The acoustic consequences of phonation and tone interactions in Jalapa Mazatec

Marc Garellek and Patricia Keating

Journal of the International Phonetic Association / Volume 41 / Issue 02 / August 2011, pp 185 - 205
DOI: 10.1017/S0025100311000193, Published online: 12 July 2011

The acoustic consequences of phonation and tone interactions in Jalapa Mazatec

Marc Garellek & Patricia Keating

Phonetics Laboratory, Department of Linguistics, UCLA

marcgarellek@ucla.edu

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 Cepstral Peak Prominence are important as well. This is true regardless of tone and speaker sex. The phonation type contrasts are strongest in the first third of the vowel and then weaken towards the end. Although the tone categories 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. Consistent with phonological work on languages that cross-classify tone and phonation type (i.e. 'laryngeally complex' languages, Silverman 1997), 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 type 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 (Gordon 2005). Jalapa Mazatec has a five-vowel system with length and nasalization contrasts. In addition, there are three tone levels (low, mid, and high) and three phonation type contrasts (breathy, modal, and laryngealized). The laryngealized phonation has in the past been referred to as 'creaky' (Kirk, Ladefoged & Ladefoged 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).

Journal of the International Phonetic Association (2011) 41/2
doi:10.1017/S0025100311000193

In their survey of phonation types in the world's languages, Gordon & Ladefoged (2001) cite few languages with phonation type 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. However, at least some dialects of Hmong belong here too, and other varieties of Chong that combine three phonation types to contrast four registers (Huffman 1985, Thongkum 1988, Silverman 1996, DiCanio 2009). A language with three phonation contrasts not included in Gordon & Ladefoged (2001) is Ekoka !Xung (König & Heine 2001). Languages with more than three contrastive phonation types are of course very rare, and none are mentioned by Gordon & Ladefoged (2001), but they include languages with four phonation type contrasts such as Bor Dinka (Edmondson & Esling 2006), and Ju|'hoansi (Miller 2007). Finally, !Xóõ (Traill 1985) has been described as having five phonation type contrasts: breathy, modal, glottalized, pharyngealized, and strident.

Jalapa Mazatec is rare in contrasting three phonation types and three tones independently. Most languages (and even other Mazatec dialects) with phonation type and tone contrasts have constraints on which phonation and tone combinations are legal. These languages are categorized by Silverman (1997) as being 'laryngeally simplex', in contrast to the 'laryngeally complex' languages that do indeed fully cross-classify phonation type and tone. The Otomanguean and Nilotic (e.g. Dinka, Andersen 1993) language families are unusual in that they include languages that distinguish several phonation types and tones independently on vowels. The independent tone and phonation type contrasts in Jalapa Mazatec make the language particularly suited for investigating how phonation type contrasts may vary by tone, speaker sex, and time. Previous studies of Mazatec have ignored or controlled tone contrasts to focus on the phonation type contrasts, with the exception of Silverman (2003), which focused on the perception of Mazatec tones in modal vs. breathy vowels.

Like previous studies (Blankenship 1997, Silverman 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 type 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 type 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 type 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 measures of voice (an argument made explicitly by Blankenship 2002); and third, that a direct relation has been found between measures related to average glottal opening and the acoustics, as will be described below.

However, clearly the activity of the vocal folds can vary in more ways than represented by glottal stricture, e.g. Laver (1980), Hanson (1995), Hanson et al. (2001), and other literature reviewed by 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 type 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 type 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 was introduced by Bickley (1982), following earlier work showing that H1 was important in the perception of breathiness, e.g. Fischer-Jørgensen (1967). H1-H2 has been shown to correlate with Open Quotient (OQ) (Holmberg et al. 1995) or, alternatively, with Contact Quotient derived from electroglottography (DiCanio 2009, Kuang 2011). Other work suggests that H1-H2 also correlates with the skew of the glottal pulse (Henrich, d'Alessandro & Doval 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 phonation types in many studies. For example, in a cross-linguistic sample of breathy versus modal phonation, Esposito (2006, 2010a) found that H1-H2 distinguished these phonation types 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, Gerratt & Antoñanzas-Barroso (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) found that H1-A3 and H1-H2 were both fairly good at distinguishing the phonation types 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, non-modal phonations, especially breathy voice, have been quantified by the presence of noise. Cepstral Peak Prominence (CPP) is thought to reflect the harmonics-to-noise ratio (Hillenbrand, Cleveland & Erickson 1994). It is similar to the harmonics-to-noise ratio measure of de Krom (1993), but differs in how the 'prominence' of the cepstral peak is calculated, i.e. it is taken as the difference in amplitude of the cepstral peak and a regression line used to normalize for window size and overall energy. A more prominent cepstral peak indicates stronger harmonics above the floor of the spectrum. This in turn can result from greater periodicity in the speech signal (thus, little jitter or shimmer). CPP has specifically 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 eight measures she considered at distinguishing modal from breathy phonation types. Recent studies have applied this or similar noise measures (e.g. harmonics-to-noise ratio, jitter, and shimmer measures) to differentiate a variety of phonation types such as breathy, glottalized, and epiglottalized voices, e.g. Andruski & Ratliff (2000) and Andruski (2006) on Mong; Blankenship (2002) on Mazatec, Chong, and Mpi; Wayland & Jongman (2003) on Khmer; Avelino (2010) on Yalálag Zapotec; and Miller-Ockhuizen (2003) 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 phonation types, 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* (where the asterisks indicate that these measures were corrected for vowel formants), each distinguished Mazatec breathy and modal phonation types. Furthermore, however, her 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, American women tend to have breathier voices than men (Klatt & Klatt 1990, Hanson & Chuang 1999), though the opposite was found to be true for Ju|'hoansi by Miller-Ockhuizen (2003) and for Yi by Kuang (2011). 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 (SADV; Esposito 2004, 2005, 2010b). She found that in this language, the three phonation types (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 type contrasts may be produced differently by men and women. In contrast, it appears from the figures in Blankenship (1997: figures 70–73) that Mazatec women made larger distinctions among the phonation types 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 timecourse of phonation has been shown to differ across phonation types and languages as well (see review in Section 4 of Gordon & Ladefoged 2001). Silverman (1997) provides a typological overview of various languages, including Mazatec, with phonation type and tone contrasts that cross-classify (laryngeally complex languages). His findings indicate that many laryngeally complex languages sequence their phonation and tone, which Silverman claims is to ensure better recoverability of both contrasts. He also reports that in some laryngeally complex languages such as Mpi and Tamang, no sequencing occurs. However, in these languages he claims that the non-modal phonation is less robust, allowing for information about phonation and tone to be encoded simultaneously.

Phonation type contrasts have been found to be most pronounced at the start of a vowel in Mazatec (Blankenship 1997). In Mazatec (see Miller-Ockhuizen 2003 for similar findings in Ju|'hoansi), it has also been found that phrase-final vowels tend to end breathy, regardless of their phonation, and this makes all the phonation types 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, but 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 (2004) found that Zapotec creaky vowels only deviate from modal ones at the ends of vowels.

There are several ways in which tone and phonation might 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, Ohala & Ewan 1979, Kingston 2005). Second, and conversely, different F0s can differ in their voice quality. Some studies (Holmberg, Hillman & Perkell 1989, Epstein 2002) have not found a strong correlation between pitch and glottal parameters or LF measures (Fant, Liljencrants & Lin 1985), but others (Iseli, Shue & Alwan 2007, Keating & 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, certain tones are often 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 & Yu 2010). Finally, and perhaps relatedly, phonation categories can be constrained to occur only with certain tone categories. For example, in Southern Yi (Kuang 2011), the phonation contrast does not occur with a high tone, whereas the opposite pattern occurs in Northern Yi. In the laryngeally-simplex language SADV Zapotec (Esposito 2010b), non-modal phonation types occur only with falling tone, and 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, where tone and phonation are orthogonal contrasts. However, it is possible that phonation type 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. Therefore, from both phonetic and typological studies, there is evidence that non-modal phonation in Mazatec is likely to vary by sex, tone and time.

2 Language materials

2.1 Recordings

The sample words come from two field recordings from San Felipe Jalapa de Díaz, 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). The four speakers from the 1984 recording are the speakers studied in Kirk, Ladefoged & Ladefoged (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). Both recordings, originally analog, were digitized at a sampling rate of 44.1 KHz, 16-bit amplitude resolution, and are available online at the UCLA Phonetics Lab Archive website.

2.2 Sample words

In keeping with Blankenship (2002), the sample words chosen for this study all had non-nasalized vowels, and most target vowels were word-final, except for the two words with breathy vowels with a high tone, which were only found on non-word-final syllables. But

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'	^{n}dja̤^{1} 'animal horn'
Mid tone	βa̰^{2} 'carries'	hæ^{2} 'finished'	^{n}da̤^{2} 'good'
High tone	βæ̰^{3} 'hits'	ha^{3} 'men'	ndʒa̤^{3}ʃu^{3} 'chocolate drink'

unlike in the previous study, the target words could have any of three level tones and any of the three phonation types. 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. Only level tones were chosen in this study, though Jalapa Mazatec has contour tones as well (Kirk 1966). Table 1 gives a sample of the Mazatec words chosen (and the rest are listed in the Appendix).

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 20% were breathy, 40% were creaky, and 40% 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 nine words from 12 speakers, for a total of 108 tokens, and Esposito (2010a), who used 16 words (eight breathy, eight modal) from each of three speakers, for a total of 48. All the phonation-tone permutations had speakers of both sexes and from both recordings, except, as already noted, the breathy high-tone 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, Keating & Vicenik 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 also obtained automatically in VoiceSauce, which uses the correction algorithm of Iseli et al. (2007). Our analysis included only the corrected difference measures, which, here as elsewhere in the literature, will be written with asterisks (e.g. H1*-H2*). Formant values were calculated using the Snack Sound Toolkit (Sjölander 2004), while F0 was calculated using the STRAIGHT algorithm (Kawahara, de Cheveigné & Patterson 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.

Table 2 Statistical results of the Linear Discriminant Analysis for the significant measures. The larger absolute correlation between each variable and either function is indicated with an asterisk.

	Correlation with discriminant functions				
Acoustic measure	Function 1	Function 2	Wilks' Lambda	F-value	Significance
H1*-H2*	.695*	−.070	.760	25.294	<.001
H1*-A1*	.776*	−.045	.816	58.399	<.001
H1*-A2*	.715*	.162	.698	20.196	<.001
CPP	−.056	.599*	.786	32.919	<.001
F0	−.224	.494*	.722	22.770	<.001
F1	.140*	.066	.685	17.755	<.001
B4	.176	−.301*	.673	15.967	<.001

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 in the 424 tokens. All the acoustic measures listed in Section 2.3 were included in the discriminant analysis. The values for these measures were 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.

The results of the LDA are shown in Table 2. Two discriminant functions were computed in the discriminant analysis because the phonation contrast has three categories (breathy, modal, or laryngealized). Each acoustic measure was assigned to one of the functions, depending on which function it correlates more highly with. Asterisks in the correlation columns of Table 2 indicate to which discriminant function a particular acoustic measure was assigned. Seven of the 16 measures were significant in the analysis: H1*-H2*, H1*-A1*, H1*-A2*, CPP, F0, F1, and B4, and the correlations of these measures with the discriminant functions are shown in Table 2.

The results of the LDA indicate that the harmonic measures H1*-H2*, H1*-A1*, H1*-A2*, as well as F1, correlate significantly with Function 1, whereas F0, CPP, and B4 correlate significantly 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 seven measures.

3.2 Timecourse during vowels

Blankenship (2002) found that the effects of phonation type on a variety of acoustic measures are strongest in about the first one-third to one-half of the vowels, are weaker later in the vowel, and are generally lost by the ends of vowels. In our study, we divided the vowels into one-third intervals and then compared the phonation types by linear mixed-effects models on the measures that were significant in the LDA, with the acoustic measure as a dependent variable and random intercepts for speaker and word. The linear mixed-effects model was chosen as the method of analysis because it allows for missing data. In this study, there were no breathy-high tokens uttered by women, yet the linear mixed-effects could still determine the effects of sex for the three phonation types, as will be shown in Section 3.4. Unlike LDA

Table 3 Pairwise modal vs. non-modal comparisons for the same seven acoustic measures at each vowel third. Asterisks indicate statistical significance at $p < .05$.

Acoustic measure	Contrast	First third	Middle third	Final third
H1*-H2*	Breathy vs. Modal	<.0001*	.0007*	.9207
	Laryngealized vs. Modal	<.0001*	<.0001*	<.0001*
	Breathy vs. Laryngealized	<.0001*	<.0001*	<.0001*
H1*-A1*	Breathy vs. Modal	<.0001*	.0693	.461
	Laryngealized vs. Modal	.0002*	.0001*	.0124*
	Breathy vs. Laryngealized	<.0001*	<.0001*	.1112
H1*-A2*	Breathy vs. Modal	<.0001*	.0099*	.7413
	Laryngealized vs. Modal	.0001*	<.0001*	.0001*
	Breathy vs. Laryngealized	<.0001*	<.0001*	.0001*
CPP	Breathy vs. Modal	.0001*	.0202*	.2329
	Laryngealized vs. Modal	.011*	.0008*	.0825
	Breathy vs. Laryngealized	.0783	.4823	.7108
F0	Breathy vs. Modal	.3848	.8807	.2397
	Laryngealized vs. Modal	.8561	.1758	.6367
	Breathy vs. Laryngealized	.4715	.2732	.4478
F1	Breathy vs. Modal	.4778	.5665	.8862
	Laryngealized vs. Modal	.4835	.7363	.3862
	Breathy vs. Laryngealized	.1659	.3642	.3343
B4	Breathy vs. Modal	.4821	.7256	.1127
	Laryngealized vs. Modal	.1967	.0016*	.0029*
	Breathy vs. Laryngealized	.0489*	.0081*	.2579

which pools all the measures together, the linear mixed-effects models here are designed to look at each acoustic measure individually, and they control for variability across speaker and item. Figure 1 shows the seven measures that were significant in the LDA by vowel-thirds, and Table 3 gives the significance of each comparison between modal and non-modal phonation types, which were calculated using linear mixed-effects models with phonation type as a fixed effect and speaker and word as random effects. The mixed-effects modeling was run in R (R Development Core Team 2008) using the *lme4* 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).

Similarly to Blankenship, we find that the phonation types 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 linear mixed-effects models indicate that these differences are not significant at $p = .05$, 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 type 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 phonation types, 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 phonation types 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 not significant, at $p = .07$.

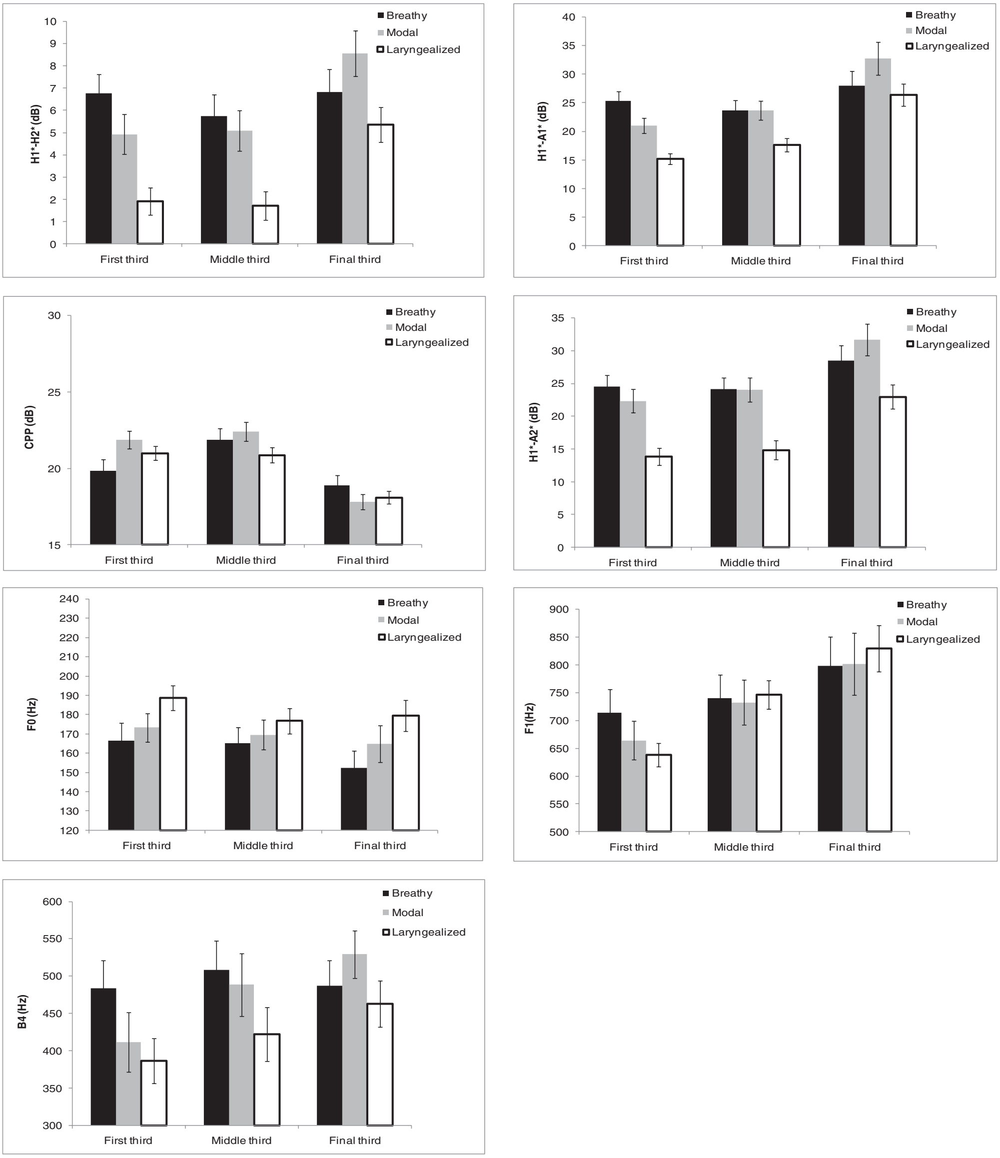

Figure 1 Acoustic measures by vowel thirds (with 95% confidence intervals), showing that differences across phonation types are greatest in the first third.

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 borne 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 and/or an interaction.

Table 4 Pairwise modal vs. non-modal comparisons for each acoustic measure by sex and phonation during the initial third. Asterisks indicate $p < .05$.

Acoustic measure	Contrast	Breathy	Modal	Laryngealized
H1*-H2*	Women vs. men	.4686	.087	.1183
H1*-A1*	Women vs. men	.0289*	.7824	.3588
H1*-A2*	Women vs. men	.0031*	.0794	.0025*
CPP	Women vs. men	.0119*	.1186	.0093*

For the four phonation measures (the LDA measures, excluding F0, F1, and B4), a significant main effect of phonation was found ($p < .001$). From Table 3 we see that both non-modal phonation types differ from modal on the four measures reported there (H1*-H2*/A1*/A2*, CPP). This finding extends those from Blankenship (2002), in which breathy vs. modal differed on all three of the parameters she tested (H1-A2/H2, 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 phonation types are usually well differentiated, even in the final third, for the spectral measures (excluding H1*-A1*) but not for CPP. However, the results for CPP are in keeping with Miller (2007), who found that Harmonics-to-noise ratios can capture the acoustic SIMILARITY of non-modal phonation types, rather than their differences. CPP can be low either because of aperiodicity (as in laryngealized phonation) or because of noise (as in breathy phonation).

3.4 Sex differences

Next we consider whether the two sexes differed significantly in how they used the four measures to distinguish phonation types. 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 ($p < .00357$ for the former, and $p = .01125$ for the latter). The direction of the differences for CPP and H1*-A2* combined may indicate that men are breathier than women. Although lower CPP values could also be a result of laryngealization, the fact that H1*-A2* is higher in men suggests that the lower CPP value is due to breathiness, since laryngealization should also result in lower H1*-A2* values. 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 $p < .09$. A similar difference was found by Blankenship (1997).

Our main interest here, however, is whether men and women differ in how they implement the three phonation types, especially with respect to the acoustic measures that best differentiate the phonation types overall as described above. Differences in how the sexes implement 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 phonation types. 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.

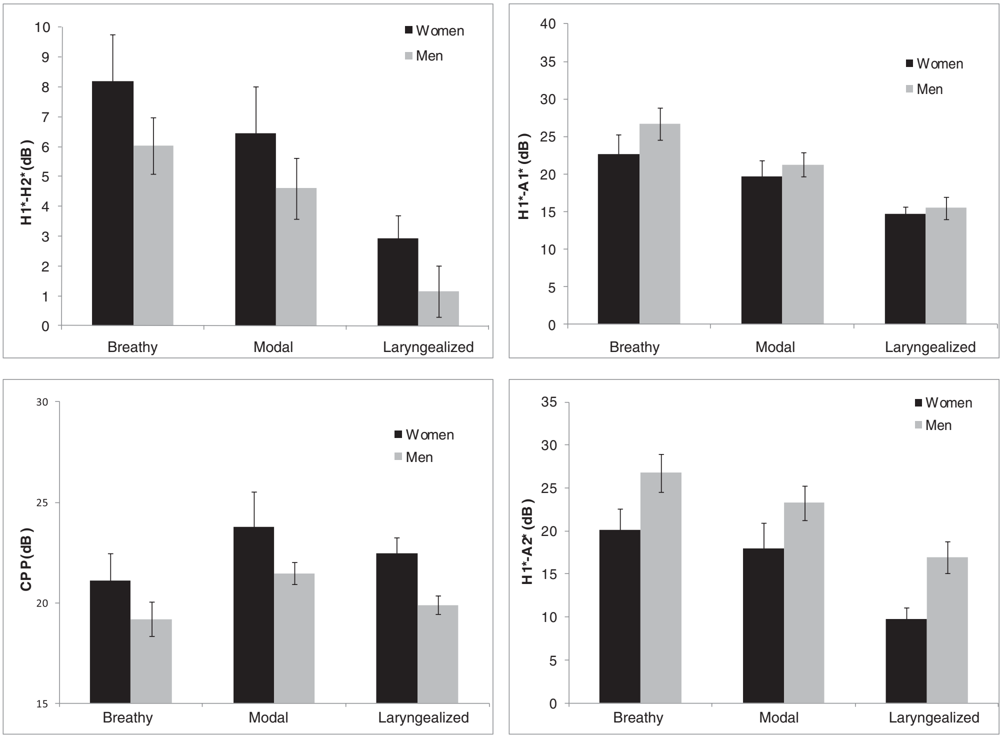

Figure 2 Acoustic measures for women vs. men compared within phonation types during the initial third. Error bars show the 95% confidence interval around the mean.

3.5 Phonation by tone interactions

Jalapa Mazatec is unusual in having independent tones and phonation types, and all nine combinations of them. Nonetheless, at least in part because acoustic measures of voice quality can vary with F0, we hypothesize that the tones, in addition to the phonation categories, will differ along one or more of our voice quality measures. Conversely, we also hypothesize that the phonation type contrasts will be more robust on some tones than on others, perhaps least robust on high tones. Finally, we hypothesize 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 ($p = .002$), 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* ($p = .02$). 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 type contrasts more robust on some tones than on others? Because the tone by phonation interactions go in different directions, sometimes there is contrast

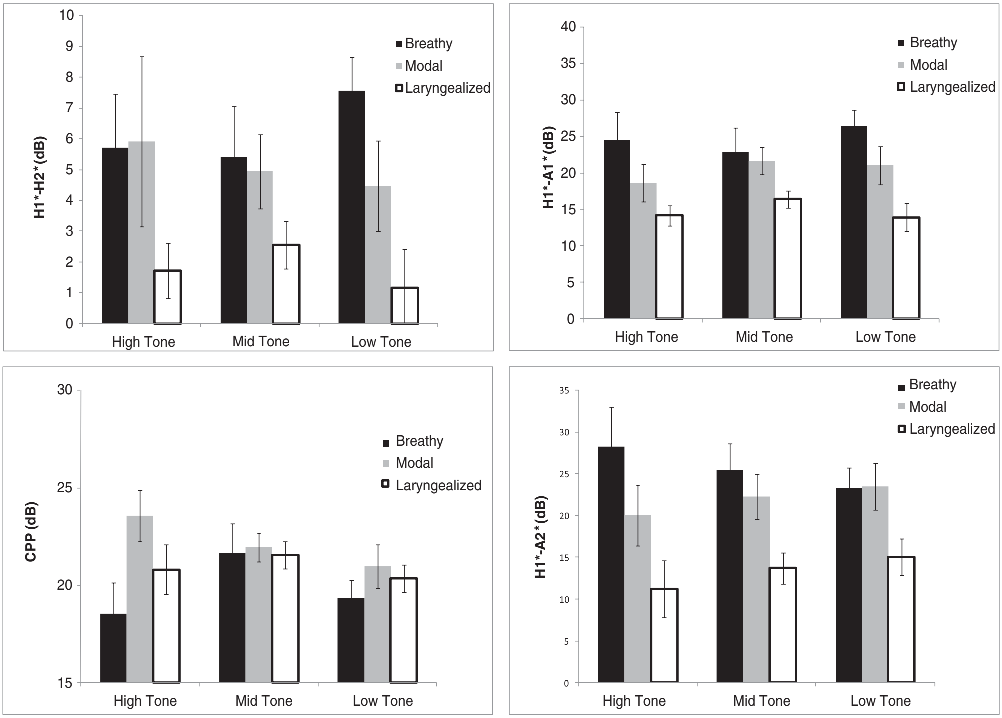

Figure 3 Acoustic measures for phonation types compared within tones during the initial third. Error bars show 95% confidence intervals around the mean.

enhancement, other times contrast reduction. Thus the phonation by 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 phonation types only on low tones, while CPP works best on high and low tones. H1*-A2* can distinguish all three phonation types only on mid tones. Thus, in terms of how well each measure (taken separately) distinguishes the phonation types within each tone category, it seems that the evidence is mixed and no single tone best supports the phonation type contrasts.

Alternatively, we can consider the robustness of phonation type 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 phonation types is distinguished by at least two of the measures, with the exception of breathy vs. modal on mid tones, where only H1*-A2* shows a significant difference. Phonation types are overall distinguished by the most measures on low tones (three out of four per contrast); the breathy vs. modal contrast is especially less distinct on high and mid tones. Thus, in terms of the set of measures which individually make significant distinctions among the phonation types within each tone category, it seems that the phonation contrast is more robust with low tones.

Third, do non-modal phonation types differ in F0 from modal phonation? That is, can F0 alone distinguish phonation types? Figure 1 appears to show such differences, and F0 was a significant measure in the initial Linear Discriminant Analysis, but no main effect of phonation type on F0 was found in the subsequent Linear Mixed Effects analysis. Pairwise

Table 5 Pairwise modal vs. non-modal comparisons for each acoustic measure by tone within the initial third. An asterisk indicates statistical significance at $p < .05$.

Acoustic measure	Contrast	High tone	Mid tone	Low tone
H1*-H2*	Breathy vs. Modal	.1718	.1032	.0011*
	Laryngealized vs. Modal	.0047*	.0006*	.0023*
	Breathy vs. Laryngealized	<.0001*	<.0001*	<.0001*
H1*-A1*	Breathy vs. Modal	.1001	.0783	.0002*
	Laryngealized vs. Modal	.1950	.0249*	.0144*
	Breathy vs. Laryngealized	.0043*	.0006*	<.0001*
H1*-A2*	Breathy vs. Modal	.0001*	<.0001*	.2882
	Laryngealized vs. Modal	.0953	.0032*	.0006*
	Breathy vs. Laryngealized	<.0001*	<.0001*	<.0001*
CPP	Breathy vs. Modal	.0005*	.2603	.0074*
	Laryngealized vs. Modal	.0131*	.0982	.3664
	Breathy vs. Laryngealized	.1761	.8725	.0491*

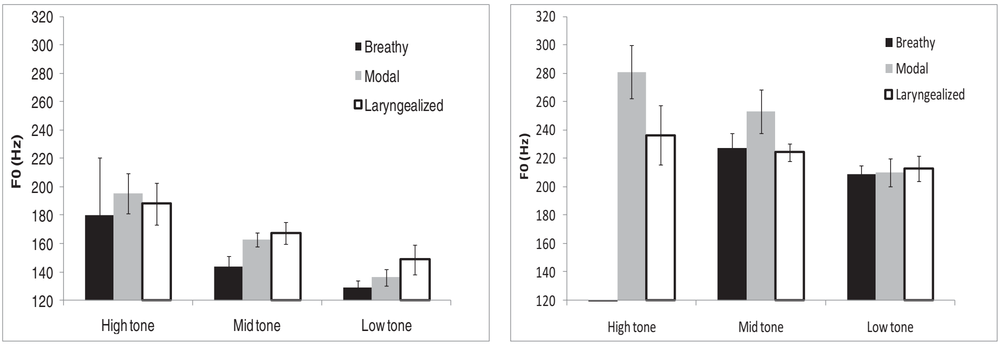

Figure 4 F0 for phonation types compared within tones during the initial third, separated by sex, with men on the left and women on the right. Error bars show 95% confidence intervals around the mean.

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 $p < .10$. 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 ($p < .0001$ during both the middle and final thirds). This suggests that tone contrasts are strongest after the initial third, at least for laryngealized vowels.

3.6 Aspirated onsets

In this study we also expected to find effects of coarticulation from aspirated onsets on the voice quality of following vowels. Such laryngeal coarticulation has been shown in a number of languages, including Swedish, French, German, Italian (e.g. in Gobl & Ní Chasaide 1999), English (Löfqvist & McGowan 1992, Garellek 2010), White Hmong (Garellek 2010), Korean (Cho, Jun & Ladefoged 2002, Garellek 2010), Navajo, Tagalog (Blankenship 1997),

Table 6 Pairwise tonal comparisons for F0 by phonation. An asterisk indicates statistical significance at $p < .05$.

Acoustic measure	Contrast	Breathy	Modal	Laryngealized
F0	High vs. Mid	.0002*	.0001*	.0143*
	Mid vs. Low	.0055*	.0014*	.0957
	High vs. Low	<.0001*	<.0001*	.0004*

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 $p < .05$.

Acoustic measure	Contrast	Modal	Laryngealized
H1*-H2*	Aspirated vs. unaspirated	.1542	.0048*
H1*-A1*	Aspirated vs. unaspirated	.0003*	.0605
H1*-A2*	Aspirated vs. unaspirated	.0018*	.0147*
CPP	Aspirated vs. unaspirated	.0148*	.9838

and Ju|'hoansi (Miller-Ockhuizen 2003, Miller 2007). Although these studies differ in the acoustic measures used to quantify the effects of laryngeal coarticulation, vowels tend to be breathier after aspirated consonants. Thus, in our study we hypothesize that H1*-H2*, H1*-A1*, and H1*-A2* should be higher in vowels following aspirated stops, whereas CPP should be lower. 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 type contrasts). A main effect of aspirated onset was significant for all measures except for CPP, where the effect was close to being significant at $p = .06$, 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 phonation types, though this is nearly so as well (at $p = .06$) for H1*-A1* within laryngealized phonation. CPP shows an effect of onset only within modal, while H1*-H2* shows an effect of onset only within laryngealized.

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 suspended in favor of what is characterized as modal phonation. However, it can be seen in Figure 5 that the values for breathy vowels after unaspirated stops (last bar on the right in each graph) are about the same as the values for modal vowels after aspirated stops (second bar from the left in each graph).

4 Discussion

4.1 Acoustics of Mazatec phonation type contrasts

Blankenship (2002), examining a small sample from the Mazatec corpus, found that the three acoustic measures she tested, H1-H2, H1-A2, and CPP, all distinguished the modal and breathy phonation types, while the first two of these measures distinguished the modal and laryngealized articulations. Esposito (2006), 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 phonation types. 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.

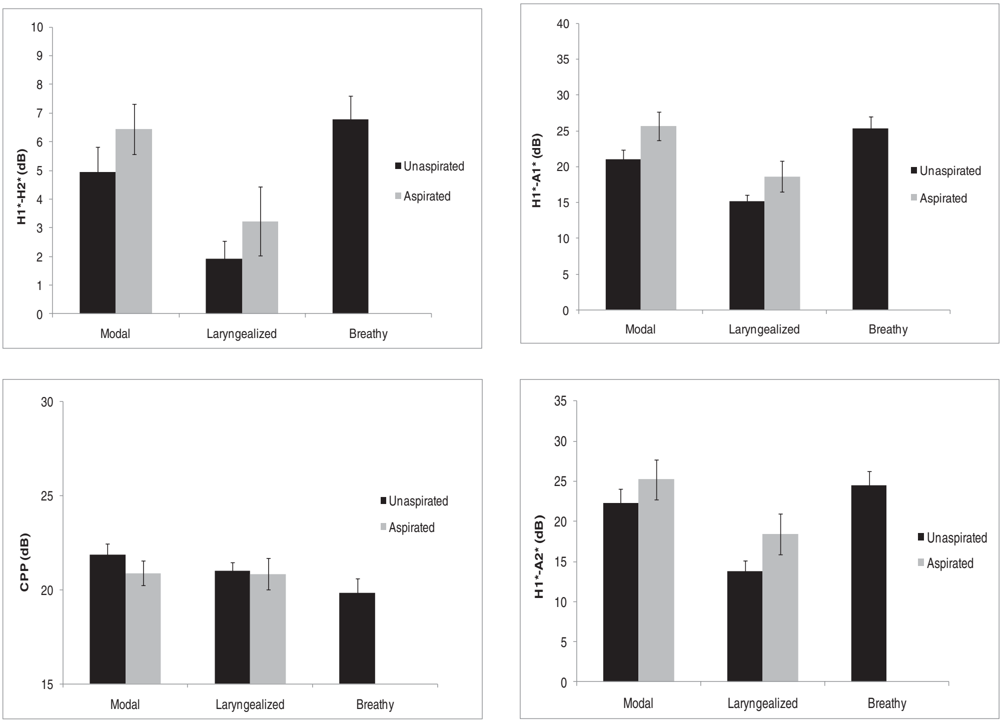

Figure 5 Influence of aspiration in onsets on acoustic measures of following vowel, compared within phonation types. Error bars show 95% confidence intervals around the mean.

The much larger sample studied here was also first examined by Linear Discriminant Analysis, to determine which acoustic measures distinguish the phonation categories. While seven tested measures were significant in the LDA, only four of them individually gave significant differences in subsequent mixed effects models, and these were the same four that Esposito (2010a) had identified. Thus it seems clear that for this set of recordings, these measures are the most useful in distinguishing the phonation types, but that other measures can weakly contribute to making these distinctions.

A focus of previous work on Mazatec, including Silverman et al. (1995) and Blankenship (2002), was the timecourse of phonation, specifically whether the phonation type 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 type 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 be even more robust over time.

Neither Blankenship (2002) nor Esposito (2006) 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 on its own in mixed effects models. That is, there is no reliable vocal tract change across a large sample of words including different tones.

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 phonation types along a continuum; suggesting that although these phonation types 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 type 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 phonation types 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 were 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 type 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 type contrasts were not equally salient. In the first third, the phonation type contrasts in low tones were only fully distinguished by H1*-H2* and H1*-A1*. In mid tones, only H1*-A2* distinguished all phonation types, and in high tones, no single measure differentiated all the phonation types from one another. It is interesting to note, however, that all pairwise phonation type 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 type contrasts might rely predominantly on a given acoustic measure to perceive such contrasts (Esposito 2010a, Kreiman, Gerratt & Khan 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, 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, on common acoustic measures of phonation, aspirated consonants can greatly alter the phonation of following vowels, resulting in suspension of a phonation contrast. This could help explain why, in languages with both these features, aspirated stops cannot be followed by breathy voice, as in Mazatec (Silverman et al. 1995) and Hmong (Fulop & Golston 2008, Esposito & Khan 2010), by modal voice, as in Gujarati (Esposito & Khan 2010), or by any non-modal phonation, as in Ju|'hoansi (Miller-Ockhuizen 2003, Miller 2007). 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.

Our results indicate that in Mazatec the effect of aspirated stops is found for all the measures investigated. The fact that modal phonation in vowels following aspirated stops and vowels with breathy phonation are acoustically similar 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 type contrasts are strongest according to the measures included in the study 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 Patricia Keating and FQRSC grant 128451 to Marc 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.

Appendix. Wordlist from UCLA Phonetics Archive

ʔã^{2}tʃ ã^{1}ndæ̤^{1} my horse
ndæ̤^{1} horse
jæ̤^{1} boil
jo^{1} there
^{n}dæ̤^{1} horse
^{n}dja̤^{1} animal horn
ngʷa̤^{1} he puts on
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}m ã^{2}ndzæ̤^{2} visible
ti^{3}βa̤^{2} he hits
ndʒa̤^{3} ʃu^{3} chocolate drink
ti^{3}βa̤^{3}ʔa^{1} weave
ʔã^{2}tʃ ã^{1} ndæ̰^{1} my buttocks
wa̰^{1}/βa̰^{1} thus
tʃa̰^{1} load
tʃʰa̰^{1} spoon
ʔi^{31} ʔja^{1} big leafcutter ants
jæ̰^{1} manure
jo̰^{1} there
kʷa̰^{1} it will happen
^{n}dæ̰^{1} buttock
tsæ̰^{1} his, hers, theirs
βa̰^{2} carries, transports
ha̰^{2} he passed
w̥a̰^{2} passes
na^{2}m̥i^{2}tʃa̰^{2} nobody
^{n}dæ̰^{2} companion, man
n̥ka̰^{2} high
ntsæ̰^{2} brother
sæ̰^{2} to exist
tʰæ̰^{2} sorcery
tʃæ̰^{2} lazy
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}tʃ ã^{1}ntʰæ^{1} my seeds
tʃʰa^{1}tæ^{1} wasp
ja^{1} kind of ant
jæ^{1} boil (noun)
kʷ̥ha^{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
kʷ̥ʰæ^{2} file
ti^{3}fi^{2}kʰæ^{2} is finished
tsa^{2} moral
tsæ^{2} much
tsʰæ^{2} spotted
tʃa^{3} old
tʃu^{1}kʰa^{3} skunk
ha^{3} men
ja^{3} tree, wood
ntæ^{3} shoes
sʰo^{3} wall
stʰæ^{3} garbage
tsʰa^{3} gives
jæ̰^{2} excrement
næ̰^{2} becomes
^{n}dæ^{2} deceased
nta^{2} soft
tæ^{2} ten
ta^{2}tʰa^{2} sticky
tʰæ^{2} itch
^{n}da^{2} good
ja̰^{2} brings, transports

References

Andersen, Torben. 1993. Vowel quality alternation in Dinka verb inflection. *Phonology* 10, 1–42.

Andruski, Jean E. 2006. Tone clarity in mixed pitch/phonation-type tones. *Journal of Phonetics* 34, 388–404.

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

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

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

Baken, Ronald J. & Robert F. Orlikoff. 2000. *Clinical measurement of speech and voice*. San Diego, CA: Singular.

Bickley, Corine. 1982. Acoustic analysis and perception of breathy vowels. *MIT Speech Communication Working Papers* 1, 73–83.

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

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

Belotel-Grenié, Agnès & Michel Grenié. 2004. The creaky voice phonation and the organization of Chinese discourse. *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, 5–8. Bejing: ISCA.

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

Cho, Taehong, Sun-Ah Jun & Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30, 193–228.

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

de Krom, Guus. 1993. A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *Journal of Speech and Hearing Research* 36, 254–266.

DiCanio, Christian T. 2009. The phonetics 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. dissertation, UCLA.

Esposito, Christina M. 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. dissertation, 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 contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40, 181–198.

Esposito, Christina M. & Sameer ud Dowla Khan. 2010. Contrastive breathiness on consonants and vowels. *Journal of the Acoustical Society of America* 128, 2476.

Fant, Gunnar, Johan Liljencrants & Qi-guang Lin. 1985. A four-parameter model of glottal flow. *STL-QPSR* 4, 1–13.

Fischer-Jørgensen, Eli. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguistics* 28, 71–139.

Fulop, Sean A. & Chris Golston. 2008. Breathy and whispery voicing in White Hmong. *Proceedings of Meetings on Acoustics* 4, 060006.

Garellek, Marc. 2010. The acoustics of coarticulated non-modal phonation. *UCLA Working Papers in Phonetics* 108, 66–112.

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

Gobl, Christer & Ailbhe Ní Chasaide. 1999. Voice source variation in the vowel as a function of consonantal context. In William J. Hardcastle & Nigel Hewlett (eds.), *Coarticulation: Theory, data and techniques*, 122–143. Cambridge: Cambridge University Press.

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/15/web.asp (July 2009).

Hanson, Helen M. 1995. *Glottal characteristics of female speakers*. Ph.D. dissertation, MIT.

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

Hanson, Helen M. & Erika S. Chuang. 1999. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America* 106, 1064–1077.

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.

Henrich, Nathalie, Christophe d'Alessandro & Boris Doval. 2001. Spectral correlates of voice open quotient and glottal flow asymmetry: Theory, limits, and experimental data. In *EUROSPEECH-2001*, 47–50. Aalborg, Denmark, September 2001.

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

Holmberg, Eva B., Robert E. Hillman & Joseph 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.

Holmberg, Eva B., Roger E. Hillman, Joseph S. 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.

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

Huffman, Franklin E. 1985. The phonology of Chong: A Mon-Khmer language of Thailand. In Suriya Ratanakul, David Thomas & Suwilai Premsrirat (eds.), *Southeast Asian linguistic studies presented to André-G. Haudricourt*, 355–388. Bangkok: Mahidol University.

Iseli, Markus, Yen-Liang Shue & Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustical measures related to the voice source. *Journal of the Acoustical Society of America* 121, 2283–2295.

Kawahara, Hideki, Alain de Cheveigné & Roy D. Patterson. 1998. An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised TEMPO in the STRAIGHT-suite. *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.

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

Kirk, Paul L. 1966. *Proto-Mazatec phonology*. Ph.D. dissertation, University of Washington.

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.

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.

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

König, Christa & Bernd Heine. 2001. *The !Xun of Ekoka: A demographic and linguistic report*. Cologne: Institut für Afrikanistik, Universität Köln.

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

Kreiman, Jody, Bruce R. Gerratt & Sameer ud Dowla Khan. 2010. Effects of native language on perception of voice quality. *Journal of Phonetics* 38, 588–593.

Kreiman, Jody, Markus Iseli, Juergen 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. 2011. Production and perception of the phonation contrast in Yi. M.A. thesis, UCLA.

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

Ladefoged, Peter. 1983. The linguistic use of different phonation types. In Diane M. Bless & James H. Abbs (eds.), *Vocal fold physiology*: *Contemporary research and clinical issues*, 351–360. San Diego, CA: College Hill Press.

Lam, Hiu Wai & Kristine M. Yu. 2010. The role of creaky voice quality in Cantonese tonal perception. *Journal of the Acoustical Society of America* 127, 2023.

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

Löfqvist, Anders & Richard S. McGowan. 1992. Influence of consonantal envelope on voice source aerodynamics. *Journal of Phonetics* 20, 93–110.

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

Miller-Ockhuizen, Amanda. 2003. *The phonetics and phonology of gutturals: A case study from Ju|'hoansi* (Outstanding Dissertations in Linguistics). New York: Routledge.

R Development Core Team. 2008. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org (14 January 2009).

Shue, Yen-Liang, Patricia Keating & Chad Vicenik. 2009. VOICESAUCE: A program for voice analysis. *Journal of the Acoustical Society of America* 126, 2221. [Program available online at http://www.ee.ucla.edu/~spapl/voicesauce.]

Silverman, Daniel. 1996. Phonology at the interface of morphology and phonetics: Root-final laryngeals in Chong, Korean, and Sanskrit. *Journal of East Asian Linguistics* 5, 301–322.

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

Silverman, Daniel. 2003. Pitch discrimination during breathy versus modal phonation. In John Local, Richard Ogden & Rosalind Temple (eds.), *Papers in laboratory phonology VI*, 293–304. Cambridge: Cambridge University Press.

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

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 Osamu Fujimura & Minoru Hirano (eds.), *Vocal fold physiology: Voice quality control*, 335–342. San Diego, CA: Singular.

Traill, Anthony. 1985. *Phonetic and phonological studies of !Xóõ Bushman* (Quellen zur Khoisan-Forschung 5). 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.