Title:
The acoustic consequences of phonation and tone interactions in Jalapa Mazatec

Author:

Garellek, Marc, University of California, Los Angeles
Keating, Patricia, University of California, Los Angeles

Publication Date:
09-20-2010

Series:
Working Papers in Phonetics

Publication Info:
Working Papers in Phonetics, Department of Linguistics, UCLA, UC Los Angeles

Permalink:
http://escholarship.org/uc/item/1gs6h5k7

Keywords:
phonation, tone, Jalapa Mazatec

Abstract:

San Felipe Jalapa de Díaz (Jalapa) Mazatec is unusual in possessing a three-way phonation contrast and three-way level tone contrast independent of phonation. This study investigates the acoustics of how phonation and tone interact in this language, and how such interactions are maintained across variables like speaker sex, vowel timecourse, and presence of aspiration in the onset. Using a large number of words from the recordings of Mazatec made by Paul Kirk and Peter Ladefoged in the 1980s and 1990s, the results of our acoustic and statistical analysis support the claim that spectral measures like H1-H2 and mid-range spectral measures like H1-A2 best distinguish each phonation type, though other measures like Cespstral Peak Prominence are important as well. This is true regardless of tone and speaker sex. The phonation contrasts are strongest in the first third of the vowel and then weaken towards the end. Although tones remain distinct from one another in terms of F0 throughout the vowel, for laryngealized phonation the tone contrast in F0 is partially lost in the initial third. This study shows that the complex orthogonal three-way phonation and tone contrasts do remain acoustically distinct according to the measures studied, despite partial neutralizations in any given measure.

UCLA Working Papers in Phonetics, No. 108, pp. 141-163

The acoustic consequences of phonation and tone interactions in Jalapa Mazatec

Marc Garellek and Patricia Keating

(marcgarellek@ucla.edu, keating@humnet.ucla.edu)

Abstract

San Felipe Jalapa de Díaz (Jalapa) Mazatec is unusual in possessing a three-way phonation contrast and three-way level tone contrast independent of phonation. This study investigates the acoustics of how phonation and tone interact in this language, and how such interactions are maintained across variables like speaker sex, vowel timecourse, and presence of aspiration in the onset. Using a large number of words from the recordings of Mazatec made by Paul Kirk and Peter Ladefoged in the 1980s and 1990s, the results of our acoustic and statistical analysis support the claim that spectral measures like H1-H2 and mid-range spectral measures like H1-A2 best distinguish each phonation type, though other measures like Cespstral Peak Prominence are important as well. This is true regardless of tone and speaker sex. The phonation contrasts are strongest in the first third of the vowel and then weaken towards the end. Although tones remain distinct from one another in terms of F0 throughout the vowel, for laryngealized phonation the tone contrast in F0 is partially lost in the initial third. This study shows that the complex orthogonal three-way phonation and tone contrasts do remain acoustically distinct according to the measures studied, despite partial neutralizations in any given measure.

1 Introduction

Mazatec is an Otomanguean language of the Popolocan branch. This study investigates the acoustics of the phonation contrasts in the San Felipe Jalapa de Díaz (henceforth Jalapa) dialect, which according to a 2005 census is spoken by approximately 24,200 people in Mexico, in North Oaxaca and Veracruz states (Ethnologue 2009). Jalapa Mazatec has a five-vowel system with length and nasal contrasts. In addition, there are three tone levels (low, middle, and high) and three phonation contrasts (breathy, modal, and laryngealized). The laryngealized phonation has in the past been referred to as ‘creaky’ (Kirk et al. 1993), but Blankenship (2002) preferred the term ‘laryngealized’, because laryngealized phonation is often used for phonation with stiffer vocal folds than modal voice but that does not involve actual creak (Ladefoged 1983; Gerfen & Baker 2005). In keeping with her work, we will also use this term. All three tones and phonation types are independent of one another and may occur on all five vowels. A thorough description of Jalapa Mazatec phonetics is available in Silverman et al. (1995).

In their survey of phonations in the world’s languages, Gordon & Ladefoged (2001) cite few languages with phonation contrasts on vowels (Gujarati, !Xóõ, Kedang, Hmong, Mpi, Bruu, Yi, Haoni, Jingpho, Wa), and only four with three contrasting categories (Jalapa Mazatec, San

Lucas Quiaviní Zapotec, Burmese, and Chong, though at least some dialects of Hmong belong here too). Languages with more than three contrastive phonations are of course very rare, but include Chong (DiCanio 2009), Bai, Bor Dinka (Edmondson & Esling 2006), Ju|’hoansi (Miller 2007), and !Xóõ (Traill 1985). Jalapa Mazatec is rare in contrasting three phonations and three tones independently. Most languages (and even other Mazatec dialects) with phonation contrasts distinguish only between two phonation types (DiCanio 2009). The independent tone and phonation contrasts in Jalapa Mazatec make the language particularly suited for investigating how phonation contrasts may vary by tone, speaker sex, and time. Previous studies of Mazatec have ignored or controlled tone contrasts to focus on the phonation contrasts.

Like previous studies (Silverman 1997; Blankenship 1997), we will consider timing and sex effects on phonation in Mazatec. The present study is thus novel in trying to account for influences of sex, tone, and time on phonation contrasts. We find notable differences in how contrasts are made across these three variables, lending further support to the notion that phonetic cues to phonation are both numerous and context-varying.

1.1 Measures of phonation

Traditionally, phonation contrasts have been distinguished using acoustic measures, though more recently there have been studies of physiological aspects of phonation production. One of the most popular models of phonation contrasts is the continuum of glottal stricture (Ladefoged 1971; Gordon & Ladefoged 2001). This model only refers to the average aperture between the vocal folds in accounting for the major differences across voice qualities. Modal voice is characterized by an average opening that allows complete closure during glottal periods (e.g. Titze 1995); breathy voice is characterized by a greater average opening, typically with only incomplete closure of the vocal folds during glottal periods; creaky or laryngealized voice is characterized by a smaller average opening, typically with a very small maximum opening during glottal periods. The major reasons for the popularity of this model are first, its simplicity; second, that breathy, modal, and creaky phonation types can usually be ordered along the various acoustic parameters of voice (an argument made explicitly by Blankenship 2002); and third, that the acoustic measure that best serves to contrast phonation cross-linguistically, H1-H2, has been shown to correlate with Open Quotient (OQ), or the proportion of a glottal period during which there is no contact between the vocal folds (Holmberg et al. 1995), or alternatively with Contact Quotient derived from electroglottography (DiCanio 2009, Esposito submitted).

However, clearly the activity of the vocal folds can vary in more ways than represented by glottal stricture, e.g. Laver 1980, Hanson et al. 2001, Baken & Orlikoff 2000. And even more strikingly, direct observation of the laryngopharynx has shown that languages may use articulators other than the vocal folds to distinguish phonation types. For example, Edmondson & Esling (2006) claim that six different ‘valves’ comprising different articulators are used in the production of voice quality: glottal vocal fold adduction, ventricular incursion, upward and forward sphincteric compression of the arytenoids, epiglotto-pharyngeal constriction, larynx raising, and pharynx narrowing. To the extent that these (or other) articulations underlie phonation contrasts, the uni-dimensional glottal stricture model is insufficient. However, this plethora of physiological dimensions of voice quality variation makes it all the more intriguing that the standard acoustic measures tend to define continua of phonation contrasts.

Because in this study we collected no articulatory data, our analysis of the phonation types in Mazatec can only be based on the acoustic measures of the recorded sound files. The most widely used acoustic measure of phonation is H1-H2, i.e. the difference between the amplitudes of the first harmonic (the fundamental) and the second harmonic in the spectrum. H1-H2 has been shown to correlate with OQ (Holmberg et al. 1995) but also with the skew of the glottal pulse (Henrich et al. 2001), and thus the relation between H1-H2 and OQ can be weak (Kreiman et al. 2008). Despite the continuing debate as to its articulatory correlates, H1-H2 has been found to distinguish among contrastive phonations in many studies. For example, in a cross-linguistic sample of breathy versus modal phonation, Esposito (2010a) found that H1-H2 distinguished these phonations in eight out of the 10 languages or dialects. Moreover, Hanson (1997) showed that H1-H2 is not well-correlated with other acoustic measures in English, and Kreiman et al. (2007) found that H1-H2 accounted for the most variance in English voices out of 19 different spectral measures.

Other acoustic measures are thought to reflect other aspects of phonation. The strength of higher frequencies in the spectrum is thought to be related to the closing velocity of the vocal folds, to the presence of a posterior glottal opening, and to the simultaneity of ligamental closure (Stevens 1977; Hanson et al. 2001), among other possible influences. Higher frequency energy is usually measured as the amplitude of H1 relative to the amplitudes of F1 (A1), F2 (A2), and F3 (A3), as H1-A1, H1-A2, and H1-A3. These formant amplitude measures also reflect the bandwidths of the corresponding formants, and Hanson et al. (2001) interpret H1-A1 in particular as reflecting the effect of a posterior glottal opening.

Esposito (2006, 2010a) compared breathy and modal phonations in 10 languages/dialects and found that H1-A3 and H1-H2 were both fairly good at distinguishing the phonations within languages, while Blankenship (2002) found that H1-A2 better distinguished breathy from modal phonation in Chong than H1-H2 (and similarly DiCanio (2009) for H1-A3).

Moreover, breathy voice has been quantified by the presence of noise. Cepstral Peak Prominence (CPP) is thought to reflect the harmonics-to-noise ratio (Hillenbrand et al. 1994). A greater cepstral peak indicates stronger harmonics above the floor of the spectrum. This in turn can result from greater periodicity in the speech signal. CPP has been used in studies on phonation to distinguish breathy from non-breathy voice qualities, for both production and perception (Blankenship 2002, Esposito 2006, 2010a). Esposito (2006, 2010a) found that CPP was the best of the 8 measures she considered at distinguishing modal from breathy phonations. Other recent studies that have applied harmonic and/or noise measures to phonation contrasts include Andruski & Ratliff (2000) and Andruski (2006) on Mong; Blankenship (2002) on Mazatec, Chong, and Mpi; Wayland & Jongman (2003) on Khmer; Avelino (2006) on Yalálag Zapotec; and Miller (2007) on Ju|’hoansi.

Specifically with respect to Mazatec, Blankenship (2002) found that all three measures she tested, H1-H2, H1-A2, and CPP, were equally effective in distinguishing breathy from modal phonations, while CPP was less effective for laryngealized vs. modal. Esposito (2010a), characterizing the stimuli she used in a cross-language perception experiment, found that four measures, CPP, H1*-H2*, H1*-A1*, and H1*-A2*, each distinguished Mazatec breathy and modal phonations. Furthermore, however, linear discriminant analysis showed that H1*-A2* accounted for 53% of the variance in the Mazatec items, and thus was the most important measure of the contrast; H1*-A1* accounted for a further 20% and H1*-H2* another 14%.

1.2 Previous work on sex, time, and tone effects on phonation

It is well-known that on average, women tend to have breathier voices than men (Klatt & Klatt, 1990; Hanson & Chuang, 1999). Beyond such overall differences, differences in the acoustics of men and women in contrasting phonation types have been found in the work by Esposito on

Santa Ana del Valle Zapotec (Esposito 2003, 2005, 2010b). She found that in this language, the three phonations (breathy, modal, and creaky) were distinguished by H1-H2 for women and H1-A3 for men. These differences were further bolstered by electroglottographic data showing the same pattern with articulatory correlates of H1-H2 and H1-A3, namely contact quotient and a measure of closing/opening symmetry, respectively. While her study used data from only five speakers (three men and two women), her findings suggest that phonation contrasts may be produced differently by men and women. In contrast, it appears from the figures in Blankenship (1997: figures 70-73) that the women made larger distinctions among the phonations on all three measures (CPP, H1-H2, H1-A2) than the men did, though perhaps largest on CPP. As Blankenship reports, women produced breathier breathy phonation than men did, but this appears to have been part of a larger pattern of enhanced contrasts in women’s speech.

The time course of phonation has been shown to differ across phonations and languages as well. (See review in section 4 of Gordon & Ladefoged (2001).) Phonation contrasts have been found to be most pronounced at the start of a vowel in Mazatec (Blankenship 1997). Silverman (1997) hypothesized that the phasing of breathiness towards the beginning of the vowel in Mazatec was a means of enhancing the tone during the latter portions. In Mazatec, it has also been found that phrase-final vowels tend to end breathy, regardless of their phonation, and this makes all the phonations less distinct at the ends of phrase-final vowels (Blankenship 2002). Thus, we expect our results for Mazatec to be similar to those of Blankenship (2002) for the same language and speakers, though they do not necessarily indicate a typological tendency toward phrase-final breathiness or maximal phonation contrast vowel-initially. For example, while Edmondson (1997) showed that Chong breathy phonation is stronger (in terms of glottal airflow) at the beginning of the vowel, DiCanio (2009) found that in Takhian Thong Chong, the breathy-tense and tense registers have much greater vocal fold contact at the ends of vowels than at the beginnings, and Esposito (2003) found that Zapotec non-modal phonations are strongest at the ends of vowels.

There are several ways in which tone and phonation could interact, and each aspect has its own literature. First, phonation categories can differ in F0. Generally, non-modal phonation is associated with pitch lowering effects (Gordon & Ladefoged 2001), though laryngealized phonation can be associated with higher pitch, presumably due to glottal tension. This is especially well-documented with respect to the tonogenetic effects of consonants on adjacent vowels (Hombert et al. 1979, Kingston 2005). Second, and conversely, different F0s can differ in their voice quality. Some studies (Holmberg et al. 1989; Epstein 2002) have not found a strong correlation between pitch and glottal parameters or LF measures (Fant et al. 1985), but others (Iseli et al. 2006, Keating and Shue 2009) found that (corrected) H1-H2 increases with increasing F0 when F0 is below 175 Hz. That is, men with higher-pitched voices also had breathier voices. We will not address this possibility in the present study. However, third, and relatedly, tone categories can differ in voice quality. In languages with tonal contrasts, often certain tones are accompanied by non-modal phonation, as in the Mandarin dipping Tone 3, which has audible creak (Davison 1991, Belotel-Grenié & Grenié 2004), and similarly in Cantonese (Lam and Yu 2010). Finally, and perhaps relatedly, phonation categories can be constrained to occur only with certain tone categories. For example, in Southern Yi (Kuang 2010), the phonation contrast does not occur with high tone; in SADV Zapotec (Esposito 2010b), nonmodal phonations occur only with falling tone; only modal phonation occurs with high and rising tones. And, when the Zapotec falling tone is spoken at a higher pitch, as when under focus, then the breathy versus laryngealized phonation contrast is nearly neutralized to modal-like. In cases like this, it is unclear whether phonation accompanies tone or vice versa. This last kind of interaction does not arise in Mazatec, at least not strongly, since in Mazatec tone and phonation are orthogonal contrasts (though there may well be differences in the lexical frequency of each tone-phonation combination). However, it is possible that phonation contrasts are more vs. less robust when combined with the various tones of the language; in particular, the Mazatec contrast might be more difficult to maintain with a high tone.

2 Language materials

2.1 Recordings

The sample words come from two field recordings from San Felipe Jalapa Diaz, Oaxaca. The first recording was made by Paul Kirk in December 1982. Words without a carrier sentence were spoken by four male speakers. The second recording was made by Paul Kirk and Peter Ladefoged in April 1993. Using a different wordlist, words without a carrier sentence were spoken by six male speakers and six female speakers. Two of the male speakers participated in both recordings. Thus, 14 speakers in total were included in this study. Most of the males were bilingual in Mazatec and Spanish, while the females were mostly monolingual (Blankenship 2002). Both recordings, originally analog, were digitized at a sampling rate of 44.1 KHz, 16-bit sound depth, and are available online at the UCLA Phonetics Archive website. Blankenship (2002) used sample words uttered by the 12 speakers from the second recording. The four speakers from the 1984 recording are the speakers studied in Kirk et al. 1984, whereas the twelve speakers from the 1993 recording were used in subsequent studies of Jalapa Mazatec (e.g. Silverman et al. 1995, Blankenship 2002, Esposito 2010a).

2.2 Sample words

In keeping with Blankenship (2002), the sample words chosen for this study all had non-nasal vowels. But unlike the previous study, the target words could have any of three tones and any of the three phonations. Most target vowels were syllable-final in keeping with Blankenship (2002), except for the two words with breathy vowels with a high tone, which were only found on non-word-final syllables. Only mid and low vowels [a], [æ], and [o] were chosen, due to their greater proportion in the wordlist and the fact that a high F1 is unlikely to influence H2. Table 1 gives a sample of the Mazatec words chosen (and the rest are listed in the Appendix).

Table 1 Examples of larger set of Mazatec words used in this study. Tone 1 is low, 2 is mid, and 3 is high.

	Laryngealized	Modal	Breathy
Low tone	βa̰^{1} ‘thus’	ja^1 ‘kind of ant’	ⁿdja̤^{1} ‘animal horn’
Mid tone	βa̰^{2} ‘carries’	hæ^2 ‘finished’	ⁿda̤^{2} ‘good’
High tone	βæ̰^{3} ‘hits’	ha^3 ‘men’	ndʒa̤^{3}ʃu^3 ‘chocolate drink’

Tokens with audible background noise were discarded. Because two different recordings were used, not all tokens are the same for all 14 speakers. A total of 80 words were sampled across all speakers. Of these, roughly twenty percent were breathy, forty percent were creaky, and forty percent were modal. In a few cases, multiple tokens of a word were analyzed, for a total of 424 tokens of the 80 words. This is in contrast to Blankenship (2002), who used only 9 words from 12 speakers, for a total of 108 tokens, and Esposito (2010a), who used 16 words (8 breathy, 8 modal) from each of 3 speakers, for a total of 48. All the phonation-tone permutations had speakers of both sexes and from both recordings, except the breathy high-toned tokens, which were uttered only by men (these words were only present in the 1982 recording).

Except for Section 3.1 and 4.4, where we discuss the specific effects of aspirated onsets on a following vowel's phonation, all words with aspirated stops preceding the target vowel were excluded in the analyses. This was done to reduce the effect of neighboring sounds on a vowel's phonation.

2.3 Obtaining acoustic measurements

The vowel portion of each word was labeled in Praat (Boersma & Weenink 2008). The vowel onset was set at the first glottal pulse following the onset, and the vowel end was set at the last glottal pulse. The selected portion was labeled for vowel, phonation, and tone using a Praat labeling script. VoiceSauce (Shue et al. 2009), a MATLAB-implemented application, was then run on the labeled audio files, providing the following measurements over time: the first, second and fourth harmonics (H1, H2, H4), the difference between the first and second harmonics (H1-H2) and the second and fourth harmonics (H2-H4), the difference between the first harmonic and the first, second and third formants (H1-A1, H1-A2, H1-A3), energy, Cepstral Peak Prominence (CPP), F0, as well as the first four formants and their bandwidths. Corrected versions of the harmonics and formant amplitudes were obtained automatically in VoiceSauce, which uses the correction algorithm of Iseli et al. (2007). Formant values were calculated using the Snack Sound Toolkit (Sjölander 2004), while F0 was calculated using the STRAIGHT algorithm (Kawahara et al. 1998). For each input .wav file, VoiceSauce produced a MATLAB file with values every millisecond for all the measures mentioned above, over the vowel portion delimited by the Praat textgrid. VoiceSauce then averaged the results by thirds of the vowels' duration and output these values in a text file.

3 Results

3.1. Significant measures of phonation

Using the results of the acoustic analysis, a linear discriminant analysis (LDA) was conducted to determine which measures are most important for distinguishing phonation types. The acoustic measures included in the discriminant analysis were the following: H1*-H2*, H2*-H4*, H1*-A1*, H1*-A2*, H1*-A3*, CPP, Energy, F0, and the first four formants and their bandwidths. The values for these measures was taken over the first third of the vowel's duration, because it has been shown (and will be corroborated below) that the phonation contrast in Mazatec is manifested early in the vowel (Silverman 1997, Blankenship 2002). The measures were input in a stepwise manner. In total, 424 tokens were included in the analysis (including words with aspirated onsets), consisting of roughly 40% breathy, 40% modal, and 20% laryngealized vowels.

The results of the LDA are shown in Table 2. Two discriminant functions were included because the phonation contrast has three possibilities. The coefficients indicate the relative importance of the measures in predicting phonation for the function. An asterisk indicates to which discriminant function a particular acoustic measure was assigned. Seven measures were significant in the analysis: H1*-H2*, H1*-A1*, H1*-A2*, CPP, F0, F1, and B4.

Table 2. Statistical results of the linear discriminant analysis. The largest absolute correlation between a variable and a function is indicated with an asterisk.

Acoustic measure	Correlation with discriminant functions		Wilks' Lambda	F value	Significance
	Function 1	Function 2			
H1*-H2*	0.695*	-0.070	0.760	25.294	<.001
H1*-A1*	0.776*	-0.045	0.816	58.399	<.001
H1*-A2*	0.715*	0.162	0.698	20.196	<.001
CPP	-0.056	0.599*	0.786	32.919	<.001
F0	-0.224	0.494*	0.722	22.770	<.001
F1	0.140 *	0.066	0.685	17.755	<.001
B4	0.176	-0.301*	0.673	15.967	<.001

The results of the LDA indicate that the harmonic measures (H1*-H2*, H2*-H4*, H1*-A1*, H1*-A2*, and H1*-A3*) all correlate with Function 1, whereas the other measures like F0, CPP, Energy and the formant frequencies and bandwidths correlate with Function 2. The most important predictors of Function 1 are (in order) H1*-A1*, H1*-A2*, and H1*-H2*. The most important predictors of Function 2 are (in order) CPP, F0, and B4. The following analysis will therefore focus specifically on these measures.

3.2. Timecourse during vowels

Blankenship (2002) found that the effects of phonation type on a variety of acoustic measures were strongest in about the first one-third to one-half of vowels, are weaker later in the vowel, and are generally lost by the ends of vowels. Correspondingly, when our vowels are divided into one-third intervals and the phonation types are compared by linear mixed- effects (LME) models on the measures that were significant in the LDA, with the acoustic measure as a fixed effect and random intercepts for speaker and word, the phonations are most often distinct in the first two thirds, and least often distinct in the last third. Although the phonation differences by third for F0, F1, and B4 look in Figure 1 as though they are trending towards significance, the results of the LME models indicate that these differences are not significant at $p = 0.05$, even in the first third. Therefore, although these measures are correlated to some degree with the discriminant functions found in section 3.1, the phonation contrasts are not distinguished from each other according to F0, F1, and B4. For this reason, we will focus the subsequent discussion on the four measures which do show significant differences across phonations, H1*-H2*, H1*-A1*, H1*-A2*, and CPP. Even though differences are weaker in the final third, the distinction between modal and laryngealized holds throughout the vowel (except for CPP in the final third), whereas breathy and modal are neutralized in the final third on all measures, and in the middle third for H1*-A1*.

Differences between breathy and laryngealized phonations are significant throughout the entire vowel duration for the spectral measures, whereas for CPP no significant differences were found. During the first third, breathy vowels had a lower CPP mean than laryngealized vowels, but this difference was only moderately significant at $p = 0.07$.

Figure 1 shows the four measures that were significant in the LDA by vowel-thirds, and Table 3 gives the significance of each comparison between modal and non-modal phonations, which were calculated using linear mixed-effects models with phonation as a fixed effect and speaker and word as random effects. The mixed-effects modeling were run in R using the *lmer* function from the *languageR* package, and the p-values were obtained using the pvals.fnc function from the same package, with 10,000 simulations. This function estimates the p-values of the model's coefficients from the posterior distributions (Baayen, Davidson, & Bates 2008).

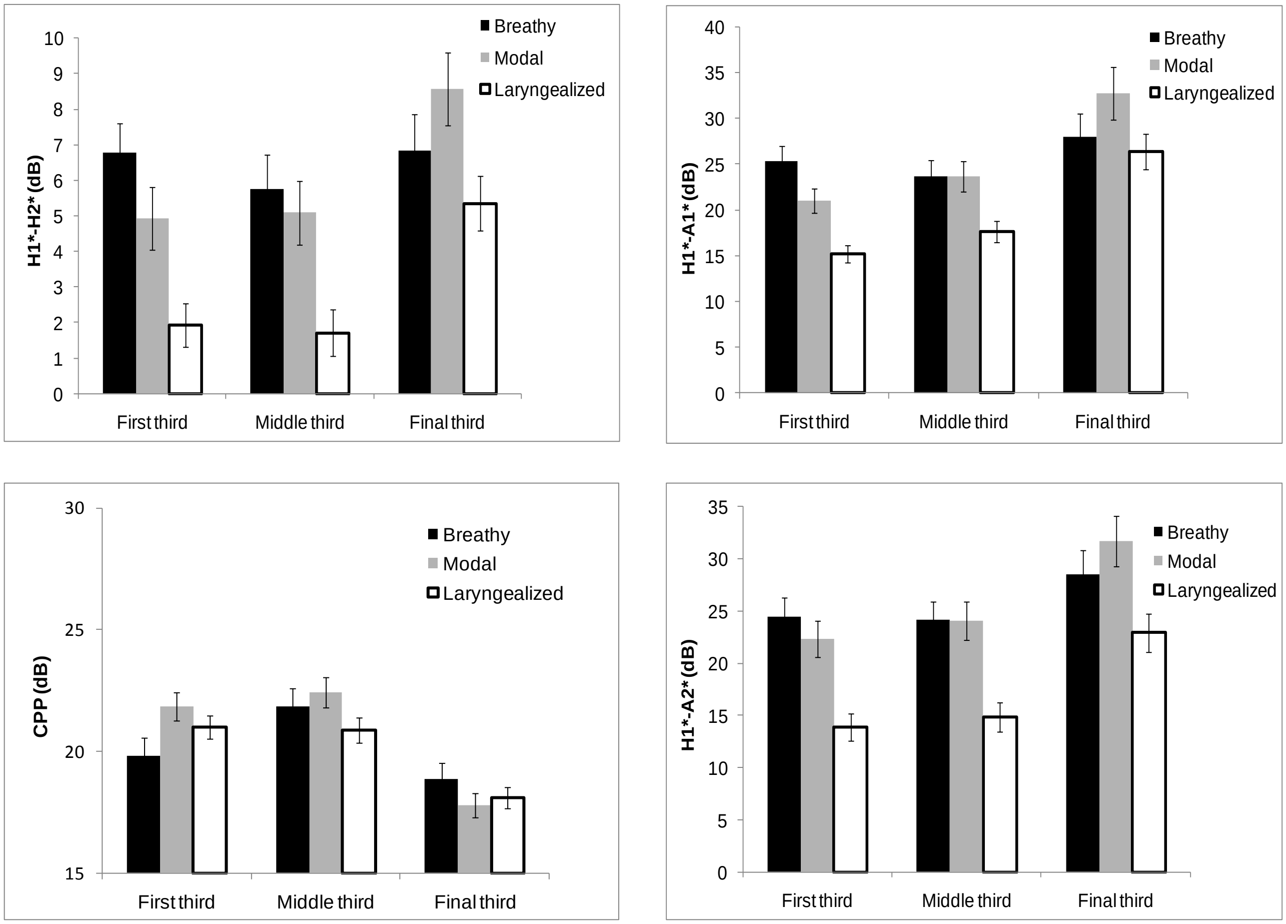
Table 3. Pairwise modal vs. non-modal comparisons for each acoustic measure at each third. Asterisks indicate statistical significance at $p < 0.05$.

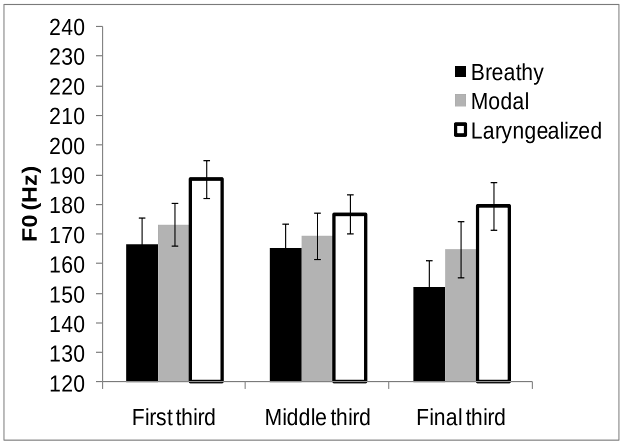
Acoustic measure	Contrast	First third	Middle third	Final third
H1*-H2*	Breathy vs. Modal	<0.0001*	0.0007*	0.9207
	Laryngealized vs. Modal	<0.0001*	<0.0001*	<0.0001*
	Breathy vs. Laryngealized	<0.0001*	<0.0001*	<0.0001*
H1*-A1*	Breathy vs. Modal	<0.0001*	0.0693	0.461
	Laryngealized vs. Modal	0.0002*	0.0001*	0.0124*
	Breathy vs. Laryngealized	<0.0001*	<0.0001*	0.1112
H1*-A2*	Breathy vs. Modal	<0.0001*	0.0099*	0.7413
	Laryngealized vs. Modal	0.0001*	<0.0001*	0.0001*
	Breathy vs. Laryngealized	<0.0001*	<0.0001*	0.0001*
CPP	Breathy vs. Modal	0.0001*	0.0202*	0.2329
	Laryngealized vs. Modal	0.011*	0.0008*	0.0825
	Breathy vs. Laryngealized	0.0783	0.4823	0.7108
F0	Breathy vs. Modal	0.3848	0.8807	0.2397
	Laryngealized vs. Modal	0.8561	0.1758	0.6367
	Breathy vs. Laryngealized	0.4715	0.2732	0.4478
F1	Breathy vs. Modal	0.4778	0.5665	0.8862
	Laryngealized vs. Modal	0.4835	0.7363	0.3862
	Breathy vs. Laryngealized	0.1659	0.3642	0.3343
B4	Breathy vs. Modal	0.4821	0.7256	0.1127
	Laryngealized vs. Modal	0.1967	0.0016*	0.0029*
	Breathy vs. Laryngealized	0.0489*	0.0081*	0.2579

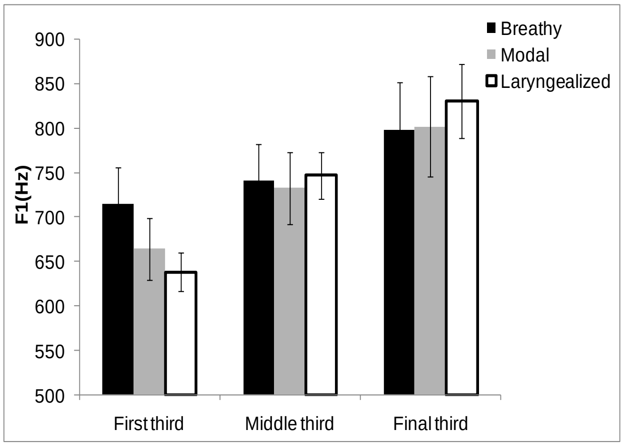
Because the phonation contrast is strongest in the first third of vowels, the analyses that follow are limited to that time interval. However, it should be born in mind that this does not mean that contrasts are made only in the first third, but simply that they are clearest there. Linear mixed-effects models were run to determine the main effects and interactions of various predictors like phonation, sex, tone, and aspiration on the four acoustic measures. The significance of main effects and interactions was established by model comparison, where the full linear mixed-effects model was compared to one lacking either a main effect or an interaction.

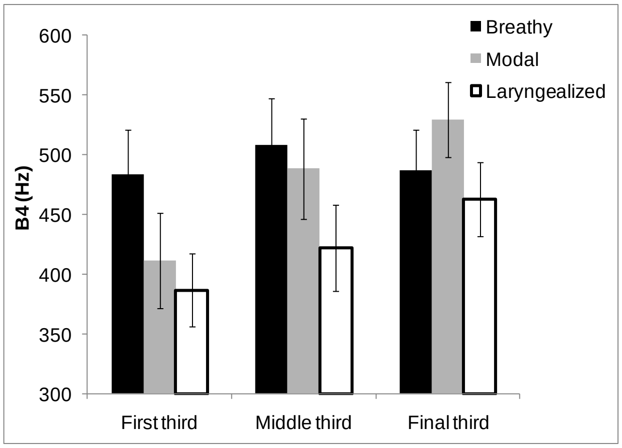
For the four phonation measures (the LDA measures, excluding F0, F1, and B4), a significant main effect of phonation was found (p <0.001). From Table 3 we see that both non-modal phonations differ from modal on the four measures reported there (H1-H2/A1/A2, CPP). This finding extends Blankenship (2002), in which breathy vs. modal differed on all three of the parameters she tested (H1-A2/H1, CPP), and laryngealized vs. modal differed more on the harmonic measures and less on CPP. Breathy phonation has the lowest CPP values, as found by Blankenship (2002), but in this study CPP for breathy phonation is only moderately lower than for laryngealized phonation. As mentioned above, breathy and laryngealized phonations are usually well differentiated, even in the final third, for the spectral measures, but not for CPP.

Figure 1. Acoustic measures by vowel thirds (with 95% confidence intervals), showing that differences between phonations are greatest in the first third. Figure continued next page.









3.4. Sex differences

Next we consider whether the two sexes differed significantly in how they used the four measures to distinguish phonations. Figure 2 shows men vs. women for each phonation, separately for each measure, and Table 4 gives the results of the tests of significance from the linear mixed-effects models. For CPP and H1*-A2*, main effects of sex were found ($p < 0.00357$ for the former, and $p = 0.01125$ for the latter). The direction of the differences for CPP and H1*-A2* would seem to indicate that men are breathier than women. A similar difference is found for H1*-A1*, although this main effect was not significant. However, for just breathy vs. modal, the difference in H1*-A1* is significant.

Interestingly, the opposite trend is found for H1*-H2*, where men seem to be less breathy than women, although only for modal phonation does this trend approach significance, at $p < 0.09$. A similar difference was found by Blankenship (1997).

Figure 2. Acoustic measures for women vs. men compared within phonations during the initial third. Error bars show the 95% confidence interval around the mean.

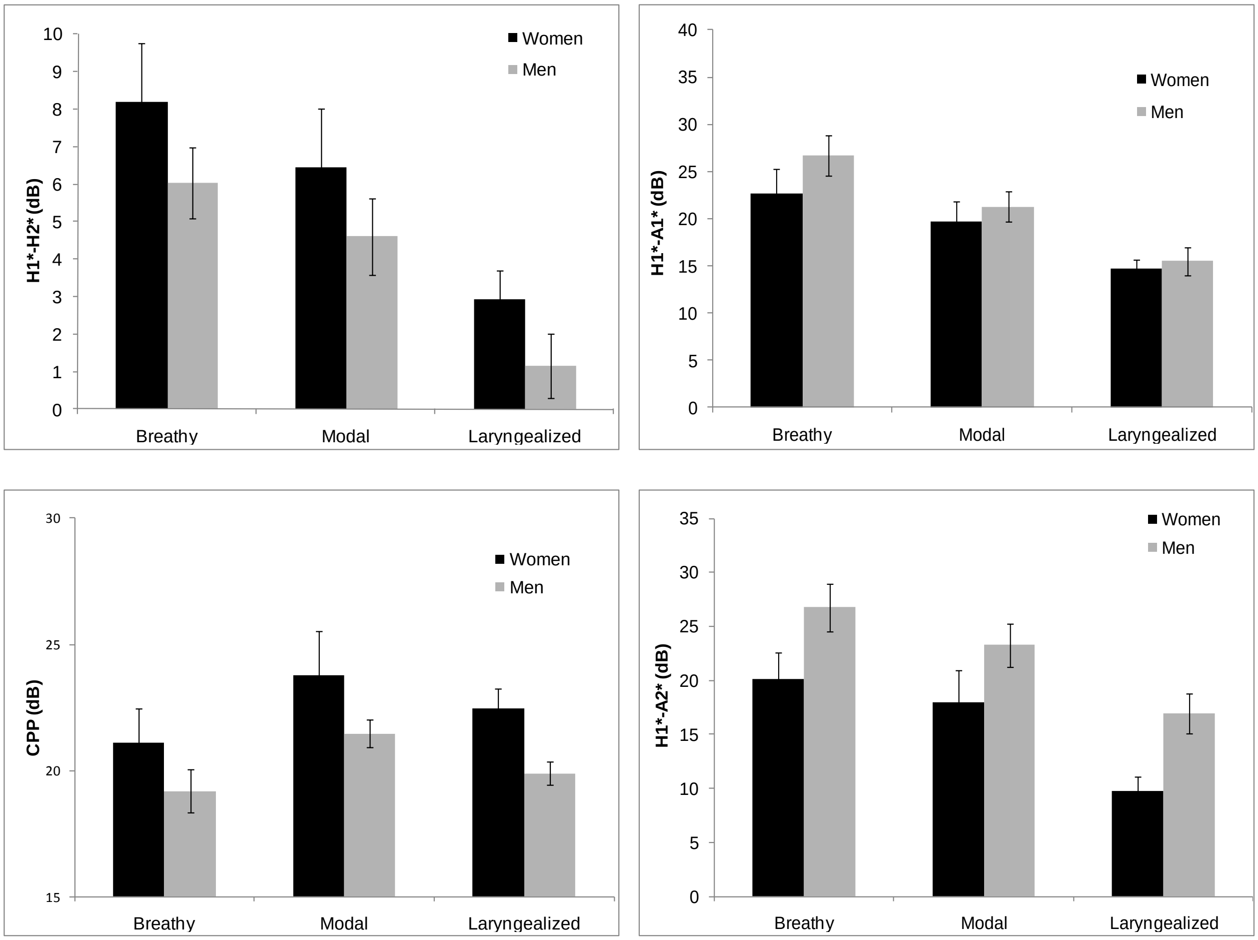


Table 4. Pairwise modal vs. non-modal comparisons for each acoustic measure by sex and phonation during the initial third. Asterisks indicate $p < 0.05$.

Acoustic measure	Contrast	Breathy	Modal	Laryngealized
H1*-H2*	Women vs. men	0.4686	0.087	0.1183
H1*-A1*	Women vs. men	0.0289*	0.7824	0.3588
H1*-A2*	Women vs. men	0.0031*	0.0794	0.0025*
CPP	Women vs. men	0.0119*	0.1186	0.0093*

Our main interest here, however, is whether men and women differ in how they distinguish the three phonation types, especially with respect to the acoustic measures that best distinguish the phonation types overall as described above. Differences in how the sexes distinguish the phonation types would result in significant sex by phonation interactions, but no such interactions were found. Thus, if men are breathier than women on a given measure, they are consistently breathier across all phonations. This result perhaps differs from Blankenship (1997), whose figures 70-73 suggest that the women's contrasts were generally larger than the men's on all three of her measures, with the exception of breathy vs. modal on H1-H2. However, she presents no statistical analyses on this point.

3.5. Phonation by tone interactions

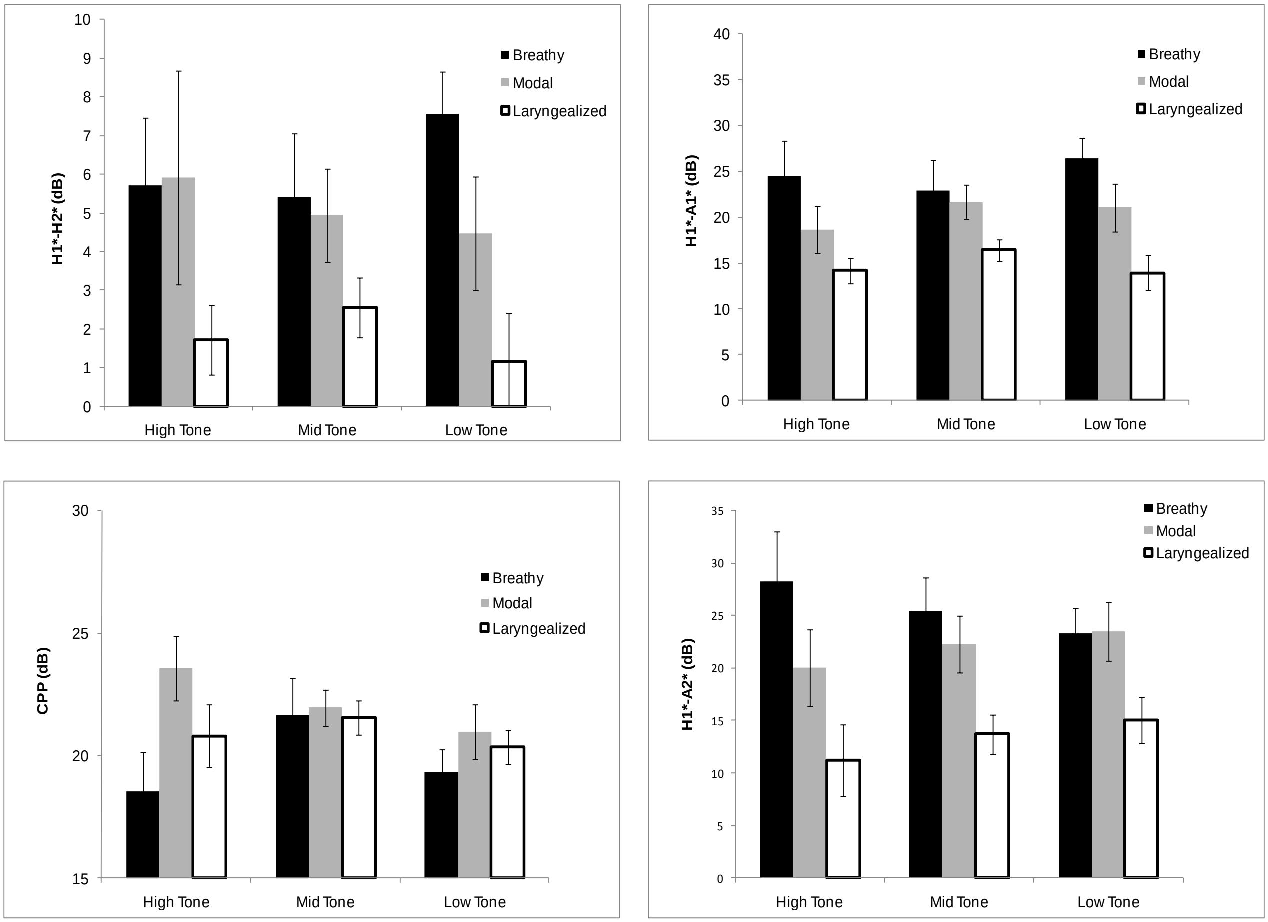
Jalapa Mazatec is unusual in having independent tones and phonations, and all nine combinations of them. Nonetheless, at least in part because acoustic measures of voice quality can vary with F0, we might expect that the tones, in addition to the phonation categories, will differ along one or more of our voice quality measures. Conversely, we might expect that the phonation contrasts will be more robust on some tones than on others, perhaps least robust on high tones. Finally, we might expect that one or both of the non-modal phonation types will differ from modal with respect to their F0 values, within the limits imposed by their lexical tones.

First, do the tone categories differ in voice quality? Most notably, is there a main effect of tone on any voice measures? For CPP, a main effect of tone was found ($p = 0.002$), with the tonal values in the order Mid > High > Low. Such a non-linear relation of CPP to tonal F0 means that this difference is unlikely to be due to any simple correlation with F0. Instead, it indicates that Mid tones, presumably spoken on the most comfortable pitches, have the most harmonic spectra.

None of the other measures showed a main effect of tone; instead, more complex interaction effects obtain, as can be seen in Figure 3. A significant phonation by tone interaction was found for H1*-A2* ($p = 0.02$). H1*-A2* decreases from High to Low tones within the Breathy category, but increases within Modal and Laryngealized. Again, such effects cannot be due to simple correlations of voice measures with F0, which, as will be presented below, did differ among the tones in the expected way.

Second, are phonation contrasts more robust on some tones than on others? Because the tone and phonation interactions go in different directions, sometimes there is contrast enhancement, other times contrast reduction. Thus the phonation and tone interaction for H1*-A2* appears to be a result of breathy vs. modal neutralization on low tones versus contrast enhancement on high tones. The contrasts also appear most robust on high tones when measured by CPP. The individual comparisons, given in Table 5, show that H1*-H2* and H1*-A1* distinguish all three phonations only on low tones, while CPP works best on high and low tones. H1*-A2* can distinguish all three phonations only on mid tones. Thus, in terms of how well each measure (taken separately) distinguishes the phonations within each tone category, it seems that the evidence is mixed and no single tone best supports the phonation contrasts.

Figure 3. Acoustic measures for phonations compared within tones during the initial third. Error bars show 95% confidence intervals around the mean.



Alternatively, we can consider the robustness of phonation contrasts in terms of how many of the individual measures support a contrast, and here we get a different picture. A closer look at the pairwise comparisons in Table 5 reveals that each pair of phonations is distinguished by at least two of the measures, with the exception of breathy vs. modal on mid tones, where only H1*-A2* makes a significant difference. Phonations are overall distinguished by the most measures on low tones (3 out of 4 per contrast); the breathy vs. modal contrast is especially less distinct on high and mid tones. Thus, in terms of how well the set of measures (taken together) distinguishes the phonations within each tone category, it seems that the phonation contrast is more robust with low tones.

Table 5. Pairwise modal vs. non-modal comparisons for each acoustic measure by tone within the initial third. An asterisk indicates statistical significance at $p < 0.05$.

Acoustic measure	Contrast	High tone	Mid tone	Low tone
H1*-H2*	Breathy vs. Modal	0.1718	0.1032	0.0011*
	Laryngealized vs. Modal	0.0047*	0.0006*	0.0023*
	Breathy vs. Laryngealized	< 0.0001*	< 0.0001*	< 0.0001*
H1*-A1*	Breathy vs. Modal	0.1001	0.0783	0.0002*
	Laryngealized vs. Modal	0.1950	0.0249*	0.0144*
	Breathy vs. Laryngealized	0.0043*	0.0006*	< 0.0001*
H1*-A2*	Breathy vs. Modal	0.0001*	< 0.0001*	0.2882
	Laryngealized vs. Modal	0.0953	0.0032*	0.0006*
	Breathy vs. Laryngealized	< 0.0001*	< 0.0001*	< 0.0001*
CPP	Breathy vs. Modal	0.0005*	0.2603	0.0074*
	Laryngealized vs. Modal	0.0131*	0.0982	0.3664
	Breathy vs. Laryngealized	0.1761	0.8725	0.0491*

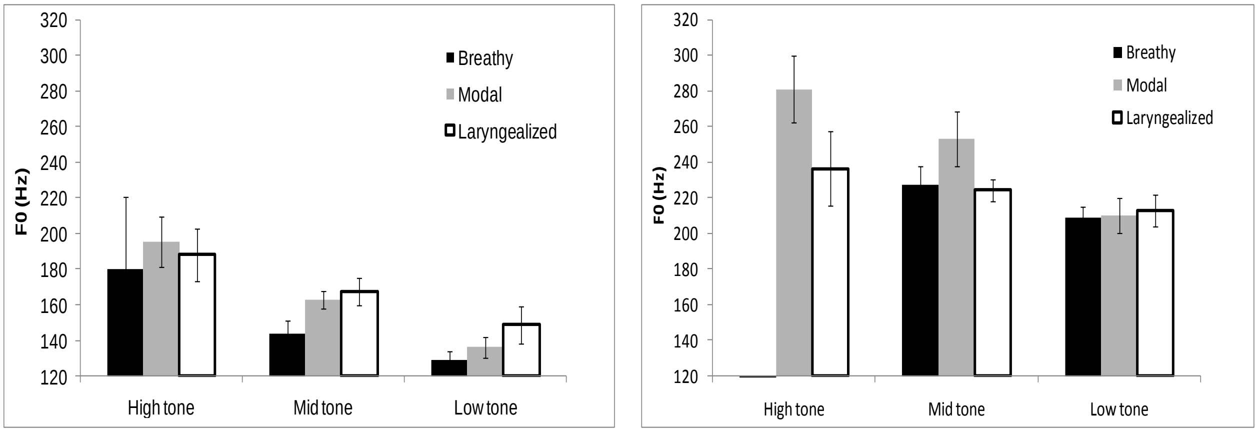
Third, do non-modal phonations differ in F0 from modal phonation? That is, can F0 alone distinguish phonations? Figure 1 appears to show such differences, and F0 was a significant measure in the initial LDA, but no main effect of phonation on F0 was found in the subsequent LME analysis. Pairwise comparisons reveal no pitch differences between modal and non-modal phonation for any of the tones. Figure 4 shows that the within-phonation variability is fairly large.

In contrast, a main effect of tone on F0 was found in the expected direction, with high tones having the highest F0, followed by mid tones, and then by low tones. Within each phonation category, this main effect holds true, as shown in Figure 4 (separated by sex). The pairwise tone comparisons for both sexes combined are given in Table 6, and show that the only non-significant difference is between mid and low tones with laryngealization, where $p < 0.10$. Recall that these results are for the first third of the vowel’s duration. During the middle and final thirds, the difference between laryngealized mid and low tones was found to be statistically significant ($p < 0.0001$ during both the middle and final thirds). This suggests that tone contrasts are strongest after the initial third, at least for laryngealized vowels.

Table 6. Pairwise tonal comparisons for F0 by phonation. An asterisk indicates statistical significance at $p < 0.05$.

Acoustic measure	Contrast	Breathy	Modal	Laryngealized
F0	High vs. Mid	0.0002*	0.0001*	0.0143*
	Mid vs. Low	0.0055*	0.0014*	0.0957
	High vs. Low	< 0.0001*	< 0.0001*	0.0004*

Figure 4. F0 for phonations compared within tones during the initial third (left for men; right for women). Error bars show 95% confidence intervals around the mean



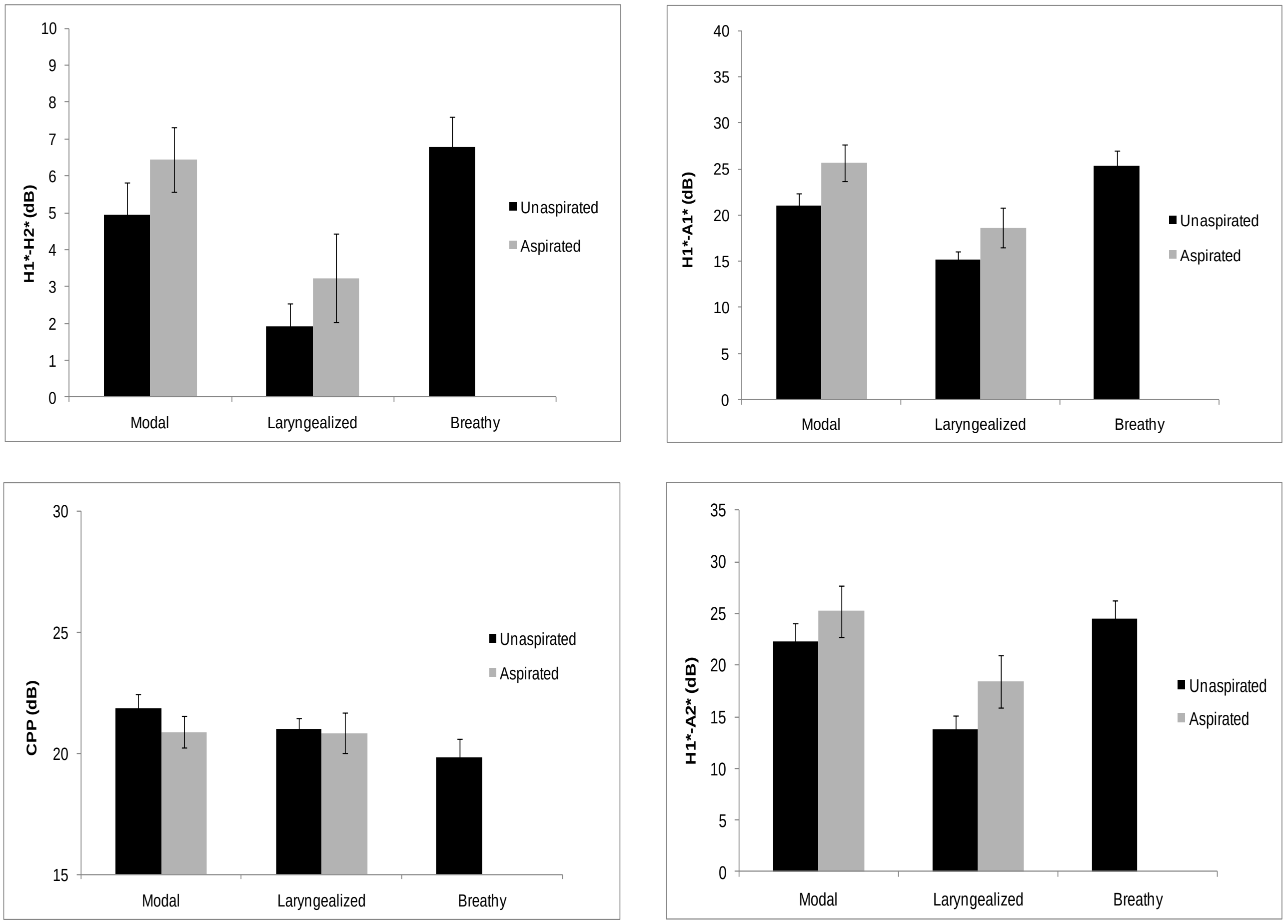
3.6. Aspirated onsets

Does aspiration in onsets affect the voice quality of following vowels? As seen in Figure 5, generally modal and laryngealized vowels following an aspirated stop are breathier than those vowels following an unaspirated stop (though without endangering the phonation contrasts). A main effect of aspirated onset was significant for all measures except for CPP, where the effect was marginally significant ($p = 0.06$), and no aspiration by phonation interactions were found. However, the pairwise comparisons in Table 7 reveal that only for H1*-A2* are the differences between the onset categories significant for both modal and laryngealized phonations, though this is nearly so as well for H1*-A1*, which trends towards significance within laryngealized phonation. CPP shows an effect of onset only within modal, while H1*-H2* shows an effect of onset only within laryngealized.

Table 7. Pairwise modal vs. non-modal comparisons for each measure by aspiration of onsets and phonation of vowel. An asterisk indicates statistical significance at $p < 0.05$.

Acoustic measure	Contrast	Modal	Laryngealized
H1*-H2*	Aspirated vs. unaspirated	0.1542	0.0048*
H1*-A1*	Aspirated vs. unaspirated	0.0003*	0.0605
H1*-A2*	Aspirated vs. unaspirated	0.0018*	0.0147*
CPP	Aspirated vs. unaspirated	0.0148*	0.9838

Figure 5. Influence of aspiration in onsets on acoustic measures of following vowel, compared within phonations. Error bars show 95% confidence intervals around the mean.



There are no comparisons shown for breathy vowels after aspirated vs. unaspirated stops because breathy vowels occur only after unaspirated consonants. After aspirated consonants, the contrast is taken to be neutralized in favor of modal phonation. However, it can be seen in Figure 5 that the values for breathy vowels after unaspirated stops (last bars on the right in each graph) are about the same as the values for modal vowels after aspirated stops (second bars from the left in each graph).

4 Discussion

4.1 Acoustics of Mazatec phonation contrasts

Blankenship (2002), examining a small sample from the Mazatec corpus, found that the three measures she tested, H1-H2, H1-A2, and CPP, all distinguished the modal and breathy phonations, while the first two of these measures distinguished the modal and laryngealized articulations. Esposito (2010a), examining a different small sample of just modal and breathy tokens, but more potential acoustic measures, found that four measures, H1*-H2*, H1*-A1*, H1*-A2*, and CPP, distinguished these two phonations. She also found that in linear discriminant analysis using all the measures, H1*-A2* accounted for fully 53% of the variance, much more than any other measure.

The much larger sample studied here was first examined by linear discriminant analysis, to determine which acoustic measures distinguish the phonation categories. While 7 tested measures were significant in the LDA, only four of them gave significant differences in subsequent mixed effects models, and these were the same four that Esposito (2010) had identified.

A focus of previous work on Mazatec, including Silverman et al. (1995) and Blankenship (2002), was the timecourse of phonation, specifically whether the phonation contrasts are temporally restricted to some sub-part of each vowel. Silverman et al. (1995) had proposed that breathy vowels are breathy only during (approximately) their first half. Blankenship tested this proposal quantitatively, and while she found that laryngealized phonation is distinct from modal on H1-H2 only during the first half of vowels, breathy phonation in fact is distinct on H1-H2 for the whole vowel, even during the middle of the vowel, when the breathiness is somewhat reduced. She also found that all the vowels, which were utterance-final, became breathier over time, and that this effect was a reason for the reduced contrasts at the ends of the vowels.

In our data, the phonation contrasts were strongest in the first third of each vowel; in this portion all modal vs. nonmodal distinctions were significantly different on all four of the reliable acoustic measures. Still, the phonation categories often remain distinct in the middle thirds of vowels, and in the case of modal vs. laryngealized contrasts, even in the last third. Thus our results extend Blankenship's with respect to acoustic measures of phonation, including the temporal extent of phonation, though in our sample the contrasts seem to have been even more robust over time.

Neither Blankenship (2002) nor Esposito (2010a) included formant frequency measures, but Kirk et al. (1993) had shown, in yet another small sample from the corpus, that F1 values were higher for laryngealized phonation, attributed to larynx raising. In our sample, however, while F1 made a significant contribution to the LDA, again it was not significant in mixed effects models. That is, across a large sample of words, including different tones, there is no clear evidence for vocal tract change.

In the first third, the three-way phonation contrast can be fully distinguished using either H1*-H2*, H1*-A1*, or H1*-A2*. These measures differentiate the phonations along a continuum; suggesting that although these phonations may be produced using multiple articulations, a single continuum of glottal states can adequately represent the phonation contrast in Mazatec.

4.2 Effects of speaker sex

Main effects of speaker sex were found for some, though not all, of the important cues to phonation contrasts in Mazatec. Surprisingly, these main effects suggested that in some ways the men's voices were generally breathier than the women's voices: men's values for H1*-A2* and to some extent H1*-A1* were overall higher than women's. However, previous observations about gender differences are typically based on differences in values for H1-H2. In our data, H1*-H2* did not differ significantly (in either direction) between the sexes, indicating that on this key measure, men were neither breathier nor creakier. Finally, men's values for CPP were overall lower, meaning that their voices were less modal – less periodic and/or noisier, for example. These variations in how the sexes differ along the different measures underscores that non-modal phonations can be articulated in different ways, so that potentially men and women phonate in ways that can appear both breathier or creakier, depending on the measure and its articulatory correlate.

Our inspection of the figures in Blankenship (1997), which was based on a small subset of the Mazatec corpus, suggested that in her data, there was no overall difference between the sexes. Instead, the women made larger contrasts than the men did. Their breathier phonation was breathier than the men's, but their laryngealized phonation was less breathy than the men's. Such contrast enhancements are not seen in our larger selection from the corpus. Instead, there are overall differences in scale along the voice measures, preserving the phonation contrasts on each measure, but at different absolute values.

4.3 Effects of tone

Generally, there was no main effect of tone on the acoustic measures included in this study (CPP the only exception, with mid tone the most modal). However, within a given tone, the phonation contrasts were not equally salient. In the first third, the phonation contrasts in low tones were only fully distinguished by H1*-H2* and H1*-A1*. In mid tones, only H1*-A2* distinguished all phonations, and in high tones, no single measure differentiated all the phonations from one another. It is interesting to note, however, that all pairwise phonation contrasts were made for each tone by at least one of the acoustic measures in this study, and more measures support contrasts on low tones than on the other tones. This has implications for perceptual studies of phonation, in that while speakers of languages with phonation contrasts might rely predominantly on a given acoustic measure to perceive such contrasts (Esposito 2010a; Kreiman et al 2010), speakers of those languages might use different acoustic measures depending on the pitch or tone. In addition, mid and low tones were not distinct in laryngealized vowels during the first third, but were distinct in subsequent thirds. This suggests that tonal distinctions are more robust towards the end of the vowel, in contrast to the phonation distinctions, which were found to be most salient during the initial third. This finding supports the claim by Silverman (1997) that tone information may not be recoverable in portions of the vowel with laryngealization. However, contra, we find that tone information in the first third is still salient in breathy phonation.

4.4 Effects of initial consonant

This study also demonstrates that, for common acoustic measures of phonation, aspirated consonants can greatly alter the phonation on following vowels, resulting in neutralization of a phonation contrast. This could help explain why breathy voice does not occur after aspirated stops in languages with both these features, like Mazatec (Silverman et al. 1995) and Hmong (Fulop and Golston 2008). If laryngealized phonation after aspirated stops is more modal, and modal phonation is more breathy, then breathy phonation after aspirated stops would likely be confused for modal phonation.

Even though aspirated stops have been found to induce different breathiness than breathy phonation in Hmong (Fulop & Golston 2008), our results indicate that in Mazatec the effect of aspirated stops is found for all the measures investigated. The similarity between modal phonation after aspirated stops to breathy phonation suggests that aspiration and breathy voice in Mazatec are produced in a similar manner. In this sense, after aspirated consonants, the modal-breathy contrast can be said to be neutralized in favor of breathy phonation, rather than in favor of modal as the traditional description has it. This finding is relevant for all studies of vowel phonation, in that it shows that the type of consonant can have significant effects on the following vowel.

5 Conclusion

Jalapa Mazatec is unusual in possessing a three-way phonation contrast and a three-way level tone contrast independent of phonation. For this reason, it is particularly suited for studying how a three-way phonation contrast is maintained across variables like speaker sex, tone, and vowel time course. With the aid of the VoiceSauce program for voice analysis, in this study we have examined a larger portion of the extensive recordings of Mazatec made by Kirk and Ladefoged in the 1980s and 1990s, comprising all tokens with low vowels and level tones. The results of our acoustic and statistical analysis support the claim that spectral measures like H1-H2 and mid-range spectral measures like H1-A2 best distinguish each phonation type, though other measures like CPP are important as well. This holds true regardless of tone and speaker sex. In Mazatec, the phonation contrasts are strongest in the first third of the vowel and then weaken towards the end of the vowel (which is in utterance-final position in this corpus), but even in the latter third of the vowel some distinctions are maintained. This study shows that using multiple measures, the complex and typologically-rare orthogonal three-way phonation and tone contrasts do remain acoustically distinct, despite partial neutralizations in any given measure. This emphasizes the value of using multiple acoustic cues to characterize phonation in a given language. On the other hand, the acoustic neutralization between modal vowels after aspirated stops and breathy vowels is well explained, given the lack of a breathy-modal contrast following aspirates in the Mazatec lexicon.

Acknowledgments

This work was supported by NSF grant BCS-0720304 to P. Keating and FQRSC grant 128451 to M. Garellek. We would like to thank Yen-Liang Shue for VoiceSauce, Larina Luu for help with manuscript preparation, and Barbara Blankenship and Paul Kirk for discussion. An earlier version was presented at the Spring 2010 meeting of the Acoustical Society of America.

References

Andruski, Jean & Martha Ratliff. 2000. Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association* 30(1/2), 37–61.

Avelino, Heriberto. 2010. Acoustic and electroglottographic analyses of nonpathological, nonmodal phonation. *Journal of Voice* 24(3), 270-280.

Baayen, R. H., Davidson, D. J., & Bates, D. M. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390-412.

Baken, R.J. & Orlikoff, R.F. 2000. *Clinical Measurement of Speech and Voice* (Singular, San Diego), pp. 185-187.

Blankenship, Barbara. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30, 163–191.

Blankenship, Barbara. 1997. *The time course of breathiness and laryngealization in vowels*. Ph.D. thesis, UCLA.

Boersma, Paul & David Weenink. 2008. Praat: Doing phonetics by computer. www.praat.org (3 December 2009).

Belotel-Grenié, A., & M. Grenié. 2004. The creaky voice phonation and the organization of Chinese discourse. International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing.

Davison, Deborah S. 1991. An acoustic study of so-called creaky voice in Tianjin Mandarin. *UCLA Working Papers in Phonetics* 78, 50-57.

DiCanio, Christian T. 2009. The acoustic of register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39(2), 162-188.

Edmondson, Jerold A. 1997. Voice qualities and inverse filtering in Chong. *Mon-Khmer Studies* 26, 107–116.

Edmondson, Jerold A. & John H. Esling. 2006. The valves of the throat and their functioning in tone, vocal register, and stress: Laryngoscopic case studies. *Phonology* 23(2), 157–191.

Epstein, Melissa. 2002. *Voice quality and prosody in English*. Ph.D. thesis, UCLA.

Esposito, Christina. 2004. Santa Ana del Valle Zapotec Phonation. *UCLA Working Papers in Phonetics* 103, 71–105.

Esposito, Christina M. 2005. An acoustic and electroglottographic study of phonation in Santa Ana del Valle Zapotec. Poster presented at the 79th meeting of the Linguistic Society of America, San Francisco, CA.

Esposito, Christina M. 2006. *The effects of linguistic experience on the perception of phonation*. Ph.D. thesis, UCLA.

Esposito, Christina M. 2010a. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38, 306-316.

Esposito, Christina M. 2010b. Variation in constrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40, 181-198.

Esposito, C. (submitted). An acoustic and electroglottographic study of White Hmong phonation, Macalester College ms.

Gerfen, Chip, & Kirk Baker. 2005. The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics* 33, 311-334.

Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.

Gordon, Raymond, Jr. (ed.), 2005. *Ethnologue: Languages of the world*, 15th edn. Dallas, TX: SIL International. http://www.ethnologue.com/web.asp (July 2009).

Hanson, Helen. 1997. Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America* 101(1), 466–481.

Hanson, Helen M., Kenneth N. Stevens, Hong-Kwang Jeff Kuo, Marilyn Y. Chen & Janet Slifka. 2001. Towards models of phonation. *Journal of Phonetics* 29(4), 451-480.

Harrington, Jonathan (2010). *Phonetic Analysis of Speech Corpora*. Wiley-Blackwell.

Henrich, Nathalie, Christophe d'Alessandro & Boris Doval. Spectral correlates of voice open quotient and glottal flow asymmetry: theory, limits, and experimental data. In *Eurospeech 2001*, Aalborg, Denmark, September 2001.

Hillenbrand, J., Cleveland, R. A. & Erickson, R. L. 1994. Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research* 37, 769–778.

Hirose, Hajime, Hirohide Yoshioka & Seiji Niimi. 1978. A cross-language study of laryngeal adjustment in consonant production. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 12, 61–71. Tokyo: University of Tokyo.

Holmberg, Eva B., Roger E. Hillman, Joseph Perkell, Peter Guiod & Susan L. Goldman. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research* 38, 1212–1223.

Holmberg, Eva B., R. E. Hillman and J. S. Perkell. 1989. Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch. *Journal of Voice* 3, 294-305.

Hombert, Jean-Marie, John Ohala & William Ewan. 1979. Phonetic explanations for the development of tones. L*anguage* 55(1), 37-58.

Hombert, Jean Marie. 1979. Consonant types, vowel height and tone. In Victoria A. Fromkin (ed.), *Tone: A linguistic survey*, 77–111. New York: Academic Press.

Iseli, M., Y. Shue, M. Epstein, P. Keating, J. Kreiman, & A. Alwan. 2006. Voice source correlates of prosodic features in American English: A pilot study. Proc. Interspeech 2006, Pittsburgh, October 2006.

Iseli, Markus, & Abeer Alwan. 2004. An Improved Correction Formula for The Estimation of Harmonic Magnitudes and Its Application to Open Quotient Estimation. Proc. ICASSP, Montreal, May 2004, pp. 669-672.

Kawahara, Hideki, Alain de Cheveigne, & Roy D. Patterson. 1998. An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: revised TEMPO in the STRAIGHT-suite. In *ICSLP-1998*, paper 0659.

Keating, Patricia & Yen-Liang Shue. 2009. Voice quality variation with fundamental frequency in English and Mandarin. *Journal of the Acoustical Society of America* 126, 2221-2226.

Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. In Sharon Hargus & Keren Rice (eds.), *Athabaskan prosody*, 137–184. Amsterdam: John Benjamins.

Kirk, Paul L., Jenny Ladefoged & Peter Ladefoged. 1993. Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. In Anthony Mattina & Timothy Montler (eds.), *American Indian linguistics and ethnography in honor of Lawrence C. Thompson*, 435–450. Ann Arbor, MI: University of Michigan.

Kirk, Paul L., Peter Ladefoged & Jenny Ladefoged. 1984. Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics* 59, 102–113.

Klatt, Dennis. 1980. Software for a cascade parallel formant synthesizer. *Journal of the Acoustical Society of America* 67, 971–995.

Klat, Dennis & Laura Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87, 820-857.

Kreiman, Jody, Bruce Gerratt & Norma Antoñanzas Barroso. 2007. Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research* 50, 595–610.

Kreiman, Jody, Markus Iseli, J. Neubauer, Yen-Liang Shue, Bruce R. Gerratt & Abeer Alwan. 2008. The relationship between open quotient and H1*-H2*. *Journal of the Acoustical Society of America* 124 (4): 2495.

Kuang, Jianjing. 2010. An acoustic and electroglottographic study of phonation contrast in Yi. *Journal of the Acoustical. Society of America* 127(3), 2022-2022.

Ladefoged, Peter. 1971.*Preliminaries to linguistic phonetics*. Chicago: University of Chicago.

Ladefoged, Peter. 1983. The linguistic use of different phonation types. In V*ocal fold physiology*: *Contemporary research and clinical issues* (D. Bless & J. Abbs, eds), pp. 351-360. San Diego: College Hill Press.

Laver, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Miller, Amanda. 2007. Guttural consonants and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics* 35(1), 56-84.

Ní Chasaide, Ailbhe & Christer Gobl. 1997. Voice source variation. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences,* 427–461. Oxford: Blackwell.

R Development Core Team. 2007. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org (14 January 2009).

Samely, Ursula. 1991. *Kedang* (*Eastern Indonesia*), *some aspects of its grammar*. Hamburg: Helmut Buske Verlag.

Shue, Yen-Liang, Patricia Keating & Chad Vicenik. 2009. VOICESAUCE: A program for voice analysis. *Journal of the Acoustical Society of America* 126, 2221(A).

Silverman, Daniel, Barbara Blankenship, Paul Kirk & Peter Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* 37(1), 70-88.

Silverman, Daniel. 1997. *Phasing and recoverability* (Outstanding Dissertations in Linguistics). London: Routledge.

Sjölander, Kåre. 2004. Snack sound toolkit, KTH Stockholm, Sweden. http://www.speech.kth.se/snack.

Stevens, Kenneth N. 1977. Physics of laryngeal behavior and larynx modes. *Phonetica* 34, 264-279.

Thongkum, Theraphan. 1988. Phonation types in Mon-Khmer languages. In Osamu Fujimura (ed.), *Vocal physiology: Voice production, mechanisms and functions*, 319–333. New York: Raven Press.

Titze, Ingo R. (1995) Definitions and nomenclature related to voice quality. In O. Fujimura & M. Hirano (eds.), V*ocal fold physiology*: *voice quality control*, 335-342. San Diego: Singular.

Traill, Anthony. 1985. *Phonetic and phonological studies of the !Xóõ Bushman*. Hamburg: Helmut Buske Verlag.

Wayland, Ratree & Allard Jongman. 2003. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31, 181–201.

Appendix: Wordlist

ʔã²tʃã¹ ndæ̤¹	my horse
ndæ̤¹	horse
jæ̤¹	boil
jo¹	there
ⁿdæ̤¹	horse
ⁿdja̤¹	animal horn
ngʷa̤¹	he puts on
na²m̯i²tʃa̰²	nobody
ⁿdæ̰²	companion, man
n̯ka̰²	high
ntsæ̰²	brother
sæ̰²	to exist
tʰæ̰²	sorcery
tʃæ̰²	lazy
ti³fi²kʰæ²	is finished
tsa²	moral
tsæ²	much
tsʰæ²	spotted
tʃa³	old
tʃu¹kʰa³	skunk
ha³	men

tʃu^{1}jæ̤^{1} turtle
jo̤^{2} flesh
ʔjũ^1 nda^2 very good

mæ̤^{2}na^2 I want
ngʷa̤^{2} I will put on
ti^{3}mã^{2}ndzæ̤^{2} visible
ti^{3}βa̤^{2} he hits
ndʒa̤^{3} ʃu^3 chocolate drink
ti3βa̤^{3}ʔa^1 weave
ʔã^{2}tʃã^1 ndæ̰^{1} my buttocks
wa̰^{1} / βa̰^{1} thus
tʃa̰^{1} load
tʃʰa̰^{1} spoon
ʔi^{13} ʔja^1 big leafcutter ants
jæ̰^{1} manure
jo̰^{1} there
kʷa̰^{1} it will happen
ⁿdæ̰^{1} buttock
tsæ̰^{1} his, hers, theirs

βa̰^{2} carries, transports
ha̰^{2} he passed
w̥a̰^{2} passes
w̥a̰^{2} passes
βæ̰^{3} hits
tʃa̰^{3} load, burden
ɸi^{2}kʰa̰^{3} is going to bring
tʃa̰^{3} load
βæ̰^{3} hits
ʔã^{2}tʃã^{1}ntʰæ1 my seeds
tʃʰa^{1}tæ^1 wasp
ja^1 kind of ant
jæ^1 boil (noun)
k̥ʷha^1 will happen
na^1 woman
(n)tʰæ^1 seed
ntsʰa^1 hair
n̥tsʰæ^1 kind of gourd
βo̤^{2} hungry
hæ^2 finished
ka^2 bald
ti^{3}fi^{2}kʰæ^2 is finished
ka^{2}ma^{2}ta^2 it became thick
ki^{2}kæ^2 I saw him
k̥ʷʰæ^2 file
ja^3 tree, wood
ntæ^3 shoes
sʰo^3 wall

stʰæ^3 garbage
tsʰa^3 gives
jæ̰^{2} excrement
næ̰^{2} becomes
ⁿdæ^2 deceased
nta^2 soft
tæ^2 ten
ta^{2}tʰa^2 sticky
tʰæ^2 itch
ⁿda^2 good
ja̰^{2} brings, transports