively. Each measurement was made over a 30 ms window, 50 ms before the end of the vowel. Spectrograms were used to position the 30 ms window. In the case of creaky token, this window would naturally include the silent intervals between pitch pulses. In addition, for the female speaker's breathy tokens, the window is partially during the noise interval. Measurements were taken manually from the FFT. H1 is equal to the height of f0 in the FFT; H2 equals the height of the second harmonic; F1, the amplitude of the highest harmonic near the first formant; F2, the amplitude of the highest harmonic near the second formant; F3, the amplitude of the highest harmonic near the third formant. The measure of phonation is the difference between the two measures ($H1-H2$, $H1+H2/2-F1$, $H1-F1$, $H1-F2$, $H1-F3$, and $F2-F3$) in dB.

Table 4: Words for the measurement trials and their cognates

SADVZ	Cognate in San Lucas Quiavini Zapotec
lat '(tin) can'	la't '(tin) can' (modal)
l̥at 'place'	laht 'place' (breathy)
l̥ats 'field'	l̥aa'ts 'flat area' (creaky)

Figure 6: Waveform, spectrogram and FFT of 'can' [lat] produced by Speaker 1. Pitch track fails at end.

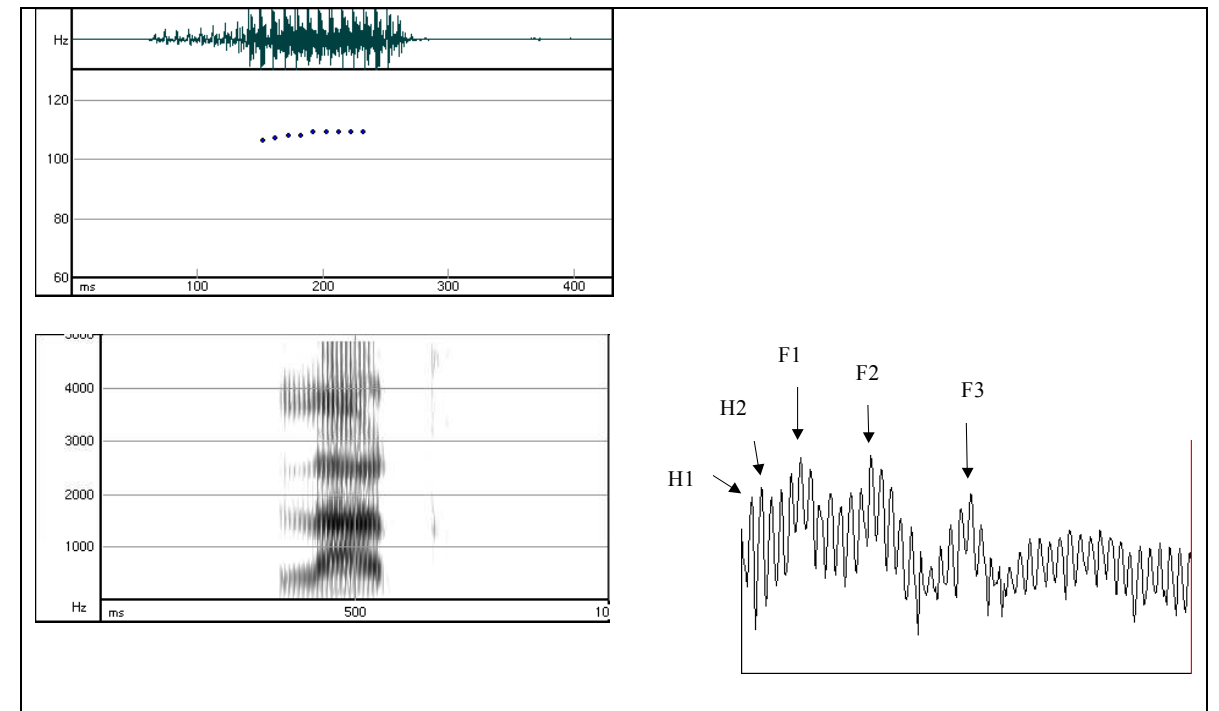


Figure 7: Waveform, spectrogram and FFT of ‘place’ [læt] produced by Speaker 1.

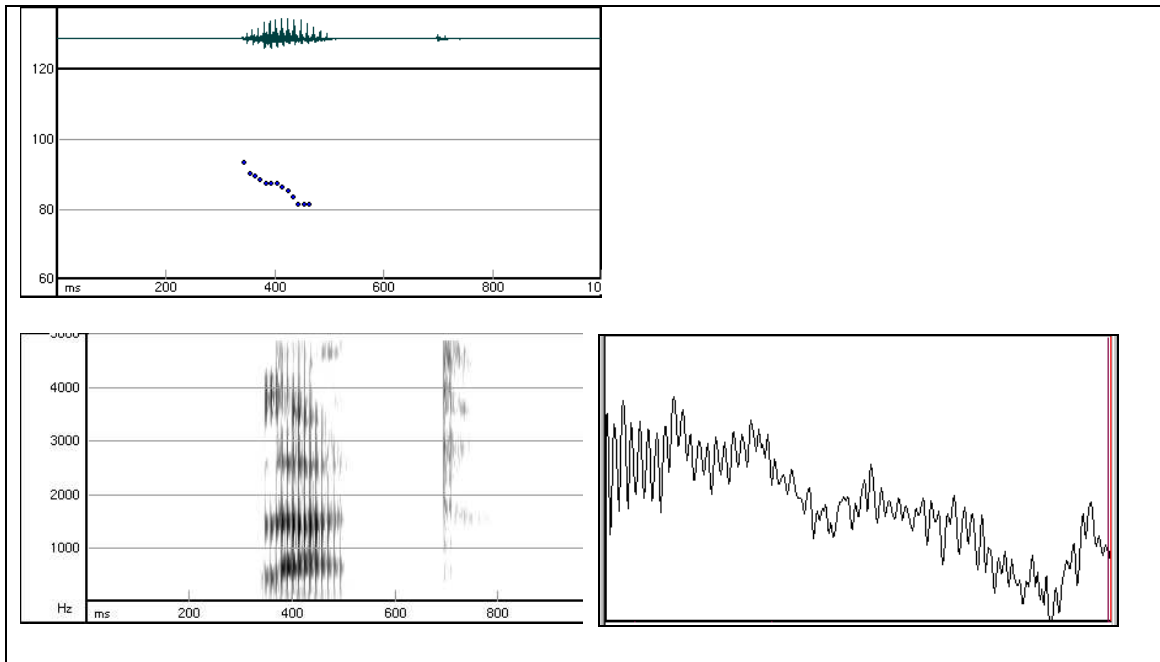


Figure 8: Waveform, spectrogram and FFT of ‘field’ [fild] produced by Speaker 1. Pitch track fails at end.

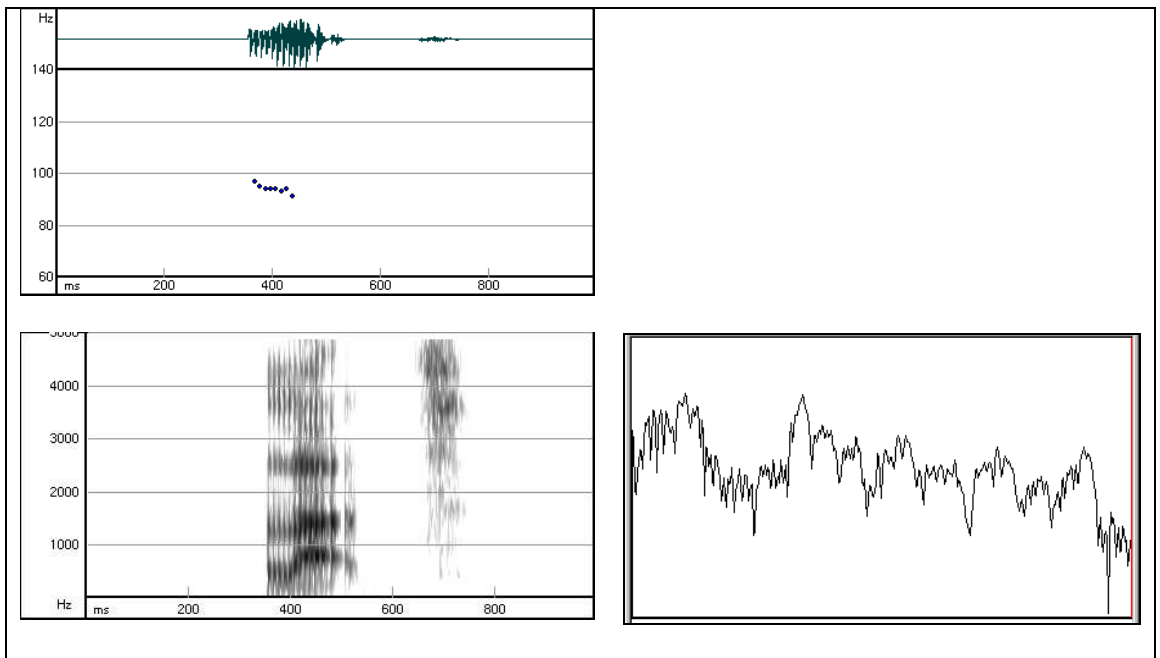


Figure 9: Waveform, spectrogram and FFT of 'can' [lat] produced by Speaker 3.

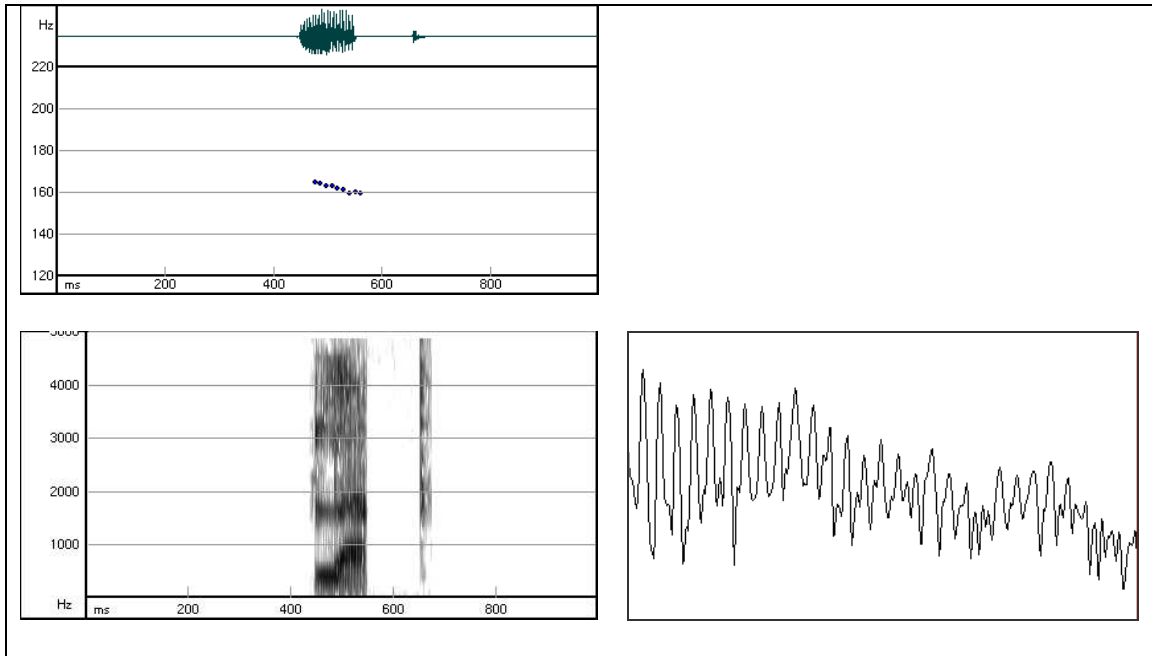


Figure 10: Waveform, spectrogram and FFT of 'place' [lat] produced by Speaker 3. Pitch track fails at end.

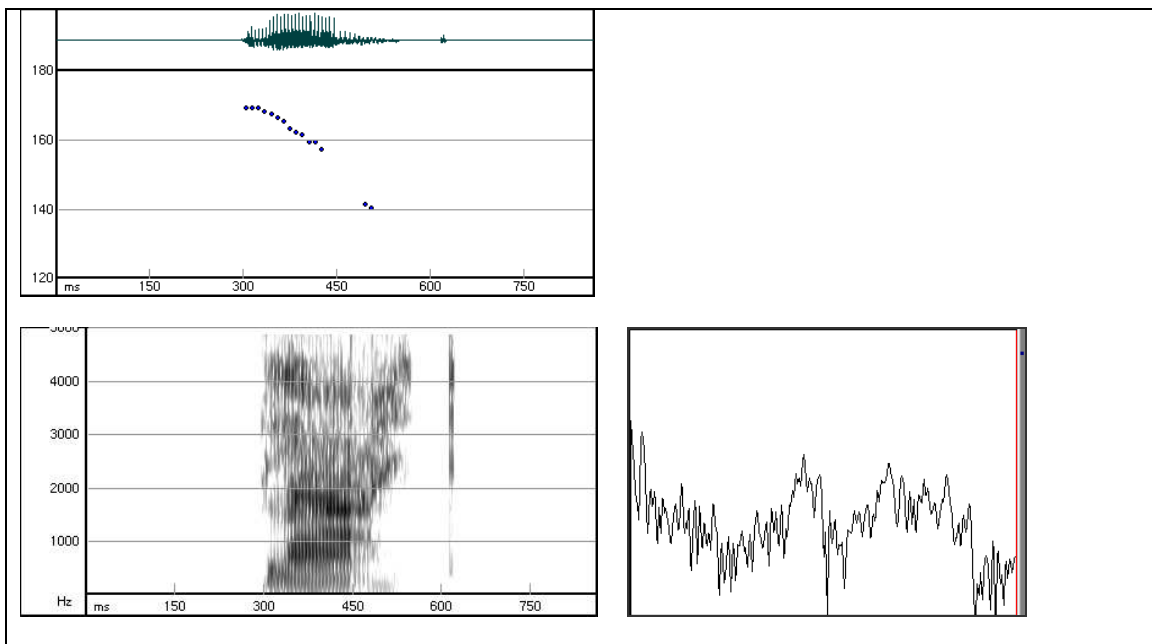


Figure 11: Waveform, spectrogram and FFT of 'field' [lɑt] produced by Speaker 3. Pitch track fails at end.

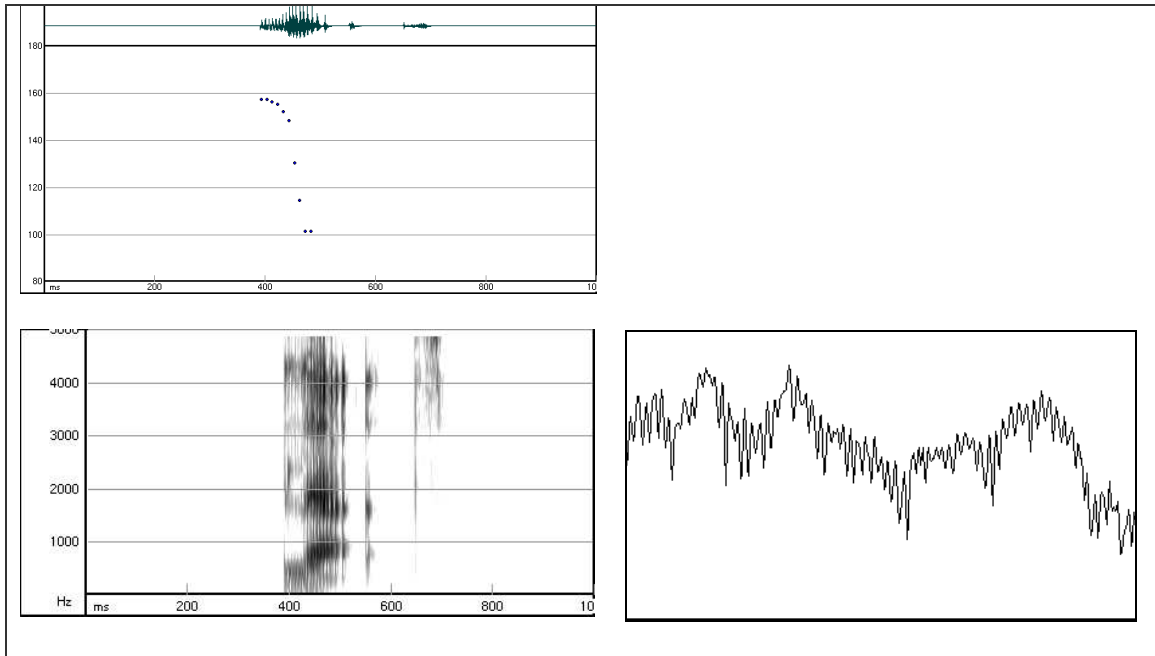


Table 5 shows the number of times the measures of phonation were successful on the near minimal triplet (the maximum is ten). The expected direction of difference is given at the top of the chart. The shaded boxes represent the most successful measures for a given speaker. A measurement was considered successful if it had any amount of difference in the expected direction. A statistical comparison of the measurements was not performed.

Table 5: Number of times the measurement succeeded

Measurement	Creaky (expected: creaky < modal)			Breathy (expected: breathy > modal)		
	Speaker 1	Speaker 2	Speaker 3	Speaker 1	Speaker 2	Speaker 3
H1-H2	4	4	10	0	0	10
H1+H2/2-F1	3	2	4	0	0	4
H1-F1	5	3	0	4	5	2
H1-F2	4	5	0	2	5	1
H1-F3	10	10	0	10	10	0
F2-F3	2	0	2	4	4	2

H1-F3 was the most successful measure of both breathy and creaky phonations for the male speakers; H1-H2, for the female speaker. The different measurements indicate a difference in the production of phonation by the speakers. For the female speaker, H1-H2 was the most successful measure of phonation. The measure of H1-H2 correlates with the open quotient, the portion of time the vocal folds are open during each glottal cycle. When the vibration of the vocal folds has a large open quotient, the fundamental frequency dominates the spectrum. This is the case for breathy phonation, which has a large open quotient and a spectrum dominated by H1. Thus, the measurement of H1-H2 is largely positive for breathy phonation.

The most successful measure of phonation for the male speakers was H1-F3. H1-F3 reflects the manner and speed at which the vocal folds close. When the vocal folds come together abruptly, it excites the higher frequencies of the vowel. For this reason, creaky phonation, which is characterized by vocal folds that close abruptly, has a higher value for F3 than H1. Thus, the phonation differences of the male speakers are apparently produced by the speed of closure of the vocal folds (which correlates with H1-F3), and not the proportion of time the vocal folds are open or closed. The opposite is true of the

female speaker. Her phonations are apparently produced by changes in the portion of time the vocal folds are open or closed (H1-H2), and not the rate of closure.

4.2 Phonation in SADVZ

Now that a measure of phonation has been determined for each speaker, it is possible to apply this measure to a wider range of words in SADVZ to determine whether a simple three-way phonation contrast, with no contours within a syllable, suffices to describe its phonations.

4.2.1 Methods

4.2.1.1 Speakers

The same three speakers from section 4 of this study were used here.

4.2.1.2 Speech Materials

A wordlist was selected based on cognates in related languages. The wordlist included tokens that could potentially exhibit a more complex phonation contrast than simply breathy, modal and creaky, and could potentially exhibit a change in phonation throughout a syllable. The wordlist was composed of monosyllabic words all with the target vowel [a] (since the first formant of low vowels does not influence the amplitude of the first or second harmonics as much as in higher vowels) and a range of consonants. (See Table 8 in appendix for wordlist.)

4.2.1.3 Procedure

The three speakers were recorded over multiple sessions in a sound booth. All words were uttered in the frame sentence [*guni?* ___ *primer*] “Say ___ first.” The sentence was selected since it was determined (through H1-F3 and H1-H2) that these two words only include modal phonation. The sentence was produced as a single intonational phrase. Each token was recorded 5 times.

As before, tokens were digitized and analyzed in PCQuirer at a sampling rate of 22050 Hz. A Fast Fourier Transform (FFT) was calculated over a 30 ms window starting 50 ms before the end of the vowel. To calculate the phonation for the male speakers, the amplitude of H1 and the amplitude of the highest harmonic near F3 was taken from the FFT; for the female speaker, the amplitude of H1 and H2 was taken. The measure of phonation is the difference between these two measurements in dB. Thus, a value of zero means that the two measured harmonics had the same value.

4.2.2 Changes in Phonation within a Syllable (Contour Phonations)

H1-F3 (for the males) and H1-H2 (for the females) was measured at three timepoints within a vowel: initial, medial and final. Figures 12-14 show the frequency of occurrence of the values of H1-F3 (in dB) for Speaker 1. (Since all three speakers displayed the same results, I will only present the histogram for Speaker 1.) The x-axis represents the measure of phonation (H1-F3) in db. The y-axis represents the number of tokens with each value.

Figure 12 represent H1-F3 (in db) for the initial portion of the vowel. Here, the data fall into two distinct groups, corresponding to two phonations. Thus, in the initial portion of the vowel, there is a distinction between modal (2-6 dB) and breathy phonation (14-20 dB).

Figure 12: Frequency of H1-F3 values for Speaker 1 (initial portion of vowel).

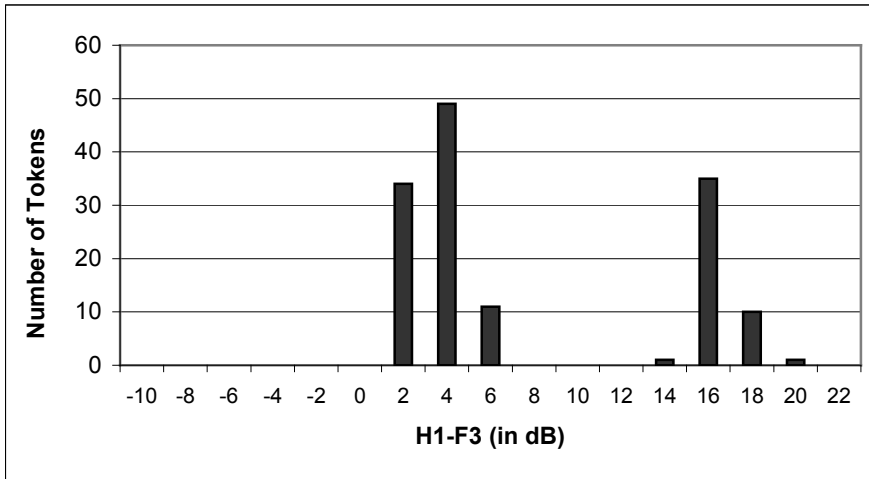


Figure 13 represents the measure H1-F3 in the middle of the vowel. Again, there are two distinct groups, representing two phonation: modal (0-6 dB) and breathy (10-16 dB).

Figure 13: Frequency of H1-F3 values for Speaker 1 (medial portion of vowel).

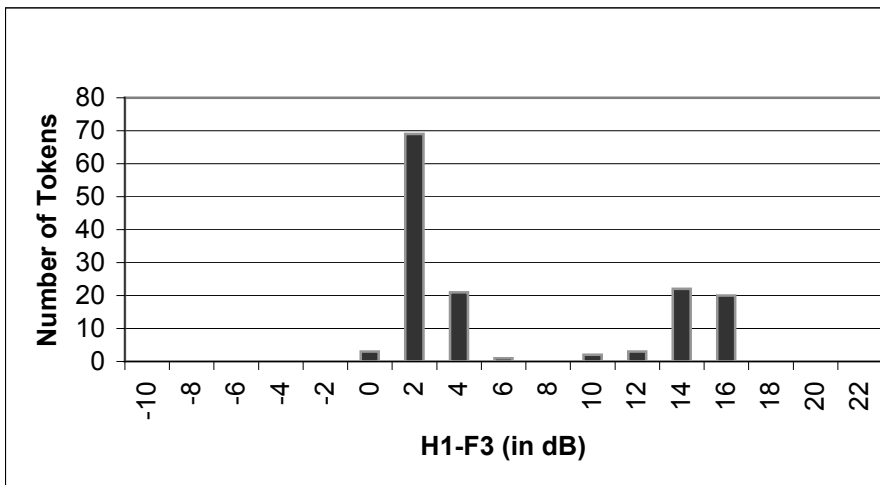
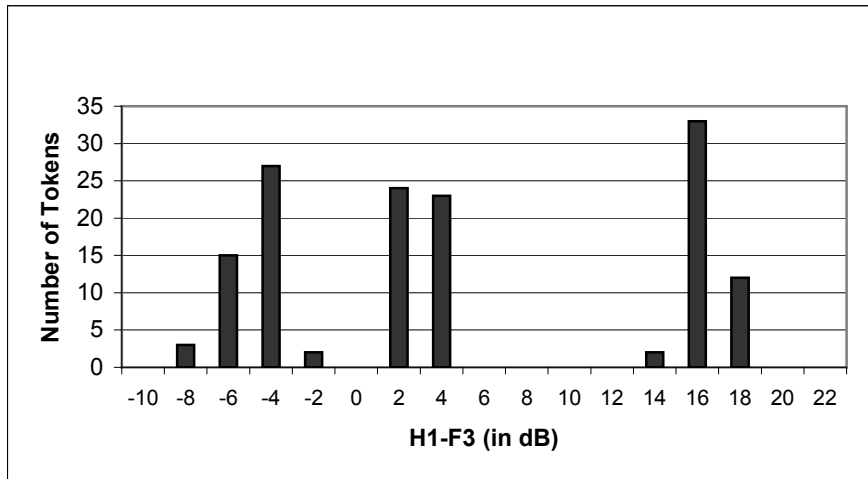


Figure 14 represents H1-F3 (in dB) for the final portion of the vowel. In this graph, three distinct groups emerge. Thus, in the last portion of the vowel, it is possible to have creaky (-8 - -2 dB), breathy (14-18 dB) or modal (2-4 dB) phonation.

Figure 14: Frequency of H1-F3 values for Speaker 1 (final portion of vowel).



Figures 12-14 show that there is not a change in phonation over the course of the vowel, except in the case of creaky phonation. (The breathiness contour in Figure 10 is not typical.) Creaky vowels begin with modal phonation and have the creaky phonation localized to the end of the vowel. (In addition, right before the creaky voicing begins, there is a brief portion of the vowel that looks breathy, though the spectral measure over the 30 ms window was creaky. For the female speaker, breathy vowels have a voiceless segment at the end, but this does not show up in the measurements, since it is not voiced.)

4.3 Three phonations (Modal, Breathy, Creaky)

In this section, detailed results and descriptions for the three phonations that have been found are presented. The results of the measures of phonation for all three Speakers can be found in Figures 15-17. For creaky vowels, the measurement is shown at both the beginning and the last 50 ms of the vowel. For modal and breathy tokens, only one point

is shown, at the last 50 ms of the vowel. The difference between the amplitudes of H1 and F3/H2 is given on the y-axis. Results for the three speakers are similar.

Figure 15: H1-F3 for Speaker 1 (sentence-medial position)

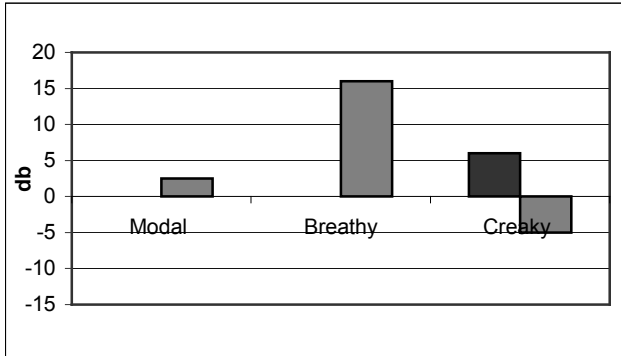


Figure 16: H1-F3 for Speaker 2 (sentence-medial position)

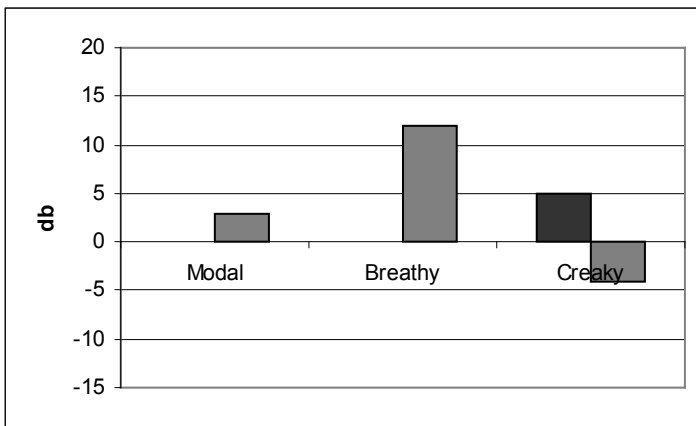
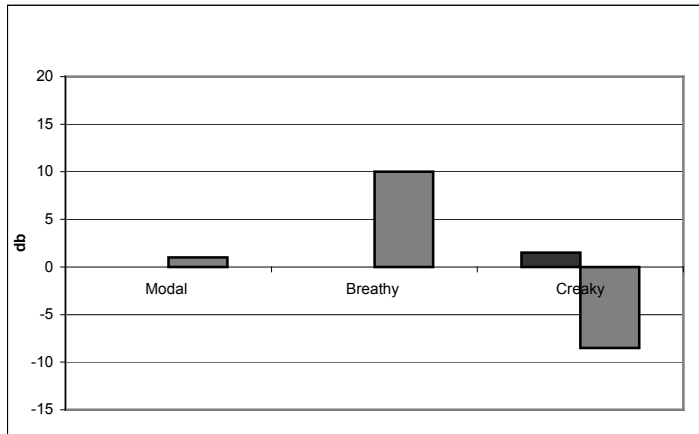


Figure 17: H1-H2 for Speaker 3 (sentence-medial position)



Thus, the measures of phonation (H1-H2 and H1-F3) distinguished three phonations in SADVZ: modal, breathy and creaky. There was no evidence for a fourth “checked” phonation. Glottal stops appeared after creaky and modal vowels, in which the phonation of the vowel was not changed. Thus, vowels that were followed by a glottal stop had the same phonation as modal or creaky vowels that were not followed by glottal stops. One possible exception to this is some creaky tokens that are a sequence of a creaky phonation with an intervening glottal stop. (These will be discussed in more detail in section 4.3.1.1.1.)

Before presenting more details on these phonations, I will first discuss other difference among the phonations such as f_0 and duration. The phonation types will be further illustrated in 4.4.

4.3.1 Other Differences Among and Within Phonations

4.3.1.1 F0

From the FFT, f_0 was measured for each token at three time points (the beginning, middle and end) of the vowel. F_0 corresponded to the frequency of H1. The results of the measurements of f_0 are presented in Figures 18-20. Timepoints are represented on the x-axis and f_0 is represented in Hz on the y-axis.

For all three speakers, a strong relation between f_0 and phonation was found, which will be described here as tone. Creaky vowels always have a large falling tone, which will be represented by a falling accent. Breathy vowels have a smaller falling tone, though this is less apparent for Speaker 1, which will be represented by a grave accent. Modal vowels have either a high level, or a high rising tone, which will be represented by an acute and rising accent, respectively.

While the relationship between high and high rising is not consistent across speakers, for all three speakers there are two distinct modal tonal categories – high level and high rising.

Figure 18: F0 for Speaker 1 (sentence-medial position)

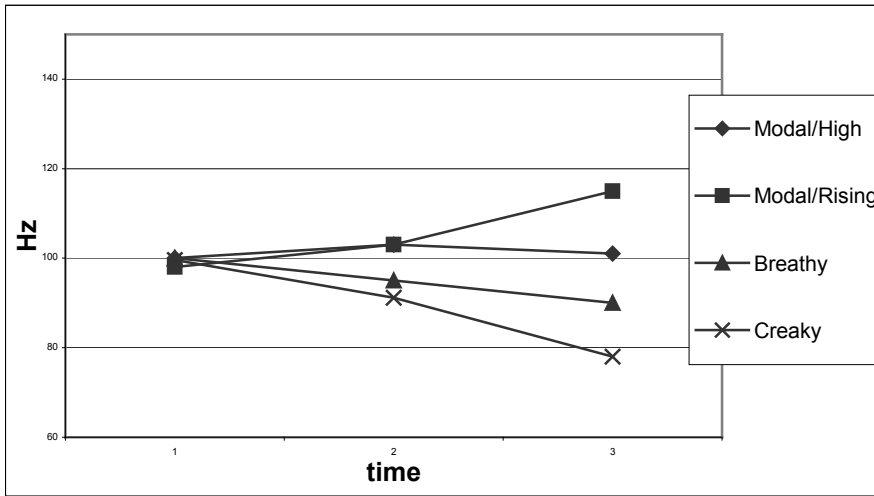


Figure 19: F0 for Speaker 2 (sentence-medial position)

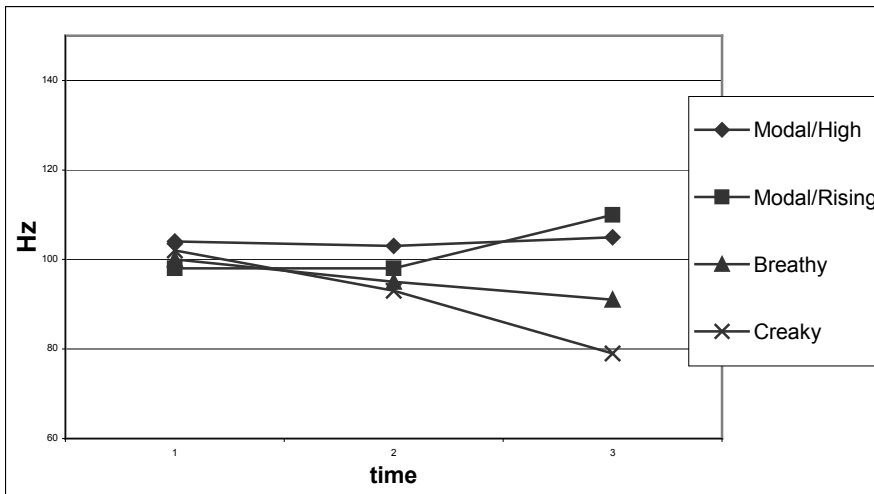
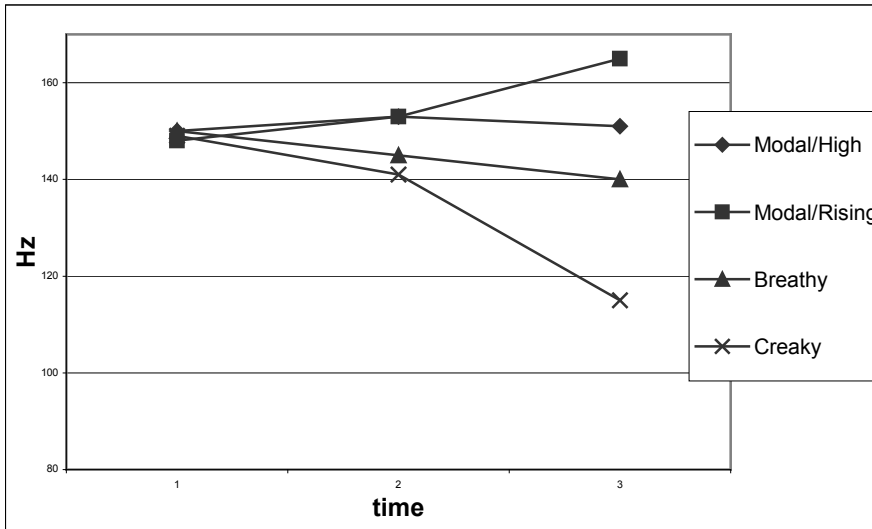


Figure 20: F0 for Speaker 3 (sentence-medial position)



Given the two tonal categories of high and high rising for modal phonation, the question arises as to whether there is a difference in phonation too. The phonation of modal vowels with high versus high rising tone is presented in Figures 21-23. Graphs represent the average of the measurements for 10 tokens per phonation (see Table 8 in the appendix for wordlist). The difference between the amplitudes of H1 and F3/H2 (in dB) is given on the y-axis.

Figure 21: H1-F3 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 1)

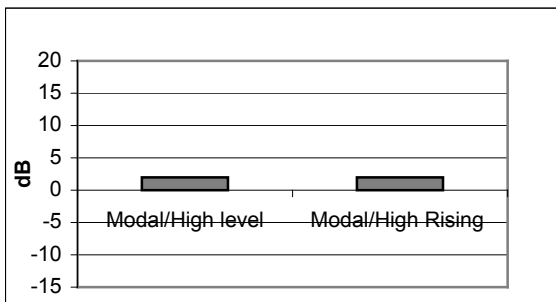


Figure 22: H1-F3 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 2)

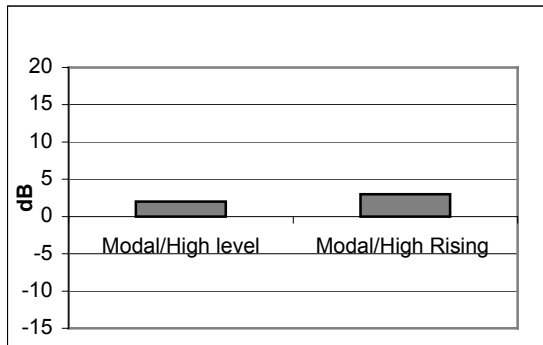
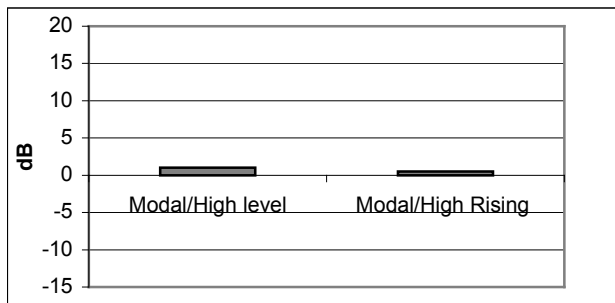


Figure 23: H1-H2 of Modal Phonation with High Level and High Rising Tone (medial position, Speaker 3)



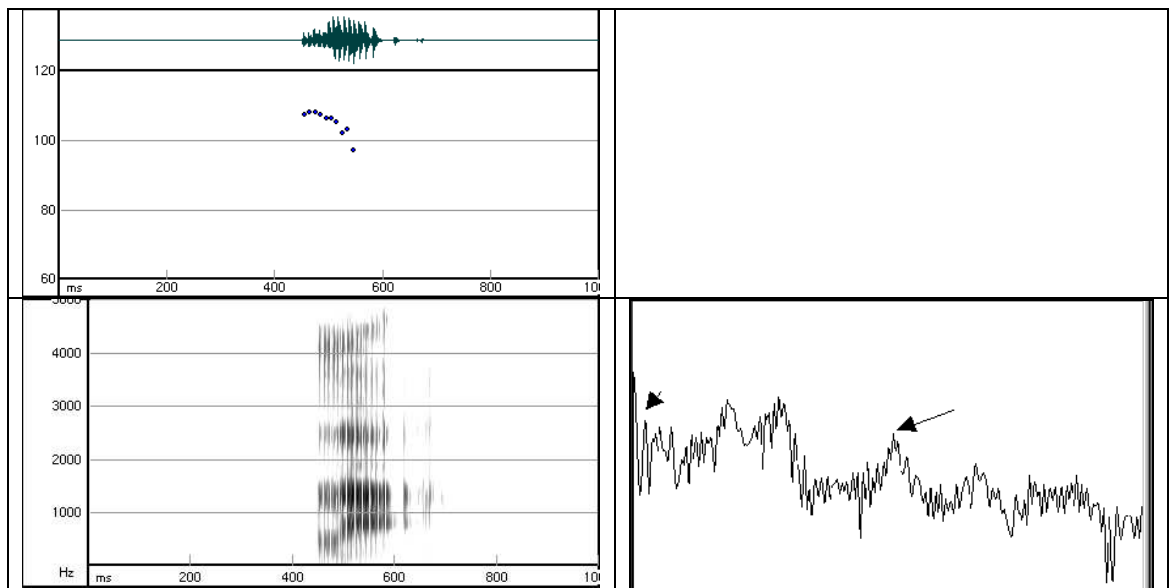
The graphs in Figures 21-23 indicate that there is not a difference in phonation between the high and high rising tone vowels.

4.3.1.1.1 Creaky phonation as a Result of Low F0

Some creaky tokens are a sequence of a creaky phonation with an intervening glottal stop. (These tokens were not included in previous figures since they have a different syllable structure, [VʔV], which is considered to be one syllable.) The portion of the vowel after the glottal stop often has weak voicing. The last portion of the vowel sequence has what sounds like creaky phonation, though often the expected spectral measurement was not obtained for the male speakers. Often, vowels that were perceptually creaky and showed visible signs of creak on the spectrogram measured as if

they were breathy or modal. In the example in Figure 24, the creaky phonation is visible at the end of the vowel, in the form of a very low f_0 . But, the measurement of H1-F3 is modal. (The other measures presented in section 4.1.1 also yielded a modal result.) From the FFT, you can see that the amplitude of H1 (the first arrow) is roughly the same as the amplitude of F3 (second arrow), yielding a more modal value for this phonation than would be expected. These are cases where the phonation has been compromised, but the f_0 is preserved. Here, the phonation (the mode of vibration of the vocal folds) is modal, but the f_0 (the rate of vibration) is as it would be expected for the creaky vowels. Thus, it is f_0 alone that is making the token sound creaky.

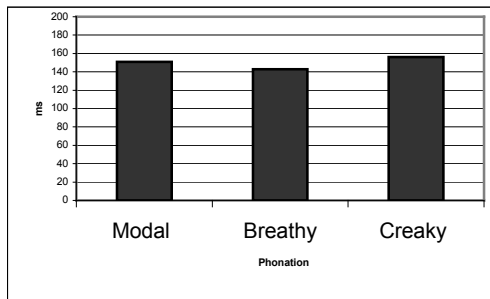
Figure 24: Spectrogram, waveform, pitchtrack and FFT of ‘mat’ [d̥ʌʔə] Pitchtrack fails at end.



4.3.1.2 Duration

Duration (in ms) was measured for the three phonations. No difference was found in the duration of modal versus non-modal phonation. I have presented the results for Speaker 1 below.

Figure 25: Duration of Modal and Non-Modal Phonation (Speaker 1)



4.4 Summary of Results

The study presented in this section has compared [a] in monosyllabic words in sentence-medial position. Results for all three speakers showed that SADVZ's three-way contrast in modal, breathy and creaky vowels is distinguished acoustically in the directions expected. No difference between checked and modal phonation was found. In addition, no change in phonation within a syllable was found, except for creaky phonation. Creaky vowels are characterized by creakiness that is limited to the end of the vowel; creaky vowels are not distinct from modal vowels until the last portion of the vowel. Sometimes, creaky vowels have no special phonation, just the characteristic falling tone (Figure 24). Breathy vowels are associated with a decrease in the higher frequency energies or noise at the end of the vowel. In addition, each of the phonations is associated with one or two tonal patterns: modal vowels can have a high or high rising tone, and breathy and creaky vowels have a falling tone. Given these results, I will now transcribe words systematically with phonation and tone diacritics.

Each of these phonation and tonal patterns is summarized in more detail in the following sub-sections. Since the measurements of phonations were similar for all three

speakers, in each section I will only provide examples from Speaker 1, unless there is a difference among speakers.

4.4.1 Modal

Within modal phonation there is a contrast between high and high rising tone.

Figures 26 and 27 show a waveform, spectrogram and pitch track of a modal vowel with a level high tone ‘lid’ [táp] and a high rising tone ‘father’ [dǎd], respectively.

Figure 26: Spectrogram, waveform and pitch track for a modal vowel with level tone ‘lid’ [táp] . Uttered in sentence-medial position by Speaker 1.

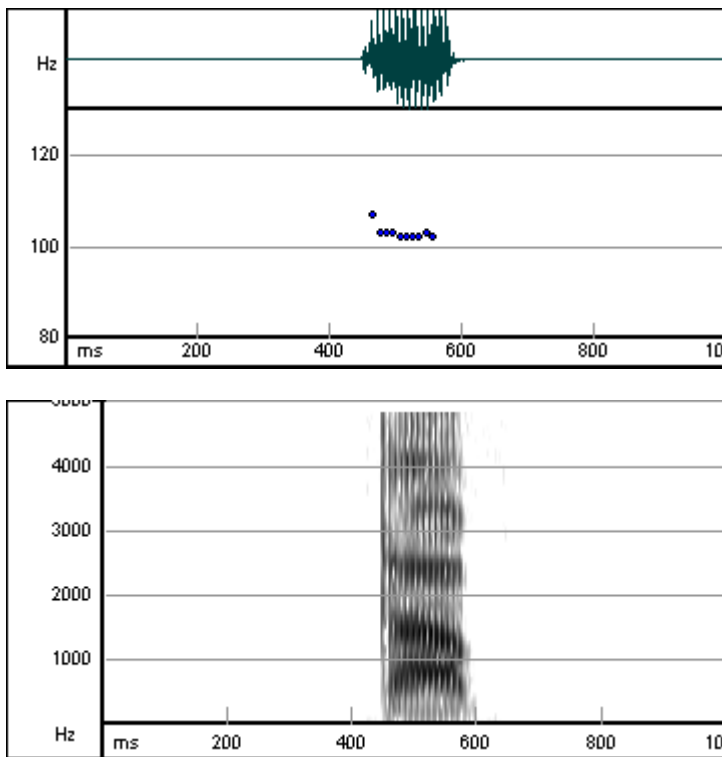
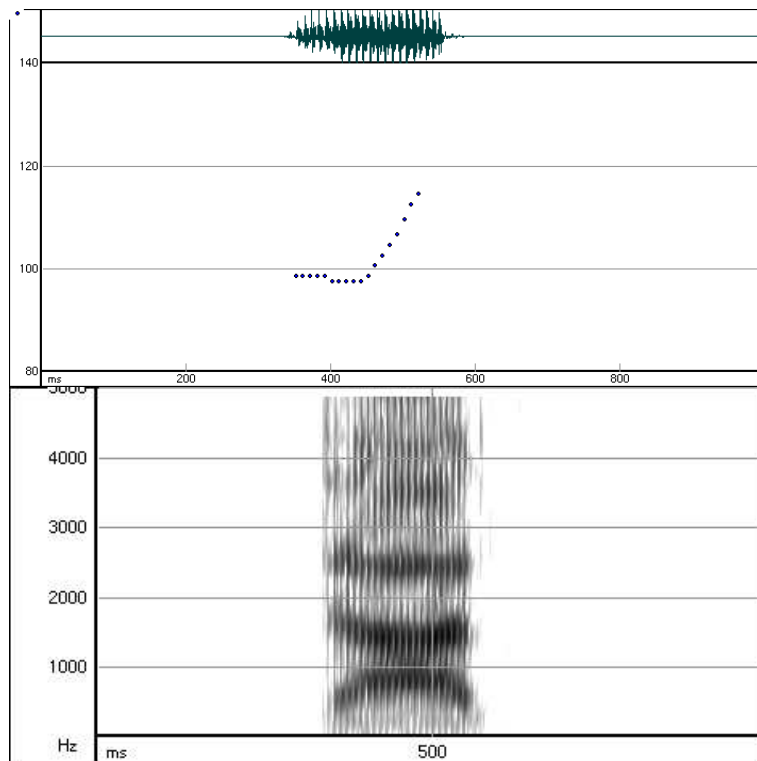


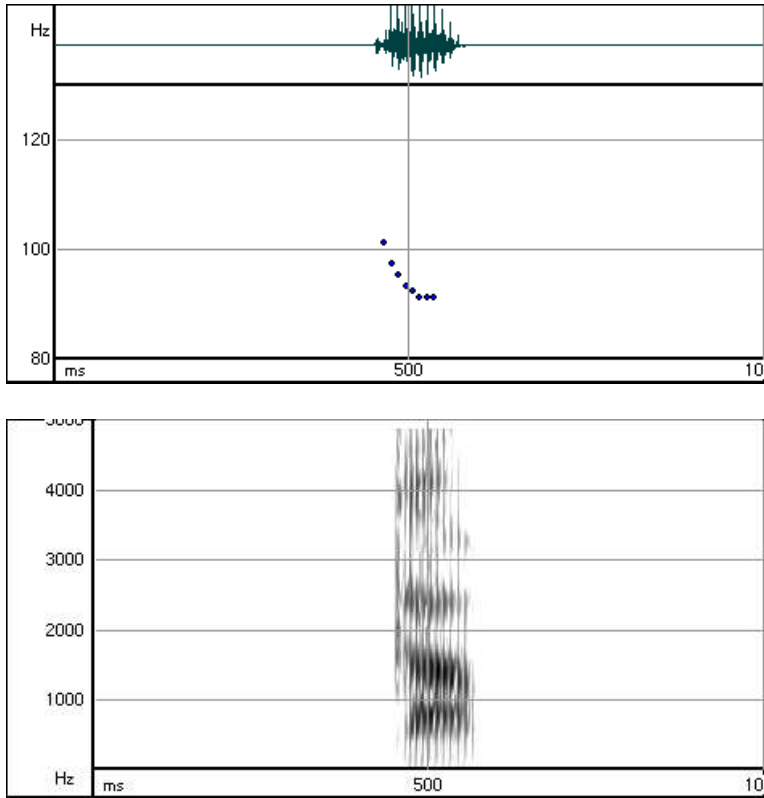
Figure 27: Spectrogram, waveform and pitch track for a modal vowel with high rising tone ‘father’ [dǎd] Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end.



4.4.2 Breathy

For the male speakers, breathy vowels are characterized by a decrease in the higher frequency energy and a slightly falling tone. A waveform and spectrogram of a breathy vowel ‘four’ [táp] appear in Figure 28. (Compare this to the waveform and spectrogram of ‘lid’ [táp] in Figure 26.) In comparison to modal phonation, the waveform for the breathy vowel is characterized by a decrease in energy.

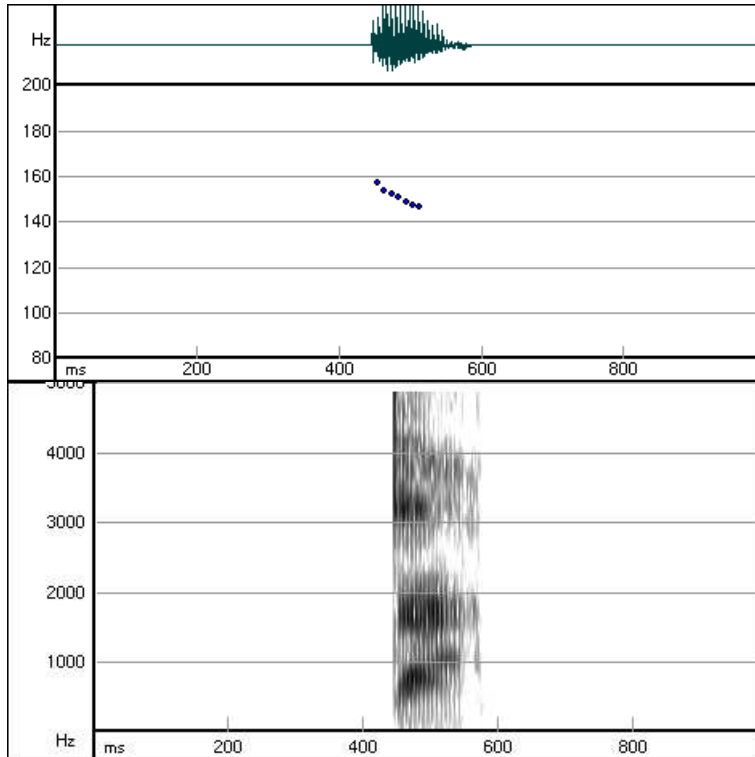
Figure 28: Spectrogram, waveform and pitch track for a breathy vowel ‘four’ [t̚əp]. Uttered in sentence medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.



4.4.2.1 Noise

The female speaker’s breathy vowels are characterized by a large open quotient, in addition to noise at the end of the vowel, which is absent in the male productions. This is evident in the spectrogram for ‘four’ [t̚əp] (Figure 29), where the end of the vowel is voiceless and marked by noise. Like the male speakers’, the female speaker’s breathy vowels are also accompanied by a falling tone.

Figure 29: Spectrogram, waveform and pitch track for a breathy vowel ‘four’ [t̪ap]. Uttered in sentence medial position by Speaker 3. Pitchtrack fails at end.



4.4.3 Creaky

Figure 30 and 31 contain a waveform, spectrogram and pitch track for the words “field” [l̪ats], for a male and female speaker, respectively. (Compare this to Figure 8, the male production of ‘can’ [lát], and Figure 11, the female production.) Creaky vowels are characterized by a falling tone. In creaky vowels, the creakiness is limited to the end of the vowel; the beginning of the vowel is modal. Thus, only the phonation measurement taken at the end of the vowel is different from modal phonation. On the spectrogram, the last portion of the vowel is marked with widely spaced vertical striations, reflecting the

low f_0 . In addition, the waveform for the creaky phonation “field” [l̥ɑts] is distinguished from the modal phonation by a decreased intensity.

Figure 30: Spectrogram, waveform and pitch track for a creaky vowel ‘field’ [l̥ɑts]. Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end.

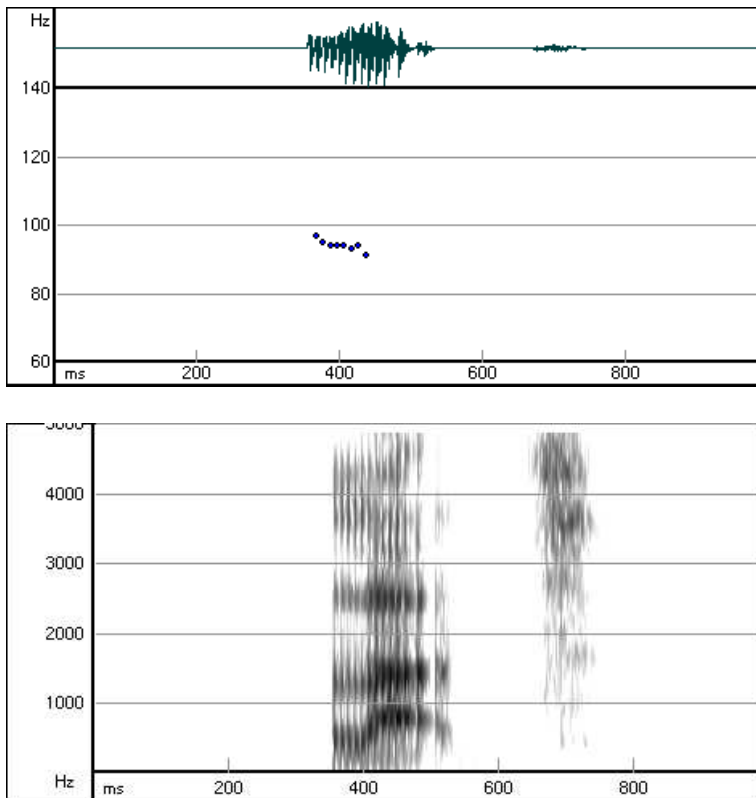
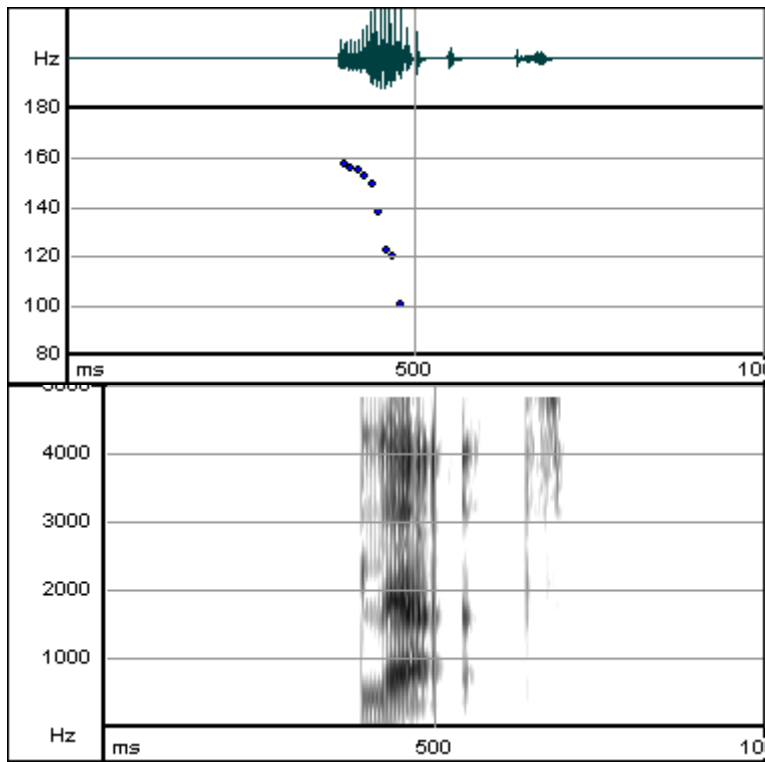


Figure 31: Spectrogram, waveform and pitch track for a creaky vowel ‘field’ [l̥aʔs]. Uttered in sentence-medial position by Speaker 3. Pitchtrack fails at end. There is a marker at 500 ms, which is not significant for this study.



5 Testing other Vowels

Since the measure H1-F3 is not affected by the formants of higher vowels, it was possible to expand the comparison to other vowels, and not limit the study to [a]. (It was not possible to examine the phonation of the other vowels for the female speaker since her measure of phonation was H1-H2. The measure H1-H2 is only successful as a measure of phonation on low vowels since the first formant of low vowels does not influence the amplitude of the first or second harmonics.) Thus, the results for [a] were checked against a subset of data from a male speaker, Speaker 1, for the other SADVZ vowels [i,e,o,u]. The vowel [ɨ] was omitted from this subset since it was difficult to find enough words that included this vowel.

5.1 Method

5.1.1 Speech Materials

Speaker 1 was asked to produce a list of words composed of five monosyllabic tokens per phonation/ tone group, per vowel [i,e,o,u] and a range of consonants. The wordlist was selected based on cognates in related languages. (See appendix for wordlist given in Tables 10-13.)

5.1.2 Procedure

The same procedure for measuring phonation and f_0 that was used in the first experiment (section 4.4) was repeated here; since the speaker is male, H1-F3 was used as the measure of phonation.

5.2 Results

The results of these measurements are presented in Figures 32, 34, 36, and 38. Just as with [a], there is a tonal distinction within modal phonation, but it is not accompanied by any significant difference in phonation (the tone groups are not distinguished in the phonation figures). The difference between the amplitudes of H1 and F3 is given on the y-axis. Figures 33, 35, 37, 39 are of the f_0 at three timepoints (initial, medial and final). Timepoints are represented on the x-axis and f_0 is represented in Hz on the y-axis.

Figure 32: H1-F3 for [e] (Speaker 1, medial position)

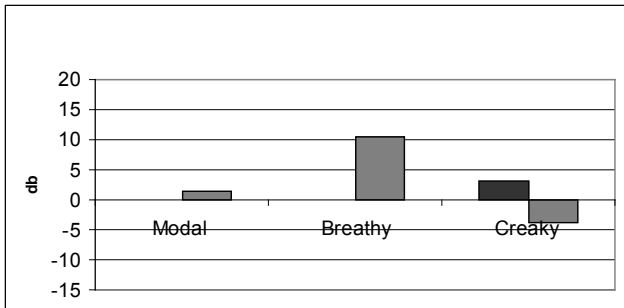


Figure 33: F0 for [e] (Speaker 1, medial position)

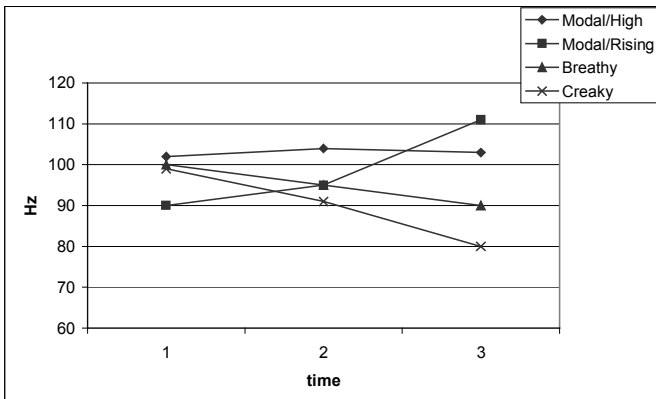


Figure 34: H1-F3 for [u] (Speaker 1, medial position)

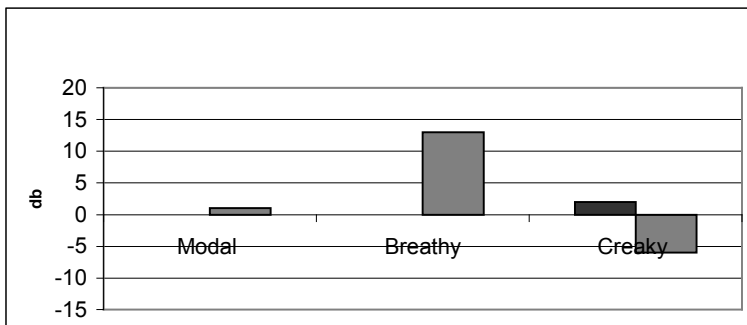


Figure 35: F0 for [u] (Speaker 1, medial position)

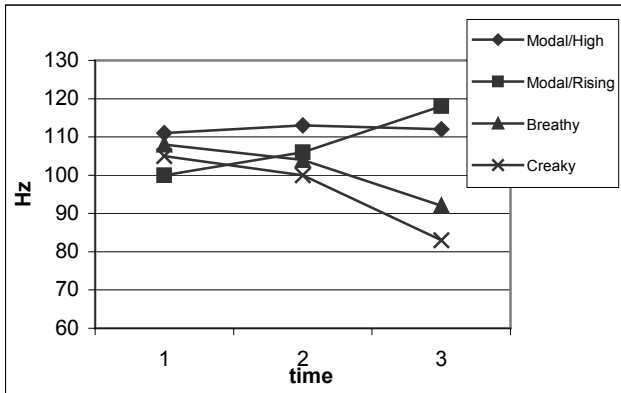


Figure 36: H1-F3 for [i] (Speaker 1, medial position)

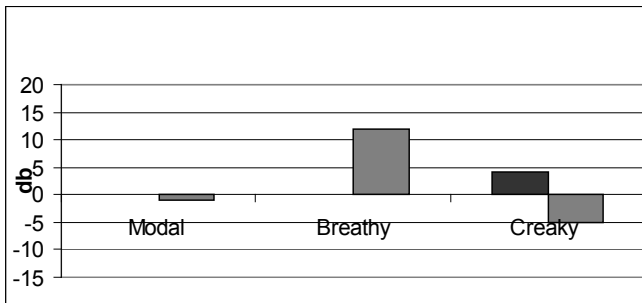


Figure 37: F0 for [i] (Speaker 1, medial position)

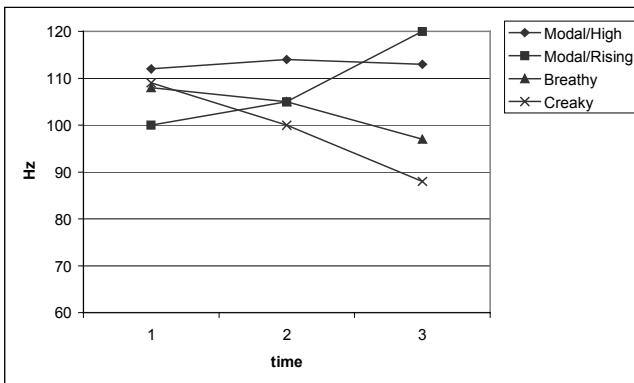


Figure 38: H1-F3 for [o] (Speaker 1, medial position)

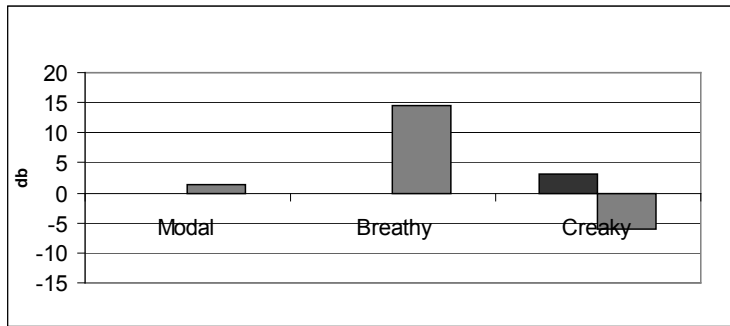
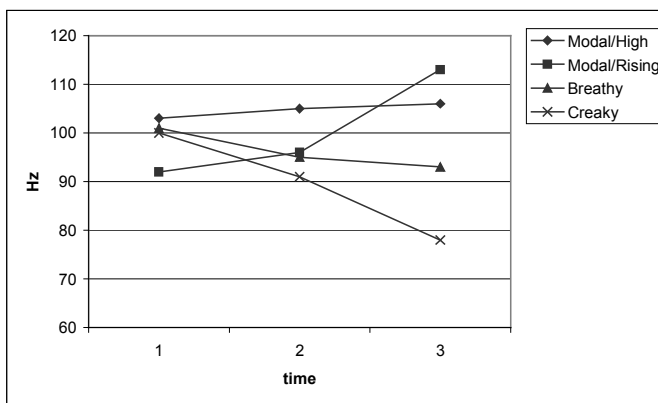


Figure 39: F0 for [o] (Speaker 1, medial position)



The measurements showed the same phonation contrasts for the vowels [i,e,o,u] as for [a]. Since the results for [i,e,o,u] were very similar, I will only describe the phonation associated with one vowel, [u].

5.2.1 Modal

Figures 40 and 41 show a waveform, spectrogram and pitch track of a modal vowel with a level high tone ‘Lucas’ [lúk] and a high rising tone ‘gopher’ [lǔʒ], respectively. Both are characterized by regularly spaced glottal pulses and normal intensity in the waveform.

Figure 40: Spectrogram, waveform and pitch track for a modal vowel with a level tone ‘Lucas’ [lúk] Uttered in sentence-medial position by Speaker 1. (The f0 is level during the vowel and low during the [l]).

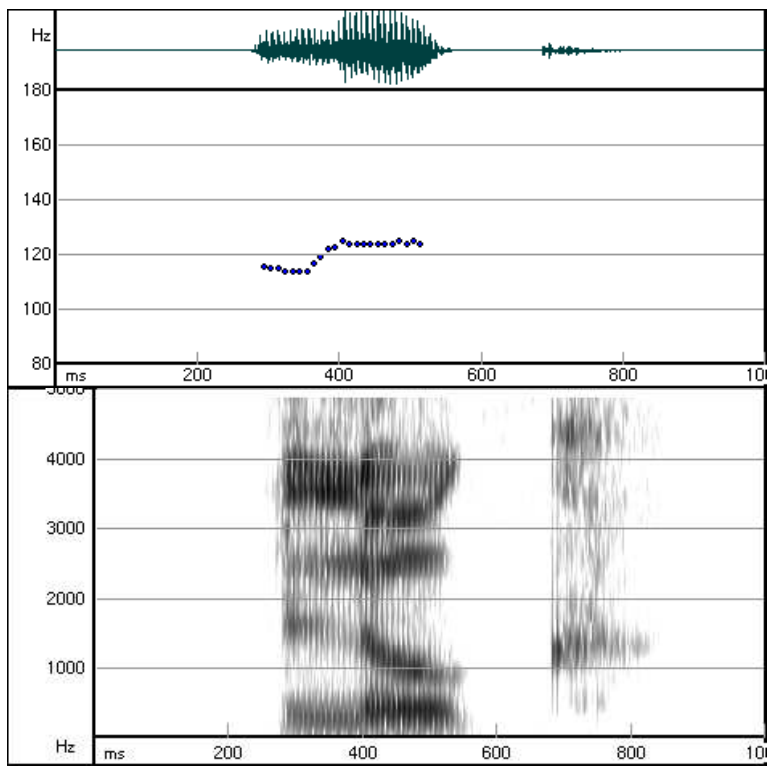
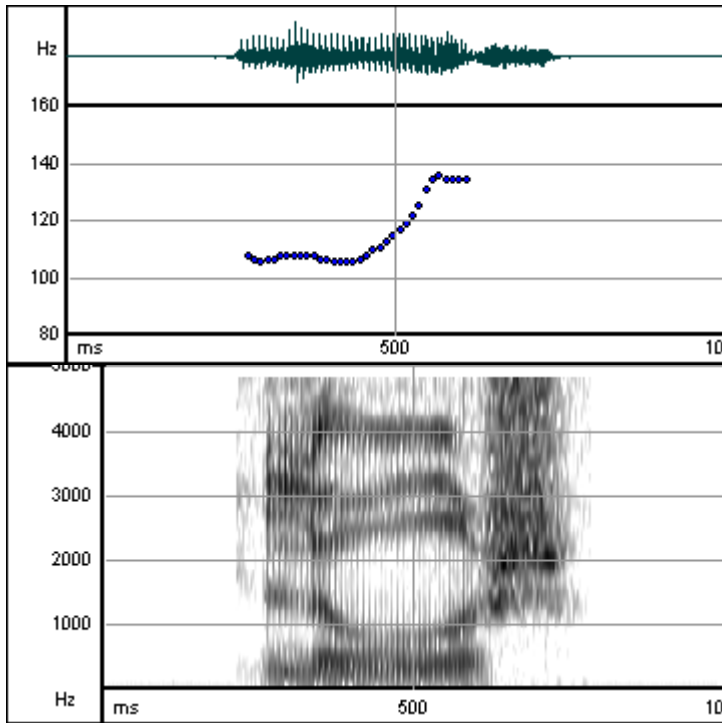


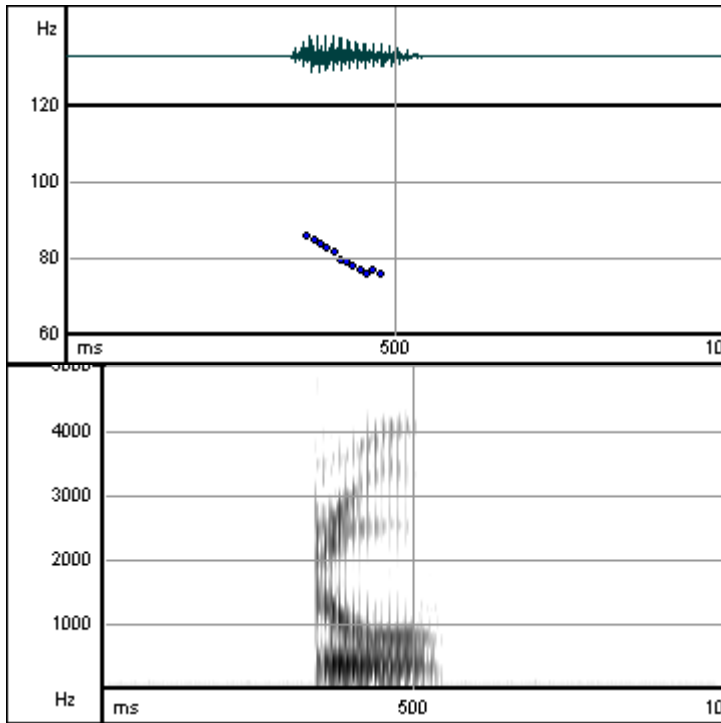
Figure 41: Spectrogram, waveform and pitch track for a modal vowel with a high rising tone ‘gopher’ [lʊʒ] Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.



5.2.2 Breathy

Figure 42 is a waveform, spectrogram and pitch track of a breathy [u] in ‘cough’ [ɾʊ̥]. As with the breathy [a], the [u] also has a slightly falling tone. In comparison to modal phonation (Figure 40), the breathy vowel is characterized by a decrease in intensity in the waveform and less high frequency energy, which is visible in the spectrogram.

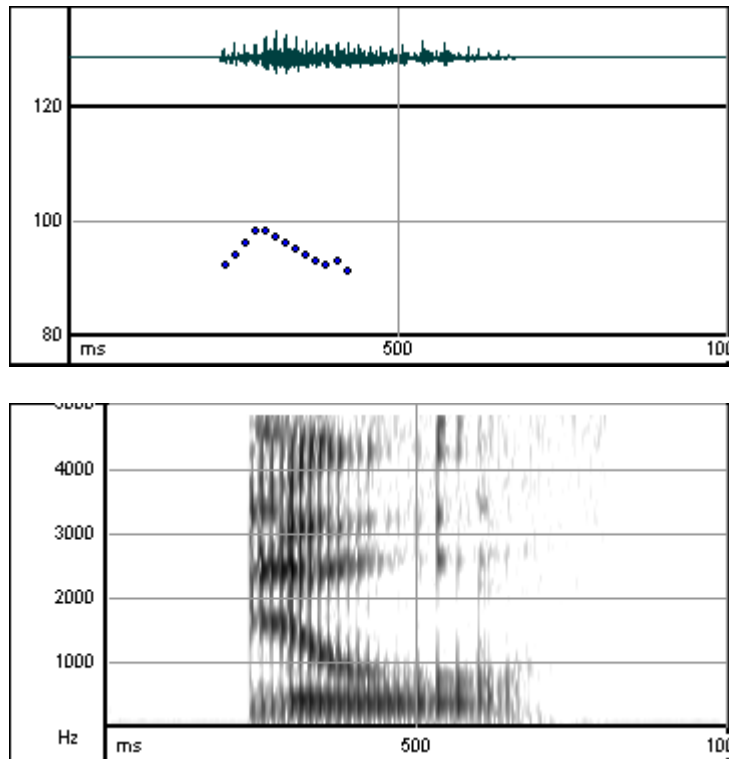
Figure 42: Spectrogram, waveform and pitch track for a breathy vowel in ‘cough’ [ɾ̥]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.



5.2.3 Creaky

Figure 43 contains a waveform, spectrogram and pitch track for “you” [l̥]. In the last portion of the vowel, the vertical striations are widely spaced, reflecting the low f_0 . Compared to modal phonation (Figure 40), the creaky vowel exhibits a decreased intensity in the waveform and a falling tone.

Figure 43: Spectrogram, waveform and pitch track for a creaky vowel ‘you’ [l̥]. Uttered in sentence medial position by Speaker 1. Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.



6 Polymorphemic, Disyllabic Words

Thus far, this study has concentrated on phonation in monosyllabic words. In this section, the experiment will be expanded to polysyllabic words to see how phonation is expressed in longer tokens and to determine if there is an effect of position-in-word on phonation.

6.1 Methods

6.2 Speech Materials

Speaker 1 was asked to produce a wordlist composed of 20 polymorphemic, disyllabic words (5 words per phonation/tone group), all with the target vowel [a] in both

syllables (see appendix for wordlist given in Table 9) and a range of consonants. In disyllabic words, the phonation contrasts are on the second syllable. The wordlist was selected based on cognates in related languages.

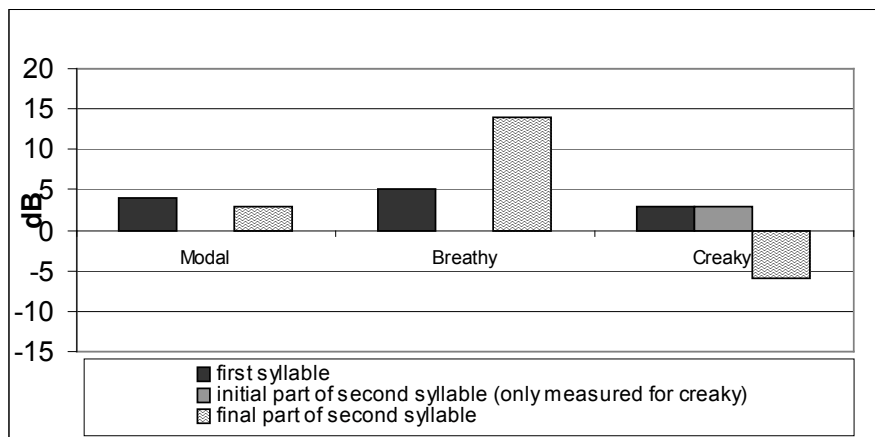
6.3 Procedure

Again, the tokens were read in the frame [*guni?* ___ *primer*] “Say ___ first.” For the first syllable, H1-F3 was measured for one point in the vowel. For the second syllable, which carries the contrast in phonation, the same procedure established in section 4 was followed here. In addition, f0 and duration were also measured following the procedure established in section 4.

6.4 Results

The results for H1-F3 are presented in Figure 44. The difference between the amplitudes of H1 and F3 (in dB) is given on the y-axis. The first syllable of the disyllabic words always has modal phonation; the second syllable has one of the three phonation contrasts. (The phonation category labeled on the x axis represents the phonation in the second syllable of the word.)

Figure 44: Graph of H1-H3 for the First and Second Syllables (Speaker 1, medial position)

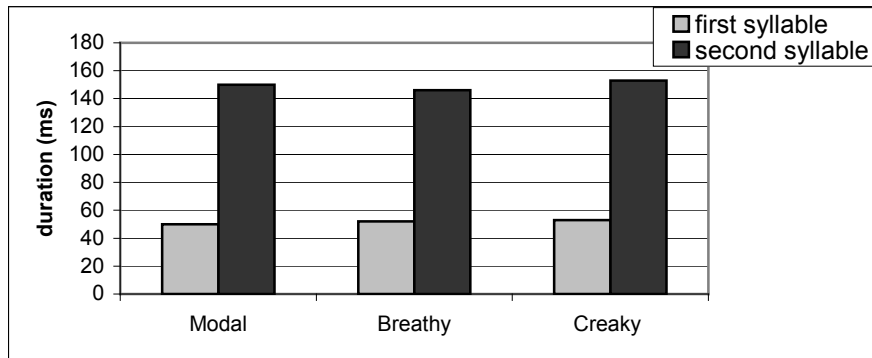


6.4.1 Other Differences

6.4.1.1 Duration

Duration was also measured (in ms) for the first and second syllable. The first syllable of the disyllabic words is always shorter than the second syllable. This is presented in Figure 45. Duration (in ms) is represented on the y-axis.

Figure 45: Duration for the First and Second Syllable (in ms) (Speaker 1, medial position)



6.4.1.2 f₀

F₀ of each syllable was measured from the FFT at three timepoints and is graphed in Figures 46- 47. Figure 46 shows the f₀ of the first syllable; Figure 47 shows the f₀ of the second syllable. Timepoints are represented on the x-axis and f₀ is represented in Hz on the y-axis. The first syllable of the disyllabic words has a fairly mid level phonation (approximately 100 Hz). The second syllable has one of the three tonal contrasts (high, high rising, or falling).

Figure 46: F0 for the First Syllable in Disyllabic Words. The phonation given in the legend corresponds to the phonation of the second syllable. (Speaker 1, medial position)

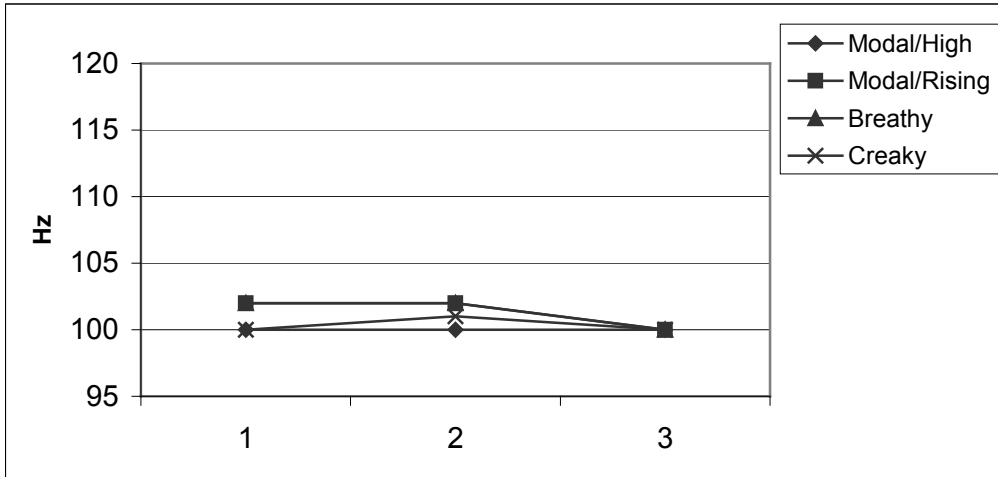
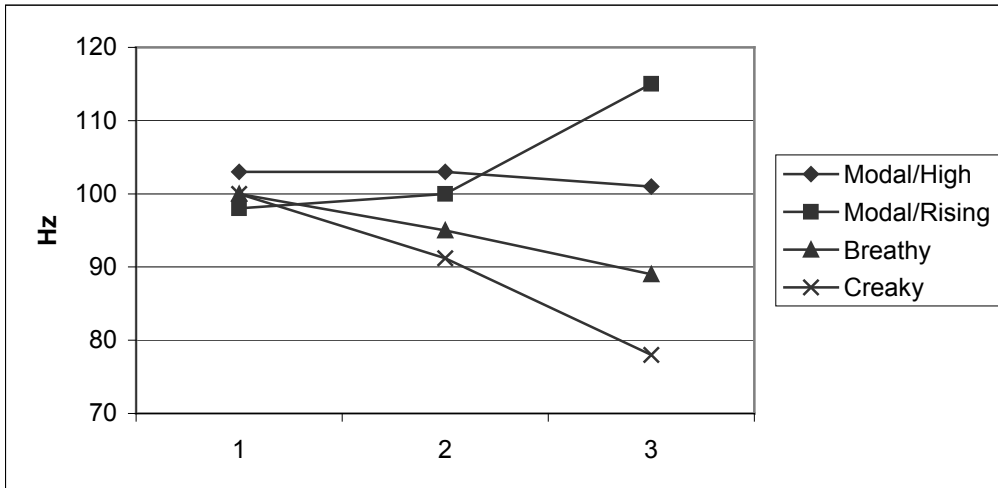


Figure 47: F0 for the Second Syllable in Disyllabic Words (Speaker 1, medial position)



The same pattern emerges for all the disyllabic words. The first syllable of the disyllabic words always has a fairly mid level tone (around 100 Hz), and is always shorter in duration than the second syllable. The first syllable always has modal phonation. The second syllable carries the phonation and tone contrast. These results are summarized by phonation in the next sub-sections.

6.4.2 Modal

In disyllabic words, the vowel of both the first and second syllable has a modal phonation, though the second vowel does have a longer duration than the first (see Figure 45).

As with the modal monosyllabic words, the second syllable of modal disyllabic tokens is also divided into two tonal patterns, either a high level or a high rising tone. Figure 48 is a pitch track, spectrogram and waveform of a modal token with a high level tone in the second syllable ‘Santa Ana’ [ʃandán]. In ‘Santa Ana’ [ʃandán], the first syllable has a level tone that is approximately 100 Hz. The second syllable is also level (100 Hz). Both vowels have a modal phonation.

Figure 49 is of a modal token with a high rising tone in the second syllable ‘cheap’ [baɾǎt]. In ‘cheap’ [baɾǎt], the first syllable has a level tone (100 Hz), but the second syllable has a high rising tone (100-120 Hz). Again, both vowels have a modal phonation.

Figure 48: Spectrogram, waveform and pitch track for ‘Santa Ana’ [ʃandán]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.

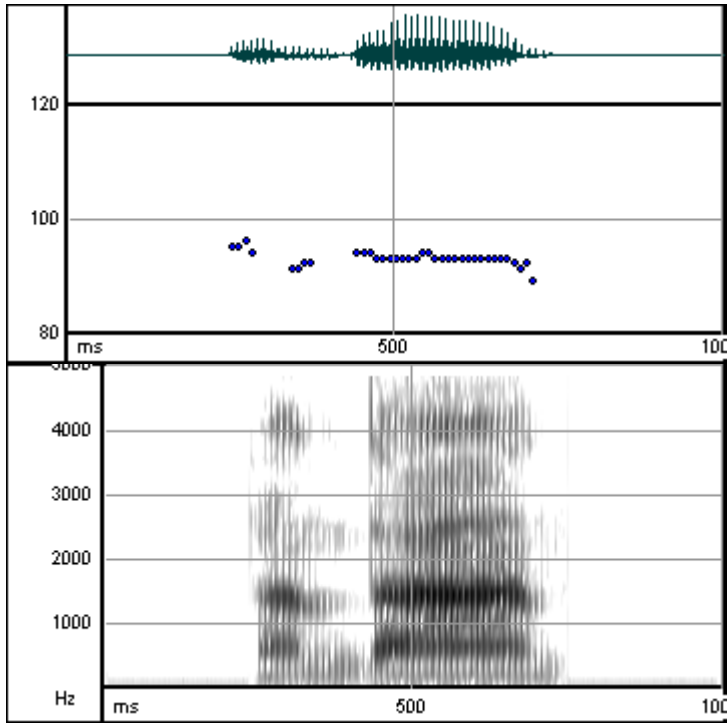
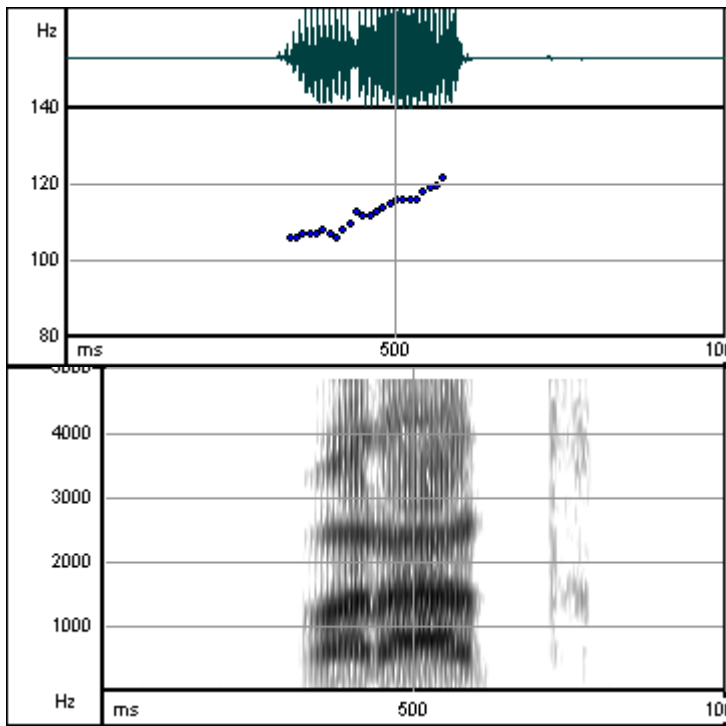


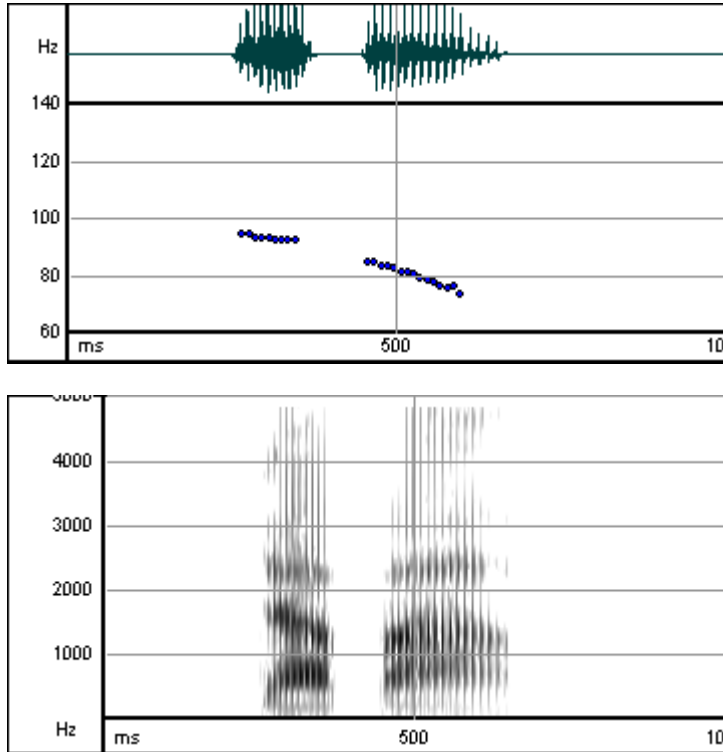
Figure 49: Spectrogram, waveform and pitch track for ‘cheap’ [ba.rǎt]. Uttered in sentence-medial position by Speaker 1. There is a marker at 500 ms, which is not significant for this study.



6.4.3 Breathy

Figure 50 is a waveform, spectrogram and pitch track for ‘itches’ [rəb̥əb]. The first syllable is characterized by a fairly modal phonation, and level pitch of approximately 100 Hz. The first syllable is also shorter than the second. The second syllable is where the breathy phonation (exemplified by the drop off in higher frequency energy) and falling tone is realized.

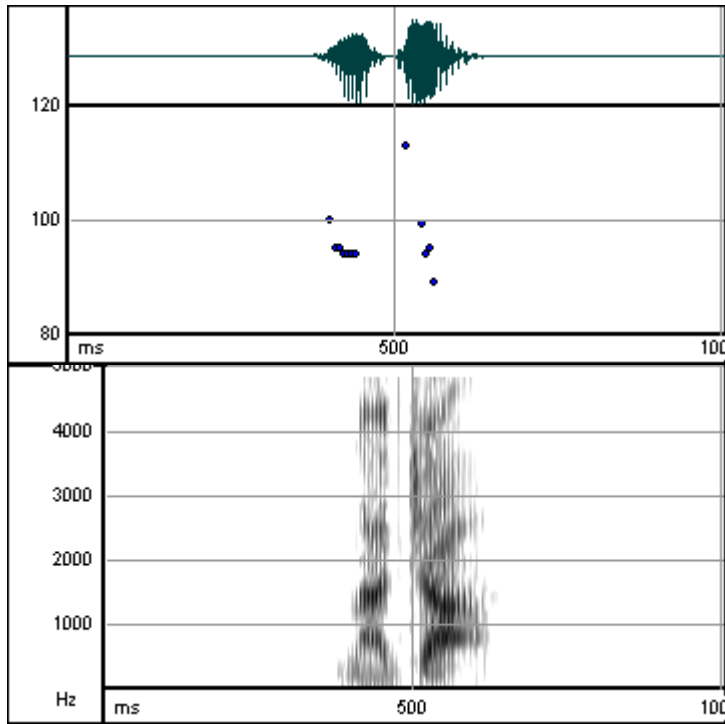
Figure 50: Spectrogram, waveform and pitch track for ‘itches’ [rabàb]. Uttered in sentence-medial position by Speaker 1. Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.



6.4.4 Creaky

Figure 51 is a waveform, spectrogram and pitch track for ‘warm’ [nadʒâ]. The first syllable, which is shorter than the second, is characterized by a fairly modal phonation, and a level f0 of approximately 100 Hz. The second syllable is where the creaky phonation and falling tone is realized. As with monosyllabic words, the creakiness is localized to the final portion of the vowel (which is evident by the widely spaced vertical striations towards the end of the vowel).

Figure 51: Spectrogram, waveform and pitch track for ‘warm’ warm’ [nadʒâ]. Uttered in sentence-medial position by Speaker 1. Pitchtrack fails at end. There is a marker at 500 ms, which is not significant for this study.



7 Summary

The measures used in this experiment have shown that SADVZ has a three-way contrast in phonation between breathy, modal and creaky vowels. In some ways, this analysis diverges from the different analyses of the Zapotec languages presented in section 2.1. No evidence was found to support a possible fourth “checked” phonation in that vowels followed by a glottal stop were not found to be distinct from modal or creaky vowels. For both monosyllabic and disyllabic words, only one phonation was found per syllable, except in the cases of creaky phonation, where the vowel begins with modal phonation and concludes with creak.

In addition, a strong relation between f_0 and phonation was found, unlike in some of the other Zapotec languages, which can have several independent tones and phonations. In SADVZ, modal phonation can be either high or rising, breathy phonation has a small falling tone, and creaky phonation, a large falling tone. Creaky vowels are sometimes just the falling f_0 , without spectral tilt.

There was an effect of position in word on phonation. In disyllabic words, the three-way phonation contrast and the tonal contrast is only seen on the second syllable. The first syllable is always modal, with a mid level tone.

8 Effects of f_0 and Position on Phonation

So far, this study has examined phonation in utterance-medial position, which has a generally mid-ranged f_0 . But, it has been shown in other languages that phonation is sensitive to changes in f_0 and position in utterance. For example, Epstein (2002) showed that position can effect phonation, independent of changes in f_0 , in English. Changes in f_0 are also likely to create changes in phonation since there is a correlation between phonation, vocal fold vibration and f_0 . A high f_0 is associated with vocal folds that have an increased length and tension; a low f_0 is linked to decreased length and tension in the vocal folds. Thus, one would expect breathy voice to be correlated with a low f_0 and tenser voice qualities to have a higher pitch. This is true of language such as Hindi, where breathy voice is associated with lowered f_0 (Ohala, 1973) and Jingpho, Lahu and Yi, where there is a higher f_0 for tenser vowels (Maddieson and Hess, 1987).

SADVZ provides an opportunity to study the effects of both f_0 and position-in-utterance on phonation, independently. As mentioned in section 1.3, SADVZ has relatively free word order. Thus, it is possible to change the position of a word in an

utterance and see if there are any effects on phonation. Since there is no phrase-final fall in f_0 , it is also possible to study the effects of position without a strong influence of f_0 . In addition, it is possible to examine the potential influence of f_0 on phonation, independently of position, through the examination of a variety of intonational contours.

In the second half of this study, I will concentrate on the effects of f_0 and position-in-utterance on phonation. But, first it is necessary to determine if f_0 , position-in-utterance, or both, could be affecting the phonation in SADVZ.

8.1 Effects of F0 versus Position

8.1.1 Method: Procedure

To determine whether f_0 or position-in-utterance influences phonation, Speakers 1 and 3 were asked to read a group of sentences that exhibited various intonational contours. (The intonation patterns of these sentence and examples of pitch tracks were provided in section 1.4.) Each sentence included a certain target word from each of the three phonations. (The phonation of the target words was determined in section 4. See appendix for a list of the target words in Table 14.) The group of sentences was composed of two sets. Set I (Table 6) was composed of sentences where the target words had different positions across the sentences, but had the same f_0 .

Table 6: Set I: Different position, but same f0

When the position is...	Compared to medial position the f0 is....	Sentence	Gloss
sentence-initial (focused)	high	[___ gunə Len]	'Elena saw the ___'
sentence-final (in a negative question)	high	[Teka gunə Len ___]	'Didn't Elena see ___?'

Set II (Table 7) was composed of sentences where the target words had the same position within the utterances, but a different f0.

Table 7: Set II: The same position, but different f0

When the position is ...	Compared to medial position the f0 is....	Sentence	Gloss
sentence-final (non-focused, in a statement)	low	gunə Len ___	'Elena saw ___'
sentence-final (in a negative question)	high	Teka gunə Len ___	'Didn't Elena see ___?'

Each sentence was repeated five times. Tokens were digitized and analyzed in PCQuirer at a sampling rate of 22050 Hz. To calculate the phonation, H1-F3 and H1-H2 were measured for the two sets, following the same procedure established in section 4.

8.1.2 Results

Figures 52-55 represent the results of H1-F3 and H1-H2. The difference between H1-F3 and H1-H2 in dB is presented on the y-axis.

Figure 52: Results of H1-F3. High f0, different position-in-utterance (Speaker 1)

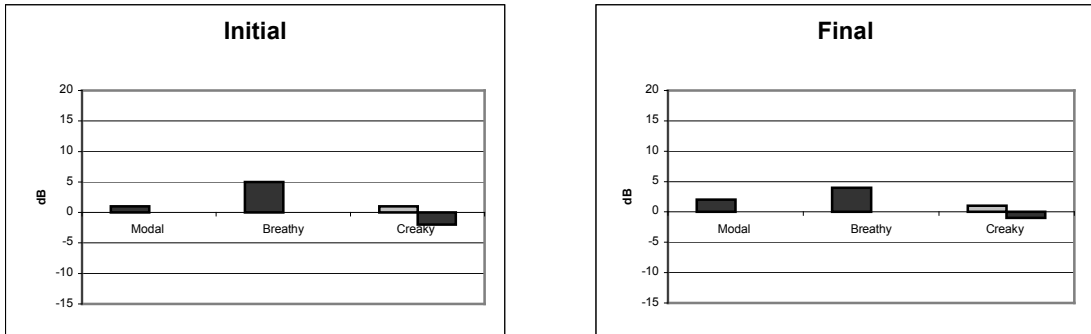


Figure 53: Results of H1-H2. High f0, different position-in-utterance (Speaker 3)

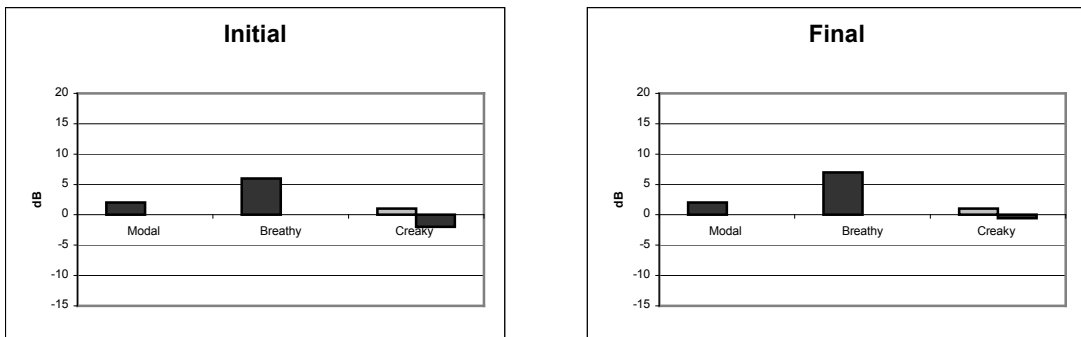


Figure 54: Results of H1-F3. Different f0, final position-in-utterance (Speaker 1)

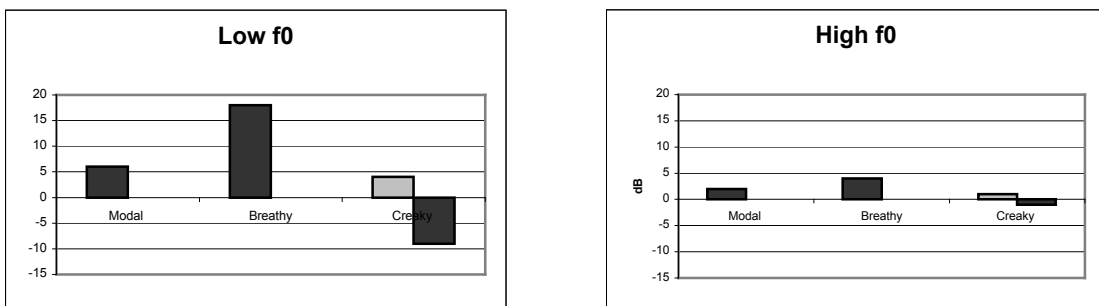
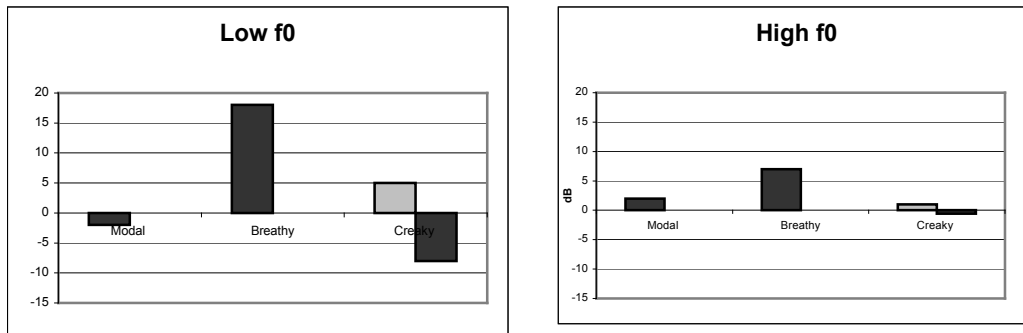


Figure 55: Results of H1-F3. Different f0, final position-in-utterance (Speaker 3)



The graphs show that there was an effect on phonation due to a change in f0, but not due to a change in position. While the three phonation categories are distinguished at low f0s, regardless of position, the three categories are not well distinguished at high f0s, regardless of position.

9 The Effects of F0 on Phonation

In the previous section, it was determined that f0 has an effect on phonation. In this next experiment, the effects of f0 on phonation will be examined in more detail.

9.1 Methods

9.1.1 Speakers and Speech Materials

All three speakers were asked to produce words from each of the phonation/tone groups. The phonation/tones of these words was determined in section 4. (See appendix for wordlist in Table 14.) This time, each token was read in three prosodic positions (isolation, sentence-initial position and sentence-final position). Tokens uttered in isolation have an overall higher f0 than sentence-medial position. In sentence-initial, focused position tokens also have a higher f0 than sentence medial position. These

tokens were uttered in the frame sentence [___ *guni?*] “___say (it)”. Sentence finally in the frame [*guni?* ___] “Say ___”, tokens had a lower f_0 than sentence medial position.

9.1.2 Procedure

Each sentence was recorded 5 times. Tokens were digitized and analyzed in PCQuirer at a sampling rate of 22050 Hz. H1-F3 and H1-H2 were measured for each token following the same procedure discussed in section 4. (Since the measures of phonation can be sensitive to changes in f_0 , spectrograms were used to confirm the results of the measures.)

9.2 Results for Isolation and Initial Position

Results showed that words read in isolation and in sentence-initial (focused) position had the same f_0 and the same phonations. I will begin by presenting the data for words produced in isolation, and will then present words produced in sentence-initial, focused position.

9.2.1 Isolation

In isolation, the three-way contrast in phonation is minimized and the f_0 is higher than in medial position. The results of these measurements for all three speakers can be found in Figures 56, 58, 60. The difference between the amplitudes of H1 and F3/H2 is given on the y-axis. The results of the measurements of f_0 are presented in Figures 57, 59, 61. Timepoints are represented on the x-axis and f_0 is represented in Hz on the y-axis.

Figure 56: H1-F3 (isolation, Speaker 1).



Figure 57: F0 (isolation, Speaker 1).

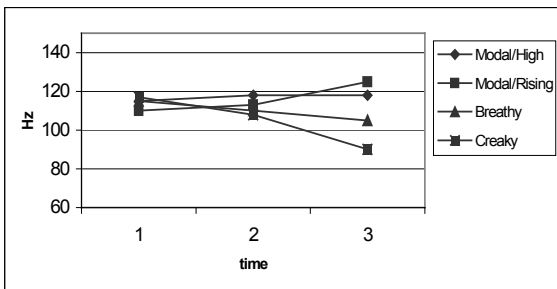


Figure 58: H1-F3 (isolation, Speaker 2).

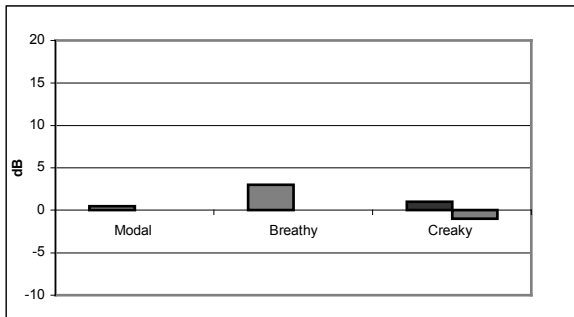


Figure 59: F0 (isolation, Speaker 2).

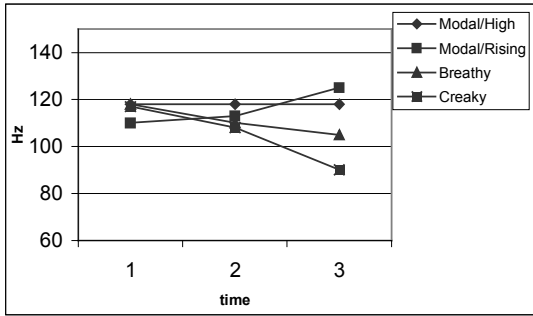


Figure 60: H1-H2 (isolation, Speaker 3).

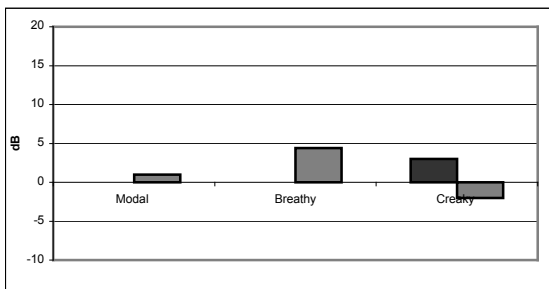
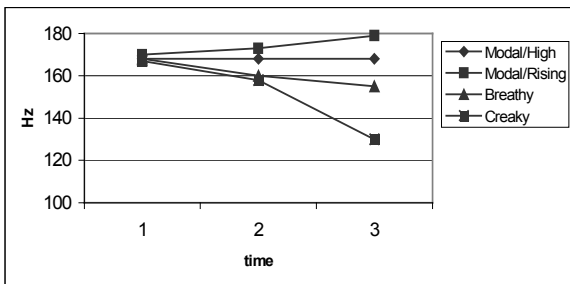


Figure 61: F0 (isolation, Speaker 3).



9.2.2 Initial (Focused) Position

In sentence-initial focused position, the three-way contrast in phonation is minimized and the f0 is higher than in medial position. The results of these measurements for all three speakers can be found in Figures 62, 64, 66. The difference between the amplitudes of H1 and F3/H2 is given on the y-axis. The results of the measurements of

f0 are presented in Figures 63, 65, 67. Timepoints are represented on the x-axis and f0 is represented in Hz on the y-axis.

Figure 62: H1-F3 (initial (focused) position, Speaker 1)

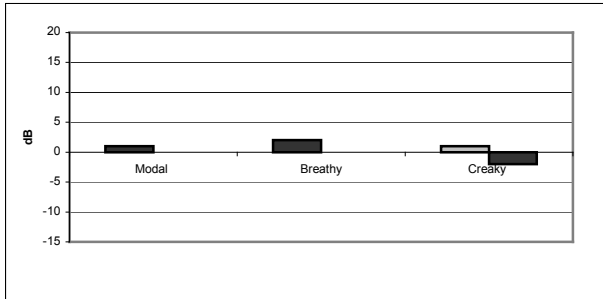


Figure 63: F0 (initial (focused) position, Speaker 1)

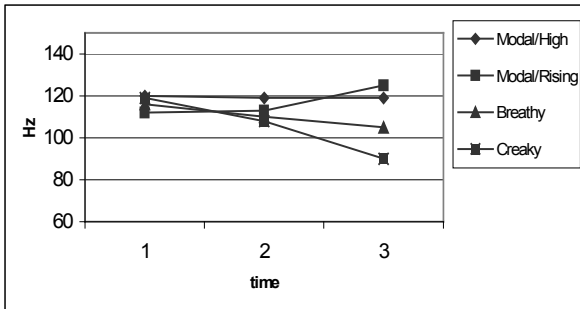


Figure 64: H1-F3 (initial (focused) position, Speaker 2)

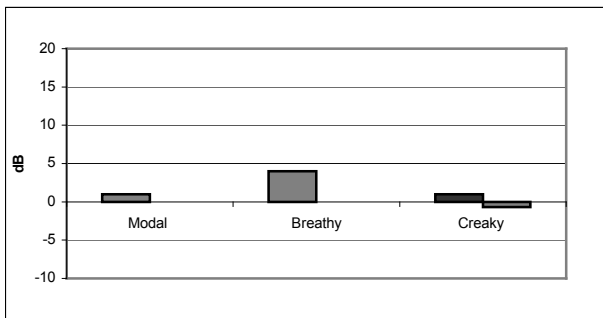


Figure 65: F0 (initial (focused) position, Speaker 2)

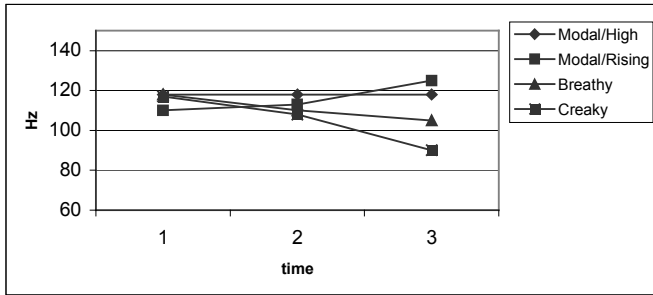
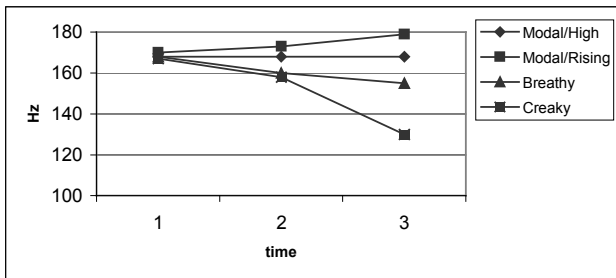


Figure 66: H1-H2 (initial (focused) position, Speaker 3)



Figure 67: F0 (initial (focused) position, Speaker 3)



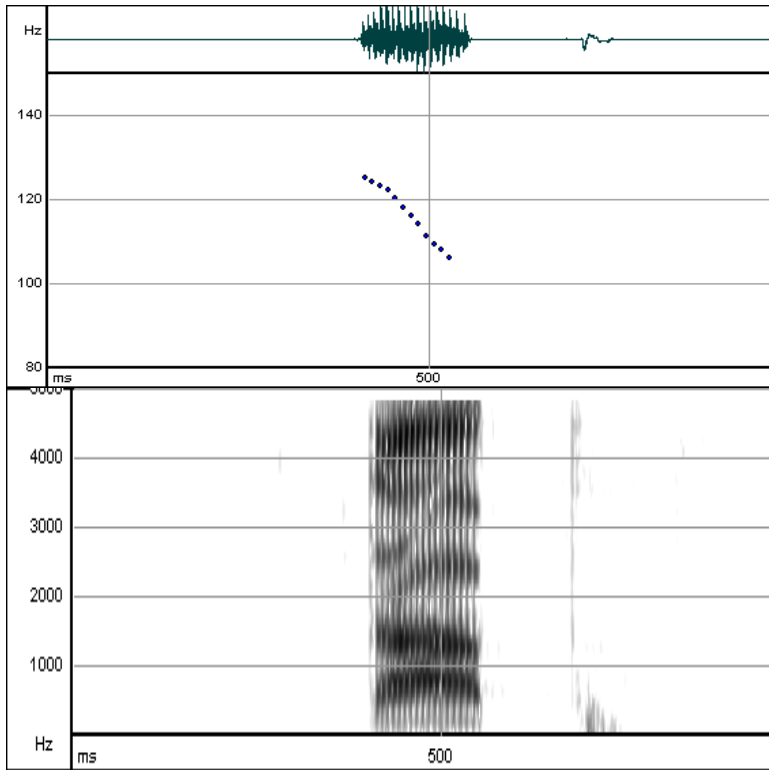
It was determined that tokens read in isolation have phonation and f0 most similar to tokens in sentence-initial position. Nouns read in isolation or in sentence-initial position are associated with an overall high pitch. The higher pitch minimizes the three-way contrast in phonation. In these cases, breathy and creaky words have a much more modal phonation, than when read in sentence-medial position. Modal phonation

remained relatively the same: no final fall in f_0 , the same phonation and either a high level or high rising tone. Thus, it will not be discussed in detail.

9.2.3 Breathy

In sentence-initial position, or isolation, there was an overall higher f_0 , which made breathy phonation more modal. Below is a waveform, spectrogram and pitch track of a breathy token 'four' [t̤əp] (Figure 68). (Compare this to Figure 28, 'four' [t̤əp] in medial position produced by the same speaker). In the example below, the vowel now has a more modal phonation. There is no longer a drop off in the higher frequency energies associated with breathy voice in medial position, nor the diminished intensity in the waveform. However, the falling tone is preserved, but overall higher than in sentence-medial position.

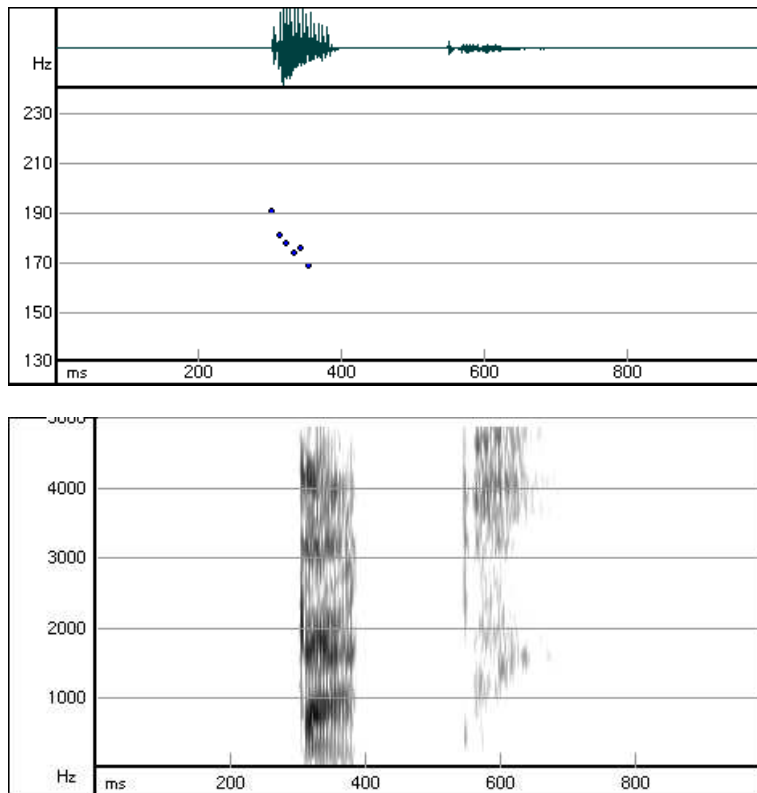
Figure 68: Spectrogram, waveform and pitch track for a breathy token ‘four’ four’ [t̚əp] pronounced in isolation (Speaker 1). Pitch track fails at end. There is a marker at 500 ms, which is not significant for this study.



9.2.3.1 Female Speaker

The female speaker (Speaker 3) also exhibits a change from breathy to modal phonation due to an increase in f_0 . Figure 69 is a spectrogram, waveform and pitch track for the word ‘four’ [t̚əp]. This token does not have the characteristic noise found in the breathy token when uttered with a mid range f_0 (Figure 29), but still has the falling tone.

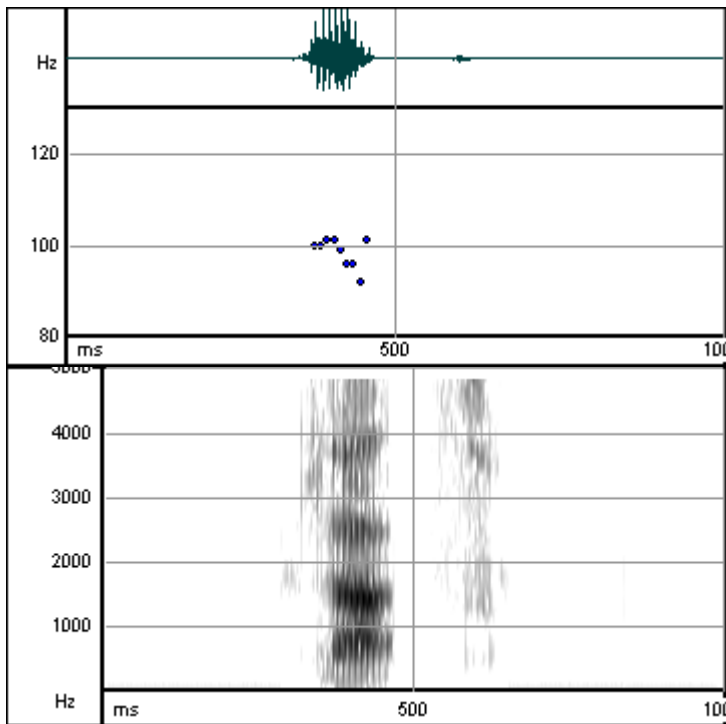
Figure 69: Spectrogram, waveform and pitch track for a breathy token ‘four’ [t̚ap] pronounced in isolation (Speaker 3).



9.2.4 Creaky

When a creaky vowel is pronounced with an overall higher f_0 , it results in a more modal phonation. Figure 70 is a waveform, spectrogram and pitch track of a creaky token [l̚ats] ‘field’ uttered in isolation. The vowel has a more modal phonation than when the token was pronounced in medial position (compare this to Figure 30). In the spectrogram in Figure 70, the widely spaced vertical striations typically associated with creaky voice are absent. In addition, the waveform has an overall greater energy than the creaky token uttered in medial position (Figure 30). However, the falling tone associated with the creaky vowels in medial position is preserved.

Figure 70: Spectrogram, waveform and pitch track for a breathy token ‘field’ [l̥âts] pronounced in isolation (Speaker 1). There is a marker at 500 ms, which is not significant for this study.



9.3 Results for Final Position

In sentence-final position, tokens are pronounced with a generally lower f0 and the three-way contrast in phonation is preserved. The results of the measurements H1-F3 and H1-H2 are presented in Figures 71, 73, and 75. The difference between the amplitudes of H1 and F3/H2 is given on the y-axis. The results of the measurements of f0 are presented in Figures 72, 74, and 76. Timepoints are represented on the x-axis and f0 is represented in Hz on the y-axis

Figure 71: H1-F3 for [a] (final position, Speaker 1)

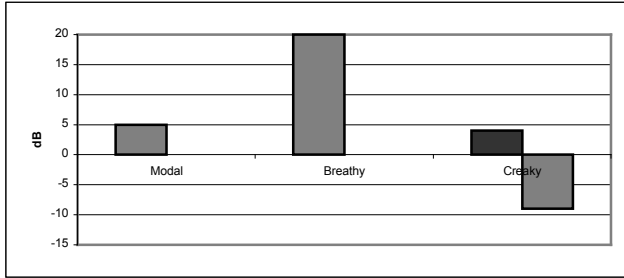


Figure 72: F0 (final position, Speaker 1)

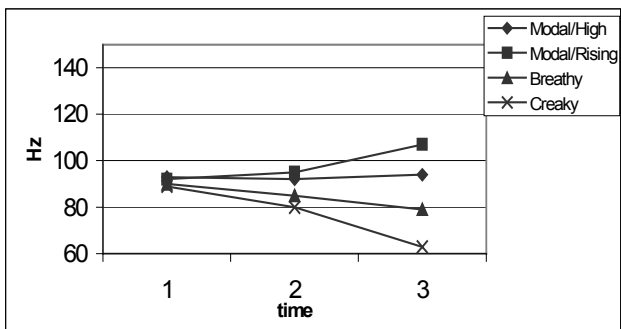


Figure 73: H1-F3 for [a] (final position, Speaker 2)

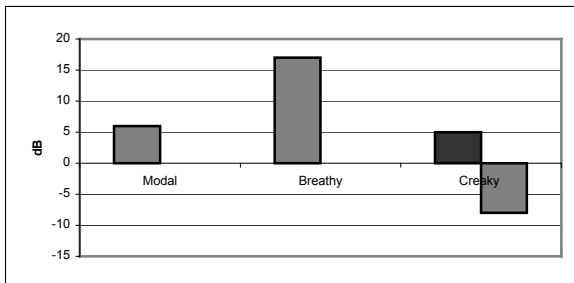


Figure 74: F0 (final position, Speaker 2)

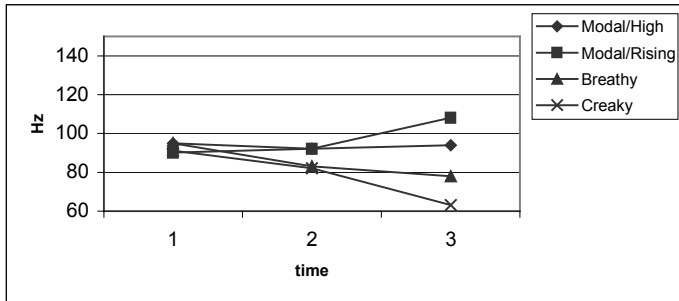


Figure 75: H1-H2 for [a] (final position, Speaker 3)

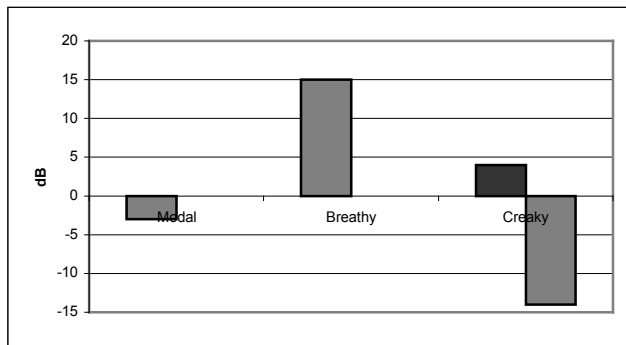
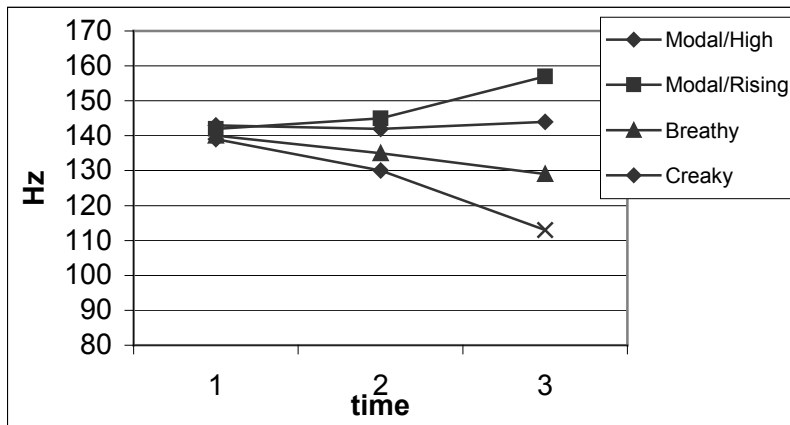


Figure 76: F0 (final position, Speaker 3)



Words in sentence-final position are associated with an overall lower pitch than words read in sentence-medial position. For the male speakers (1, 2), creaky vowels with lower f_0 were creakier and the modal and breathy vowels with lower f_0 were breathier. For the female speaker (3), with the lower f_0 the breathier vowels were breathier, and modal and creaky vowels were creakier.

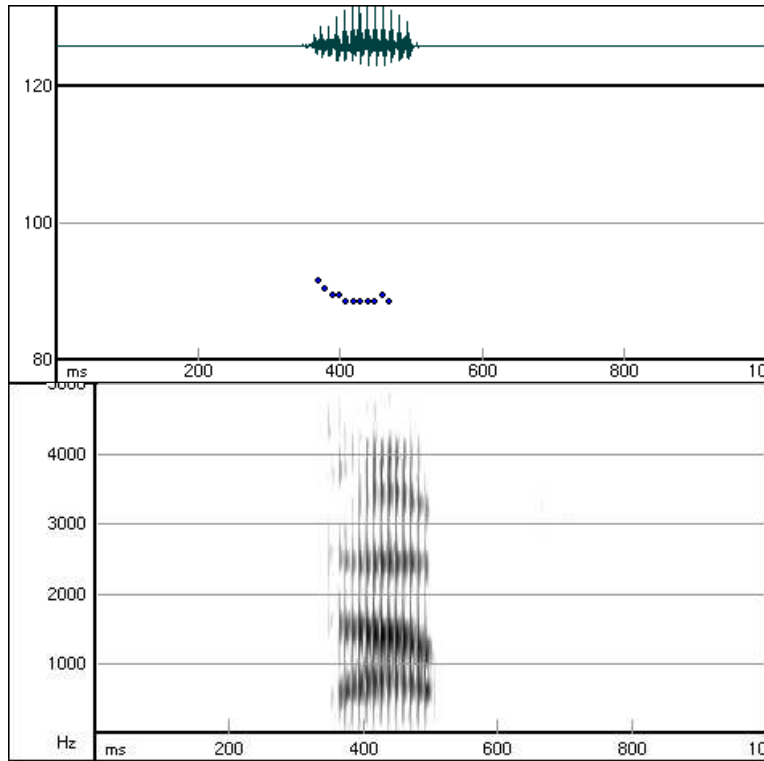
9.3.1 Modal

9.3.1.1 Male Speakers (Speakers 1 and 2)

When a modal³ vowel is pronounced with a lower f_0 , it has a breathier phonation, though the level tone associated with modal phonation remains. Figure 77 is a waveform, spectrogram and pitch track of a modal token [táp] ‘lid’ from the end of a sentence, and thus produced with a lower f_0 . This token has a breathier phonation than when pronounced in medial position (compare this to Figure 26). In the spectrogram in Figure 77, there is a fall off in the high frequency energy, a typical characteristic of breathy voice. There is also a reduction of energy in the waveform, more typical of non-modal phonation.

³ The lowered f_0 has the same influence on modal vowels with level tone and high rising tone. Since the results are the same, I will only discuss the modal phonation with the level tone in detail here.

Figure 77: Spectrogram, waveform and pitch track for a breathy token ‘lid’ [táp] pronounced in final position (Speaker 1). Pitch track fails at end.

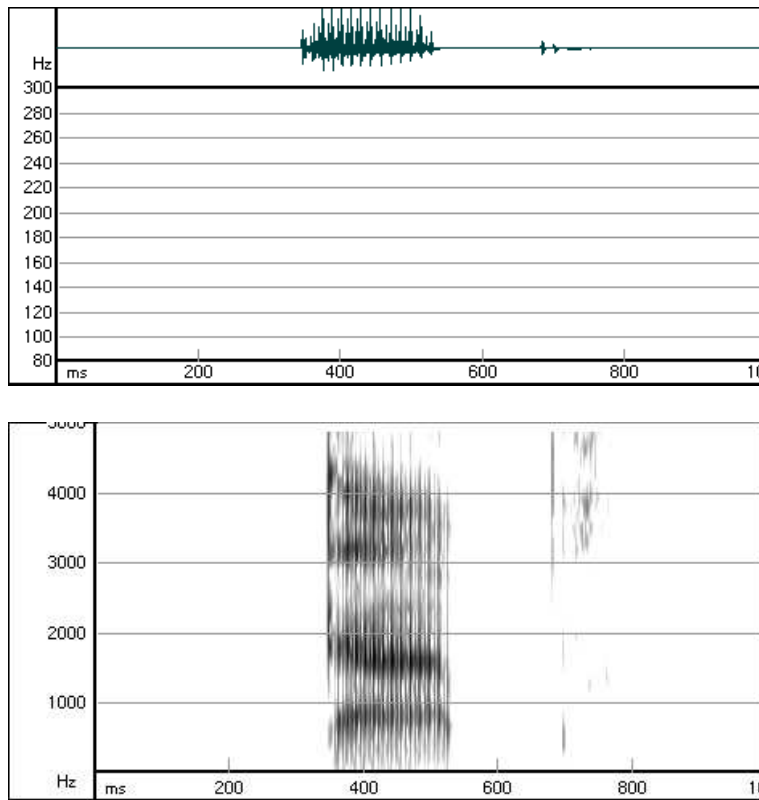


9.3.1.2 Female Speaker (Speaker 3)

For the female speaker (3), modal phonation becomes creaky with a lowered f_0 .

Figure 78 is a waveform, spectrogram and pitch track of a modal token taken from the end of a sentence, and thus produced with a lower f_0 . This token has a creakier phonation than when pronounced in medial position. The vowel is marked by widely spaced vertical striations. There is also visible period doubling, which is commonly associated with creaky phonation, in the waveform and spectrogram.

Figure 78: Spectrogram, waveform and pitch track for a breathy token ‘lid’ [táp] pronounced in final position (Speaker 3). Pitch track fails completely.

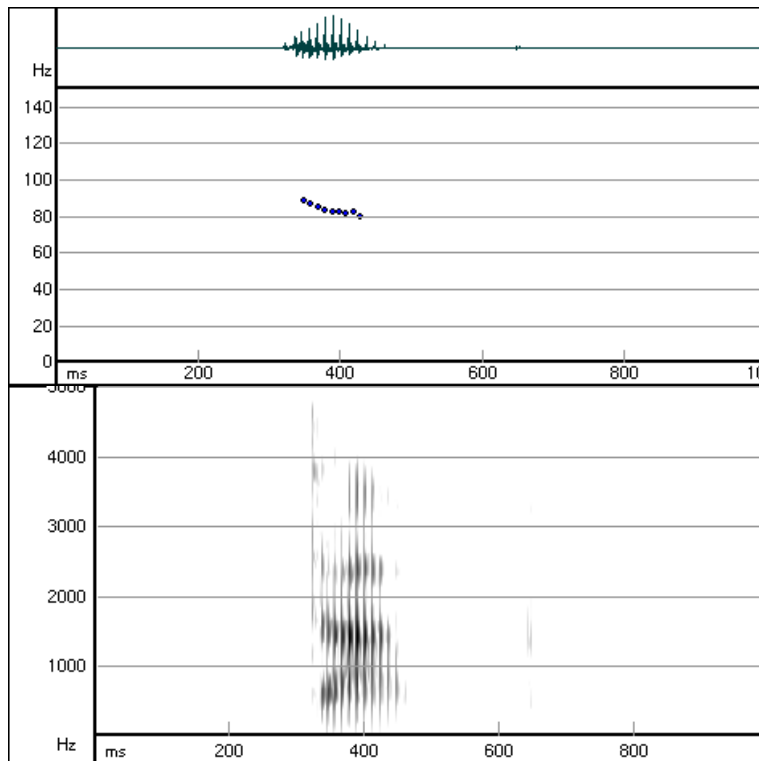


9.3.2 Breathy

In sentence final position, the lower f_0 made breathy phonation even breathier.

Below is a waveform, spectrogram and pitch track of a breathy token ‘four’ [tə̀p] (Figure 79). Compared to Figure 28, ‘four’ [tə̀p] (in medial position), the example below has an even greater drop off in the higher frequency energies, which is present for the entire length of the vowel.

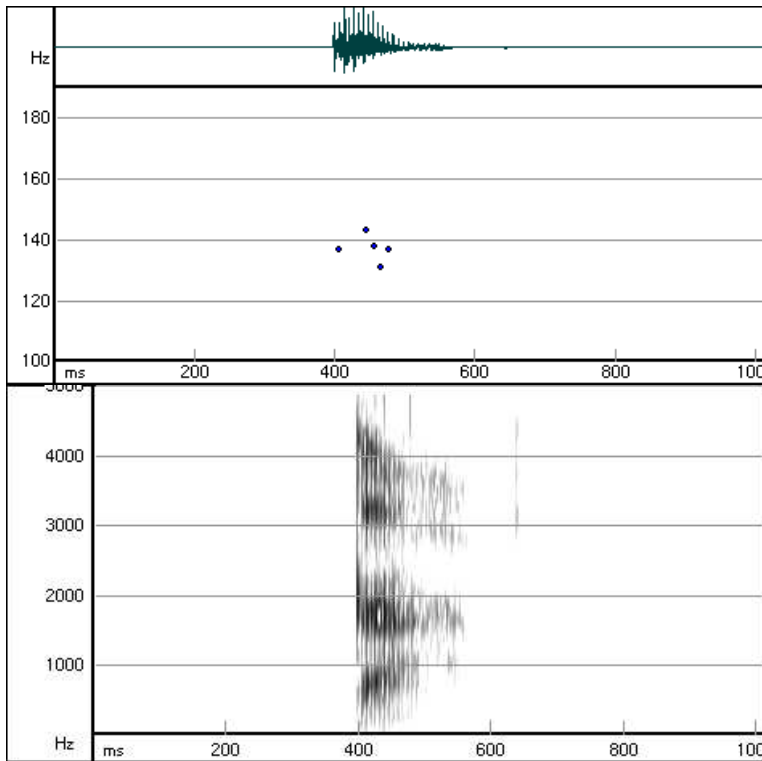
Figure 79: Spectrogram, waveform and pitch track for a breathy token ‘four’ [t̪əp] pronounced in final position (Speaker 1). Pitch track fails at end.



9.3.2.1 Female Speaker (Speaker 3)

Figure 80 is a waveform, spectrogram and pitch track of a breathy token uttered in final position for the female speaker. In the spectrogram, a greater portion of the vowel is marked by voicelessness and noise than when the same token was pronounced sentence medially (Figure 29).

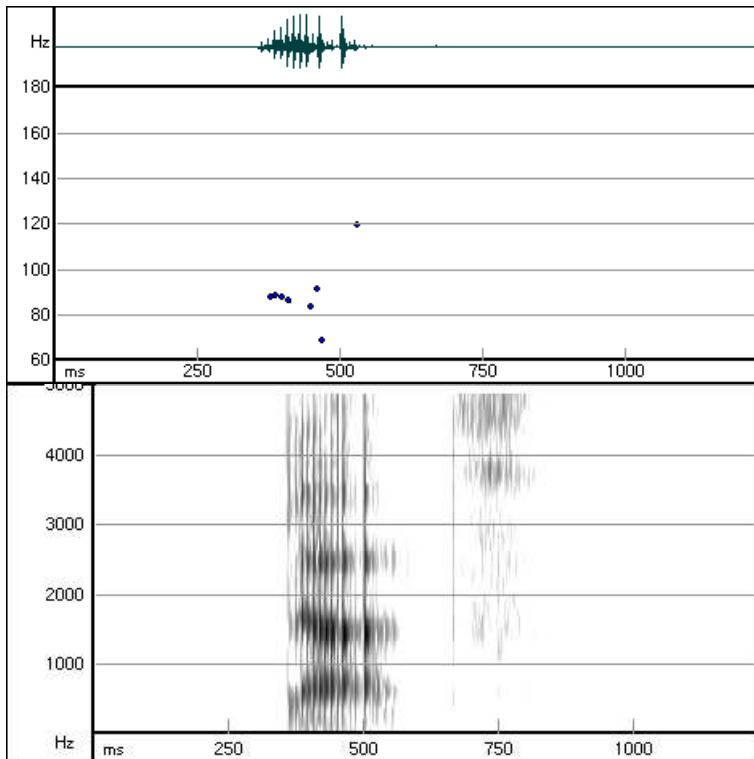
Figure 80: Spectrogram, waveform and pitch track for a breathy token ‘four’[t̪əp] pronounced in final position (Speaker 3). Pitch track fails at end.



9.3.3 Creaky

When a creaky vowel is produced with a lower f_0 , it results in an even creakier phonation. Figure 81 is a waveform, spectrogram and pitch track of a creaky token ‘field’ [l̪əts] uttered in final position. In the spectrogram, the vertical striations are more widely spaced and more numerous than when the same token was pronounced sentence medially (Figure 30). In final position, a greater portion of the vowel is creaky and the individual pitch pulses are widely separated. In addition, the energy in the waveform is more diminished than in the creaky token uttered in medial position.

Figure 81: Spectrogram, waveform and pitch track for a breathy token ‘field’ [lâts] pronounced in final position (Speaker 1). Pitch track fails completely.



10 Summary and Conclusion

This thesis has investigated the issue of phonation in Santa Ana del Valle Zapotec, a member of a language family where phonation contrasts are unusual and complex.

Below, I will summarize the major findings of this study.

In the first part of this study, phonation was examined at a mid-range f_0 , in utterance-medial position. Results showed that the phonations of SADVZ can be adequately described by modal, breathy and creaky categories. No evidence was found to support a fourth (checked) phonation type. As far as could be determined, vowel quality did not have an effect on phonation. In addition, there was no change in phonation within a

syllable except in the case of creaky voice, which begins modal and concludes with creak.

Position within a word did have an effect on phonation since the initial syllable of disyllabic words did not have a phonation contrast. The first syllable of disyllabic words always has a fairly mid level tone (around 100 Hz, for the male speaker), and a shorter duration than the second syllable. The first syllable always has modal phonation. The second syllable carries one of the phonation types and tones. It remains to be seen how this applies in multi-morphemic words with affixes that create a change in phonation.

In addition, there is a strong relationship between f_0 and phonation. Modal phonation had either a high or high rising f_0 . Breathy phonation always had a slightly falling f_0 ; creaky phonation had a largely falling f_0 . These relations were investigated further in the second part of the study, where the influence of f_0 and position-in-utterance on phonation were closely examined. It was determined that position-in-utterance (independent of f_0) does not have an effect on phonation, but that changes in f_0 do alter phonation. For example, it was found that in isolation or sentence-initial position, where the f_0 is high, the three-way contrast in phonation was minimized. Words that were breathy or creaky in the mid range f_0 , now had modal phonation. In sentence-final position, when the tokens had a lower f_0 , the contrast between modal, breathy and creaky voice was preserved. For the male speakers, modal and breathy words were breathier; for the female speaker, modal and creaky words were creakier and breathy words were breathier.

The results obtained in this study raise an interesting issue for fieldworkers working on Zapotec languages. In SADVZ, when f_0 is high the phonemic three-way contrast in phonation is not well preserved. Thus, words elicited in isolation, which has a higher f_0 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 displays a full range of f_0 s. It is possible that the effect of f_0 on phonation is one of the reasons that phonations in Zapotec languages are not completely understood.

The question arises as to whether tone or phonation is in any way more basic than the other; perhaps only one should be considered to be contrastive, and the other only an enhancer. There are two arguments that tone is more basic. It seems that when the phonation is weakened, the tonal pattern remains, preserving some level of distinction. For example, in sentence-initial position, breathy phonation becomes modal, but the falling tone remains. Thus, the tonal patterns seem to act as an enhancing feature of the phonation differences. In addition, creaky vowels are sometimes just the falling f_0 , without spectral tilt. No arguments have been found that phonation is more basic, as no cases have been found where phonation is preserved when the tone is compromised.

The results obtained in this study show that there is a simple three-way contrast in phonation in SADVZ. This study was unsuccessful in its goal of describing and understanding the more complex phonations typically associated with Zapotec languages. On closer examination, SADVZ did not prove to be a good example of the unusual and fascinating phonations that are reported for other Zapotec languages.

Although the measures used in this study have shown that there is a simple three-way contrast in phonation without changes in phonation within a syllable, there are additional questions that remain to be answered for SADVZ, and Zapotec languages in general. It remains to be seen how SADVZ compares with other Zapotec languages describe in the literature, especially a language with complex phonation contours. In addition, this study has shown that there is a strong effect of sentence-level f_0 on phonation. Further research is necessary to understand the effects of f_0 in more natural discourse contexts.

Appendix

Table 8: Monosyllabic words with [a]

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[páʔp]	‘potato’	[dǎd]	‘father’	[təp]	‘four’	[dəʔa]	‘petate’
[kár :]	‘car’	[bǎd]	‘duck’	[bəd]	‘Tlacolula’	[ləʔa]	‘huaje bean’
[pág]	‘pay’	[ləʔs]	‘skinny’	[sək]	‘price’	[lâts]	‘field’
[fáb]	‘detergent’	[stǎmp]	‘stamp’	[bədʒ]	‘tepache’	[dʒəʔp]	‘girl’
[táp]	‘lid’	[kwǎʔtʃ]	‘twin’	[dǎny]	‘mountain’	[bəʔa]	‘tomb’
[lát]	‘(tin) can’	[mǎn]	‘animal’	[rəp]	‘care’	[bəd]	‘scabies’
[rán]	‘frog’	[dǎm :]	‘owl’	[n : əʃ]	‘chocolate’	[ʒən]	‘butt’
[bál]	‘stick’	[dʒǎn]	‘saint’	[lədy]	‘clothes’	[bəʔt]	‘skunk’
[bág]	‘cow’	[fǎld]	‘skirt’	[nəl]	‘thick’	[gəʔa]	‘nine’
[káʔm :]	‘bed’	[mǎrkw]	‘mark’	[nən]	‘cold (n.)’	[dətʃ]	‘empty’
[nəʔtʃ]	‘rough’	[rəp]	‘have’	[nəd]	‘to stick’	[gəts]	‘hidden’

Table 9: Disyllabic words

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[barət]	‘cheap’	[mansán]	‘apple’	[bal : əg]	‘caterpillar’	[naləz]	‘sad’
[kampǎn]	‘bell’	[ʃandán]	‘Santa Ana’	[batʃə]	‘fill’	[natʃəʔ]	‘warm’
[kansǎd]	‘tired’	[nal : á]	‘hot’	[rabəb]	‘itches’	[batʃə]	‘crow’
[kapǎd]	‘castrated’	[sabád]	‘shoe’	[nagəts]	‘yellow’	[ʒaʒən]	‘butt’
[matǎd]	‘scab’	[kanáʔst]	‘basket’	[rabən]	‘bangs’	[nadəʔ]	‘loose’

Table 10: Words with [e]

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[béd]	‘Peter’	[mēs]	‘table’	[dɛ̀ts]	‘back’	[bɛ̀l:]	‘meat’
[prés]	‘price’	[bɛ̀rd]	‘green’	[gɛ̀t]	‘tortilla’	[bɛ̀dz]	‘rabbit’
[tʃétʃ]	‘Moses’	[bɛ̀n]	‘twenty’	[gɛ̀l]	‘milpa’	[bɛ̀ʔ]	‘mushroom’
[tʃéʔk]	‘check’	[mɛ̀rg]	‘row of agave plants’	[bɛ̀l]	‘fish’	[gɛ̀l]	‘dawn’
[méskl]	‘mortar’	[mɛ̀l]	‘Ishmael’	[dɛ̀]	‘powder’	[bɛ̀tz]	‘child?’

Table 11: Words with [i]

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[líbr]	‘book’	[tʃɪntʃ]	‘bedbug’	[zɪt]	‘far away’	[rɪʔi]	‘jar’
[tíly]	‘Spanish’	[tʃɪkl]	‘gum’	[dʒɪ]	‘day’	[nɪʔi]	‘foot’
[díʒ]	‘language’	[tɪʔt]	‘Tito’	[nɪʒ]	‘delicious’	[nɪs]	‘water’
[lílm]	‘lime’	[zɪt]	‘stranger’	[ʃɪg]	‘jicara’	[zɪt]	‘far away’
[ʃíl]	‘wing’	[bɪʔtʃ]	‘small animal’	[nɪz]	‘dried corn cob’	[gɪts]	‘paper’

Table 12: Words with [o]

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[sóʔbr]	‘envelope’	[ʃóp]	‘six’	[tʃón:]	‘three’	[rɔʔ]	‘big’
[lór]	‘parrot’	[söp]	‘soap’	[dòb]	‘maguey’	[dôn]	‘if’
[ór]	‘hour’	[töp]	‘speed bump’	[kòb]	‘a type of alcoholic drink’	[bɔʔ]	‘carbon’
[kótʃ]	‘car’	[böt]	‘canister’	[lò]	‘face’	[zòb]	‘corn’
[bób]	‘drool’	[kök]	‘coconut’	[dòp]	‘short’	[dòp]	‘fart’

Table 13: Words with [u]

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[lúk]	‘Lucas’	[lũpy]	‘Lupe’	[rù]	‘cough’	[rû?]	‘mouth’
[lúdʒ]	‘tongue’	[mũsk]	‘music’	[sù]	‘earthquake’	[dû?]	‘rope’
[ʃúk]	‘upper arm’	[lũʒ]	‘edible rats’	[sùdy]	‘a cut’	[gùn]	‘bull’
[kúrs]	‘course’	[pũnt]	‘top’	[dùl:]	‘fault’	[lû]	‘you’
[lús]	‘Luis’	[grũp]	‘group’	[ʃùn]	‘eight’	[gû?]	‘sweet potato’

Table 14: Wordlist for Part II of Study

SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss	SADVZ	Gloss
[pá?p]	‘potato’	[dád]	‘father’	[táp]	‘four’	[nâ?tʃ]	‘rough’
[kár:]	‘car’	[bäd]	‘duck’	[bäd]	‘Tlacolula’	[râp]	‘have’
[pág]	‘pay’	[lã?s]	‘skinny’	[sæk]	‘cost, price’	[lâts]	‘field’
[fáb]	‘detergent’	[stämp]	‘stamp’	[bädʒ]	‘tepache’	[dʒâ?p]	‘girl’
[táp]	‘lid’	[kwã?tʃ]	‘twin’	[dâny]	‘mountain’	[nâd]	‘to stick’
[lát]	‘(tin) can’	[măn]	‘animal’	[râp]	‘care’	[bäd]	‘scabies’
[rán]	‘frog’	[dâm:]	‘owl’	[n:â]	‘chocolate’	[ʒan]	‘butt’
[bál]	‘stick’	[dʒăn]	‘saint’	[lâdy]	‘clothes’	[bâ?t]	‘skunk’
[bág]	‘cow’	[fãld]	‘skirt’	[nâ]	‘thick’	[gâts]	‘hidden’
[ká?m:]	‘bed’	[mãrkw]	‘mark’	[nân]	‘cold (n.)’	[dâts]	‘empty’

References

- Avelino, Heriberto. 2001. *Phonetic Correlates of Fortis-Lenis in Yalálag Zapotec Consonants* M.A. thesis. UCLA
- Blankenship, Barbara. 1997. *The Time Course of Breathiness and Laryngealization in Vowels*. Ph.D. dissertation. UCLA.
- Broadwell, George A. 1991 "Santa Ana del Valle Zapotec-English-Spanish Word List". MS.
- Epstein, Melissa. 2002. *Voice Quality and Prosody in English*. PhD Dissertation. UCLA.
- Gordon, Matthew and Peter Ladefoged. 2002. Phonation types: a cross-linguistic overview.
- Grimes, Barbara F., ed 1990. *Ethnologue*. Consulting editors: Richard S. Pittman and Joseph E. Grimes. Dallas, TX: Summer Institute of Linguistics, Inc. <http://www.sil.org/ethnologue>.
- Hanson, Helen. 1997. Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America* 101 (1), 446-481.
- Huffman, Marie K. 1987. Measures of phonation type in Hmong. *Journal of the Acoustical Society of America* 81 (2), 495-504.
- Jones, Ted E., and Lyle M. Knudson. 1977. "Guelavia Zapotec Phonemes." *Studies in OtoManguean Phonology*, ed., William R. Merrifield, pp. 163-80. Dallas: SIL/ University of Texas, Arlington
- Jaeger, Jeri. 1983. "The fortis/lenis question: evidence from Zapotec and Jawoñ". *Journal of Phonetics*, 11, 177-189.
- Klatt, Dennis H., and Laura C. 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.
- Ladefoged, Peter. 1983. The linguistic use of different phonation types. In Bless and Abbs, 1983: 351-360.
- Laver, John. 1980. *Phonetic Description of Voice Quality*. Cambridge: University of Cambridge Press.

- Lyman, Larry and Rosemary Lyman. 1977. "Choapan Zapotec Phonology", *Studies in Otomanguean phonology*. William R. Merrifield, editor. Summer Institute of Linguistics. University of Texas Arlington: 137-161. Summer Institute of Linguistics and the University of Texas at Arlington.
- Madiesson, I. and S. Hess. 1987. The effect of F0 on linguistic use of phonation types. *UCLA Working Papers in Phonetics*. 67, 112-118.
- Munro, Pamela and Felipe H. Lopez, with Olivia Mendez, Rodrigo Garcia, and Michael Galant. 1999. *Di'syonaary X:tee'n Diizh Sah Sann Lu'uc (San Lucas Quiavini Dictionary/ Diccionario Zapoteco de San Lucas Quiavini)*. Chicano Studies Research Center Publications, UCLA.
- Nellis, Donald G. and Barbara E. Hollenbach. 1980. "Fortis versus lenis in Cajonos Zapotec phonology" *International Journal of American Linguistics* 49.
- Ohala, John. 1973. The physiology of tone. *Southern California Occasional Papers in Linguistics*. 1,1-14.
- Silverman, Daniel D., Barbara Blankenship, Paul Kirk and Peter Ladefoged. 1995. "Phonetic structures of Jalapa Mazatec". *Anthropological Linguistics*, 37; 70-88.
- Silverman, Daniel. 1997. Laryngeal complexity in Otomanguean vowels, *Phonology* 14, 235-262.
- Stevens, Kenneth N. 1988. Modes of Vocal Fold Vibration Based on a Two-Section Model. In Fujimura, 1988; 357-371.
- Stevens, Kenneth and Helen Hanson. 1995. Classification of glottal vibration from acoustic measurements. In O. Fujimura and M. Hirano, eds., *Voice Quality Control*. Singular Publishing Group, Inc.: San Diego, CA. 147-170.
- Stubblefield, Morris, and Carol Miller de Stubblefield. 1991. *Diccionario Zapoteco de Mitla, Oaxaca*. Mexico, D.F: Instituto Linguistico de Verano.
- Swadesh, Morris. 1947. "The phonetic structure of Proto-Zapotec". *International Journal of American Linguistics*, 13, 220-230.