UCLA Working Papers in Phonetics, No. 108, pp. 188-202

Phonation Contrasts Across Languages*

P. Keating[1], C. Esposito[2], M. Garellek[1], S. Khan[3], J. Kuang[1]

[1]UCLA; [2]Macalester College; [3]Brown University

(keating@humnet.ucla.edu)

Abstract

This study compared the contrastive phonation types of four languages—Gujarati (modal vs. breathy), Hmong (modal vs. breathy vs. creaky), Mazatec (modal vs. breathy vs. creaky), and Yi (tense vs. lax)—on several acoustic measures, within and across languages. For Gujarati, Hmong, and Yi, two electroglottographic (EGG) measures were also compared; a Contact Quotient measure distinguished the within-language phonation types in all three languages. While several acoustic measures distinguished phonation types within each language, only H1*-H2* did so in all four languages. However, when each within-language phonation category was then compared across languages, each category was found to differ from language to language on multiple acoustic measures, e.g. breathy in Hmong is distinct from breathy in Gujarati. This unexpected result suggests that language/speaker differences in voice quality are larger than phonation category differences. This suggestion finds support in a Multi-Dimensional Scaling analysis of the acoustic measures. A three-dimensional space turns out to mostly distinguish the languages, less so the phonations. The phonation categories do not form clusters in this space across languages as might have been expected, but they do occupy separate regions along the third dimension of the space, a dimension correlated especially with H1*-H2*. Thus H1*-H2* is again seen to be the most important measure of phonation contrasts across languages.

1. Introduction

Across languages with phonation contrasts, the phonation categories are distinguished by a variety of measures (e.g. Gordon & Ladefoged 2001, Esposito 2010), but not by every measure in each language. In this study we ask the following questions:

- What measures distinguish phonation categories within and across languages?
- What are the dimensions of the cross-language acoustic voice quality space?
- How are the phonation categories of different languages located in this space?

---

* Expanded version of poster presented at LabPhon12 in Albuquerque NM, July 2010.

2. Methods

We compare the contrastive phonations of four unrelated languages on several acoustic measures, and for three languages on measures from electroglottographic (EGG) recordings. Summary information about the languages, their phonation contrasts, and the corpus of recordings is given in Table 1. For each language, a wordlist of words contrasting in phonation, in minimal pairs/sets whenever possible, was compiled. Although these wordlists included vowels of different heights, in the present study we selected only words with low (Gujarati, Mazatec, Hmong) or low and lower-mid (Yi) vowels. Onsets included a variety of consonants, though for most of the analyses reported here, aspirated onsets were excluded. In the three tone languages, tones were also systematically varied in the wordlists; in Mazatec, only level tones were included, and in Yi, only non-high tones were included.

For three of the languages, speakers were recorded specifically for this study, using either a computer soundcard, or PCQuirer with its external D/A box. For Mazatec, however, existing recordings made by Paul Kirk and Peter Ladefoged in the 1980s and 1990s, using reel and cassette tapes, were accessed from the online UCLA Phonetic Archive. For two of the languages, words were spoken in isolation; the Hmong words were spoken in a carrier sentence; in Gujarati, to avoid spelling pronunciations of words with breathy vowels, speakers were asked to create their own sentences beginning with the test words; see Khan (in preparation) for more details.

The Yi and Gujarati speakers as well as 11 of the 32 Hmong speakers also made simultaneous electroglottographic recordings along with the audio, using a Glottal Enterprises EG2 electroglottograph. The Mazatec recordings were audio-only.

Table 1. *Languages included in the study, their contrastive phonations, whether they also contrast tones, and information about the corpus of recordings analyzed.*

Language (variety) (family)	**Phonations**	**Tones**	**Source of recordings**	**# of speakers**	**EGG # of speakers**
Gujarati (Indo-European)	Modal, breathy	No	Fieldwork in Los Angeles CA	10 (7F, 3M)	Yes (7F, 3M)
Hmong (White) (Hmong-Mien)	Modal, breathy, creaky	Yes	Fieldwork in St. Paul MN	32 (9F, 23M)	Yes (5F, 6M)
Mazatec (Jalapa) (Otomanguean)	Modal, breathy, creaky	Yes	UCLA online phonetic archive (fieldwork in Mexico)	16 (6F, 10M)	-None-
Yi (Southern) (Tibeto-Burman)	Lax, tense	Yes	Fieldwork in SW China	12 (6F, 6M)	Yes (6F, 6M)

2.1. Acoustic Measures

Two-channel recordings with audio plus EGG signals were first split, with the audio converted to .wav format and the EGG signal converted to .wav format using a utility in EggWorks, a free UCLA program for EGG analysis by Henry Tehrani[1]. Acoustic measurements over time were made semi-automatically from the audio by VoiceSauce (Shue et al. 2009), a free UCLA program implemented in Matlab[2].

A preliminary step with VoiceSauce is to segment and label the vowels of interest from the test utterances, using Praat to make TextGrid files. A Praat script can be used to help with this task, but to some extent this is done manually. Then VoiceSauce estimates a set of acoustic parameters within these segments, automatically for all the files in a folder. The STRAIGHT algorithm (Kawahara et al. 1998) is used to estimate the fundamental frequency at 1 msec intervals. Harmonic spectra are computed pitch-synchronously over windows of three pitch pulses. Given the F0 estimate, VoiceSauce uses an optimization function to locate the harmonics of the spectrum, and finds their amplitudes. This method greatly reduces variability compared to methods that use a fixed-length window. VoiceSauce then uses the Snack Sound Toolkit (Sjölander 2004) to find the frequencies and bandwidths of the first four formants, also at 1 msec intervals. The harmonics nearest to these formant frequencies are located, and their amplitudes are taken as the amplitudes of the formants. Finally, the formant frequencies, along with stored estimates of their bandwidths, are used in an algorithm that corrects harmonic amplitudes for the effects of the formants, using Iseli et al.'s (2007) extension of Hanson's (1995) method. Corrected harmonic amplitudes are indicated by an asterisk, e.g. H1*. Harmonic amplitude differences (e.g. H1-H2) are calculated for corrected and uncorrected amplitudes. In addition, (pitch synchronous) energy and Cepstral Peak Prominence (Hillenbrand et al. 1994) are calculated. (Subsequent versions of VoiceSauce also have other Harmonic to Noise Ratio measures.) All these parameter values are stored in a Matlab file, with one Matlab file for each audio file. Finally, VoiceSauce produces an output text file for each folder containing audio and Matlab files, giving the mean value of each parameter for each segment in each file, plus (optionally) means over some specified number of sub-segments within each segment.

In the present study, the following measures were tested:

- Corrected (*) harmonic amplitude differences:
    - H1*-H2*
    - H2*-H4*
    - H1*-A1*
    - H1*-A2*
    - H1*-A3* (with H1*-An* representing any of these three)
- Cepstral Peak Prominence (CPP)
- Energy

See Blankenship (2002), Gordon and Ladefoged (2001), Hanson et al. (2001), Bishop and Keating (this volume), or Garellek and Keating (this volume) for descriptions of these measures. H2*-H4* is relatively new, introduced by Kreiman et al. (2007). Bishop and Keating (this

---

[1] EggWorks can be downloaded at http://www.linguistics.ucla.edu/faciliti/facilities/physiology/EggWorksSetup.exe.

[2] VoiceSauce can be downloaded from http://www.ee.ucla.edu/~spapl/voicesauce/.

volume) found that it contributes to listeners' identification of speaker sex, while Kuang (in preparation) found that it distinguishes low and mid tones in southern Yi.

2.2. EGG Measures

We discovered that recordings made via a laptop soundcard, as was the case with our Yi recordings, are inverted. Therefore the first step in processing the EGG signals was inversion of the Yi EGG recordings, a function provided by EggWorks. Then all EGG signals were analysed automatically using EggWorks. This program takes EGG signals in either their original PCQuirer .pmf format, or in .wav format, and calculates five measures (see below) for each glottal pulse that it can find, throughout each entire file. It then interpolates values to 1-msec intervals, to match VoiceSauce, and outputs a text file for each input EGG signal file. An option in VoiceSauce's output function specifies that the EGG outputs should be included in VoiceSauce's output text files as additional parameters.

Four of the five EGG measures calculated by EggWorks are different methods of measuring Contact Quotient CQ (sometimes called the Closed Quotient), one of which will be reported here: CQ_H, or CQ by the Hybrid method. "Hybrid" here refers to the fact that the edges of the contacting phase of the glottal cycle are defined using two different methods (Rothenberg and Mahshie 1988, Orlikoff 1991, Howard 1995, Herbst and Ternström 2006; see also http://voiceresearch.free.fr/egg/thresholdmethods.html#EGGDEGG). The beginning of the contact phase is taken to be the (positive) peak in the first derivative of the EGG signal, which occurs during the increase in contact; the end of the contact phase, however, is *not* based on the negative peak in the derivative during the decrease in contact, but instead is arbitrarily determined using a fixed threshold—in our case 25% (Orlikoff 1991)—of the difference between the minimum and maximum amplitude values in each cycle of the EGG signal. This threshold method of setting the end of contact is preferred because, while the positive peak in the EGG derivative is almost always well-defined, the negative peak is usually not, making a pure-derivative measure uncertain.

The fifth EGG measure calculated by EggWorks is here called the Peak Increase in Contact, or PIC. This is the peak positive value in the derivative of the EGG signal, presumably equivalent to the DECPA measure of Michaud (2004). This measure gives the instantaneous peak rate of change in contact, and here replaces earlier, non-derivative-based measures of average contacting "speed" (as described by Baken and Orlikoff 2000).

3. Results

3.1. Analyses of individual languages

Within each of the four languages, statistical comparisons of the two or three phonation categories were made to determine which acoustic and (where available) physiological measures distinguished the phonations. These within-language comparisons were based on means over entire vowels in Gujarati, Hmong, and Yi, but over just the first third in Mazatec (where phonation contrasts are strongest in the language). Statistical tests used were either Repeated Measures ANOVAs (Gujarati, Hmong), or linear mixed effects models (in Mazatec, with speaker and item as random effects; in Yi, with speaker as random intercept, and tone and phonation as random slopes). The results of these within-language comparisons are shown in Table 2. More details can be found in the papers reporting on individual-language analyses (Esposito, submitted; Garellek and Keating, this volume; Khan, in preparation; Kuang, in preparation).

Table 2. *Results of within-language tests of significance of phonation contrasts on each acoustic or physiological measurement. For Gujarati, the categories are modal and breathy. For Hmong and Mazatec, the categories are modal, breathy, and creaky. For Yi, the categories are tense and lax. A checkmark in a cell indicates that that measure significantly distinguished some or all of the phonations in that language in the expected direction. N/A indicates that no EGG measures are available for Mazatec.*

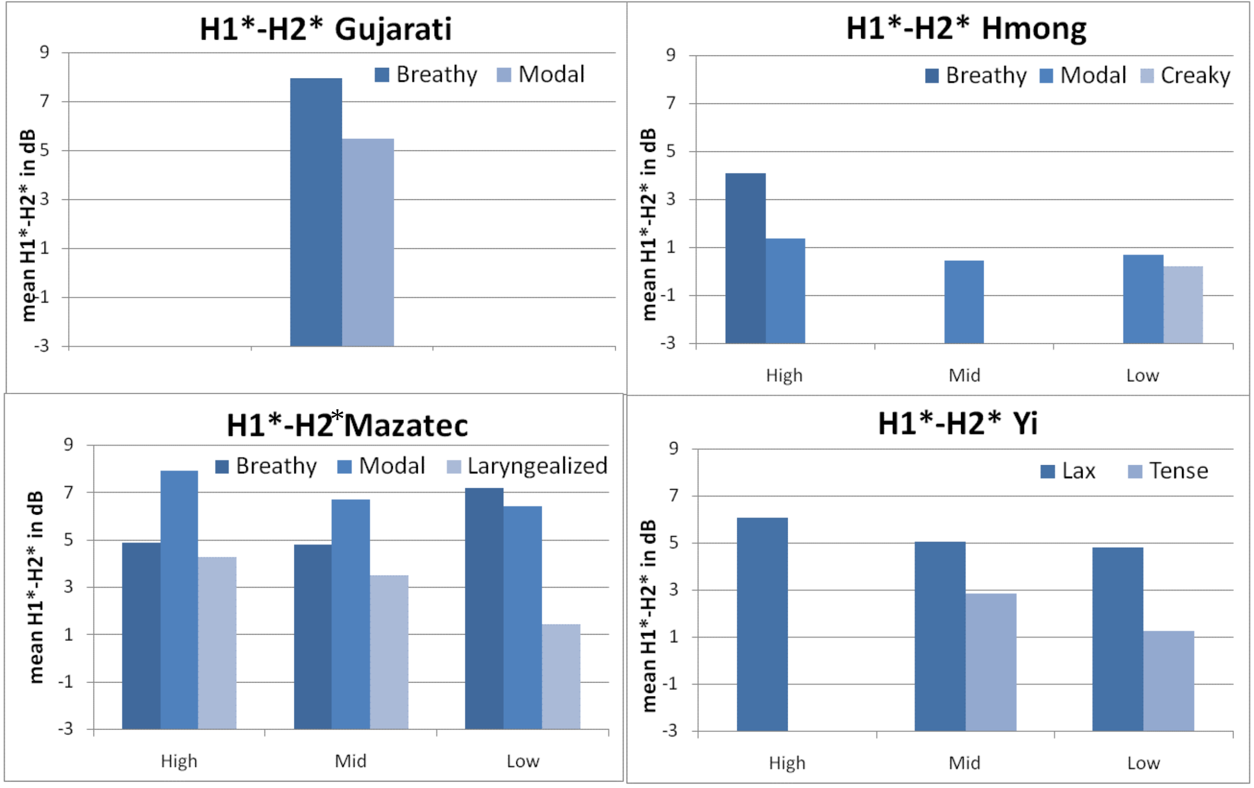
Measure	Gujarati	Hmong	Mazatec	Yi
H1*-H2*	✓	✓	✓	✓
H2*-H4*				
H1*-A1*	✓		✓	✓
H1*-A2*	✓		✓	✓
H1*-A3*	✓		✓	✓
CPP		✓	✓	✓
Energy			✓	
CQ_H	✓	✓	N/A	✓
PIC		✓	N/A	✓

3.1.1. Phonation contrasts

From Table 2, it can be seen that the only acoustic measure that distinguishes phonation categories in all four languages is H1*-H2*. Mean values are shown in Figure 1. As expected, breathy and lax phonations have the highest values, while creaky and tense phonations have the lowest values, though the average differences are often fairly small. However, not every pair-wise comparison is significant. This is partly due to the fact that means over entire vowels are compared here (with the exception of Mazatec); some comparisons are significant only over specific portions of vowels. The Hmong results presented here are different from Esposito (submitted) because the current study averages the measures across the entire vowel, while Esposito looks at three timepoints within a vowel. Another factor is that some comparisons are skewed by imbalances of male and female speakers. Also, it can be seen that the values are scaled somewhat differently from language to language. For example, the modal phonation of Gujarati has a similar mean value to the breathy phonation of Hmong.

The spectral tilt measures, H1*-An*, do not distinguish any of the phonations in Hmong, and energy differs by phonation only in Mazatec. The new spectral measure, H2*-H4*, does not distinguish any of the phonations within any language. Finally, CPP, the measure of noise and/or periodicity, does not distinguish modal from breathy phonations in Gujarati.

Figure 1. *Mean H1*-H2* for the contrasting phonations in the four languages, separately by tone in the three languages with lexical tone, but combining data from men and women. The "Laryngealized" category here in Mazatec is called "creaky" elsewhere in this paper. In Hmong, tones are grouped together into three basic levels. In Yi, tense phonation does not occur with high tone. Values come from means over entire vowels except in Mazatec.*



The last two measures in Table 2 are the EGG measures, available for only three of the four languages (i.e. Gujarati, Yi, Hmong). Contact Quotient using a Hybrid method (CQ_H) with a 25% threshold distinguished at least some of the phonations in all three languages. However, Hmong creaky and modal vowels were not distinguished by this measure. In contrast, Peak Increase in Contact (PIC) did not distinguish the phonations in Gujarati, though it did distinguish modal from creaky in Hmong. Means for CQ_H are shown in Figure 2, and means for PIC are shown for Yi and Hmong, the two languages in which differences were significant, in Figure 3. As expected, CQ_H is lower for breathy and lax phonations, and higher for creaky and tense phonations. That is, the glottis is proportionately more open in the breathier phonations.

However, the results for PIC are contrary to expectations. It might be thought that breathier phonation, typically having a more sinusoidal glottal waveform and a more gradual vocal fold closing, would have lower PIC values (if PIC reflects, even indirectly, the speed of closing of the vocal folds). However, it can be seen in Figure 3 that, in these two languages, the breathy and lax phonations have higher, not lower, PIC values. That is, the rate of change of vocal fold contact is greater for the smoother vibrations of the breathier phonations. Visual inspection of EGG signals suggests that PIC might follow a principle of "the further, the faster", i.e. the larger the amplitude of the glottal cycle, the faster the transition between the open and closed phase. Amplitude changes within glottal cycles in the EGG signal are greatest for breathy phonation, and these changes are also faster. Sample signals (EGG and its derivative) are shown in Figure 4. (Note that EggWorks does not display any signals; this display is from PCQuirerX.)

Figure 2. *EGG Contact Quotient (Hybrid method, 25% threshold) for phonations in three languages, separately by tone in the two languages with lexical tone, but combining data from men and women. In Hmong, tones are grouped together into three basic levels. In Yi, tense phonation does not occur with high tone. Values come from means over entire vowels.*

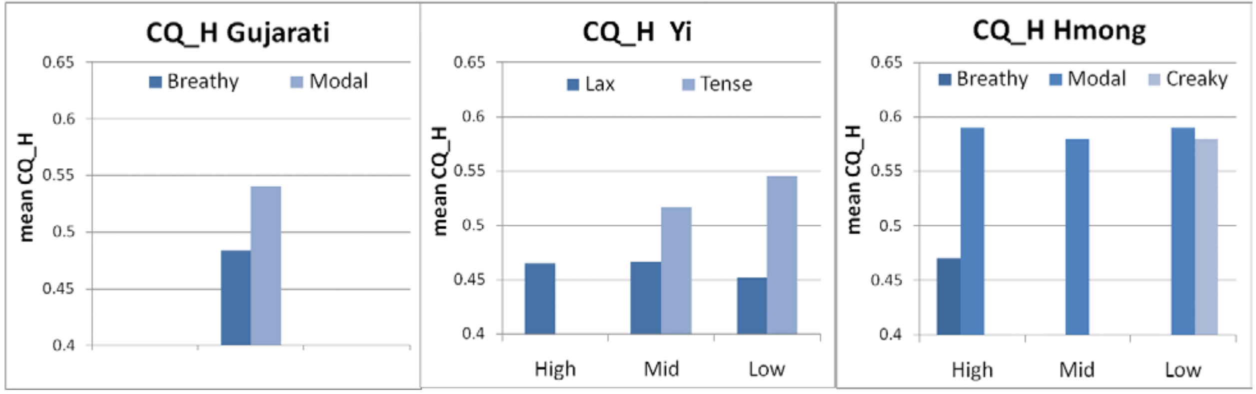


Figure 3. *EGG Peak Increase in Contact for phonations in the two languages in which significant differences were found, separately by tone, but combining data from men and women. In Hmong, tones are grouped together into three basic levels. In Yi, tense phonation does not occur with high tone. Values come from means over entire vowels.*

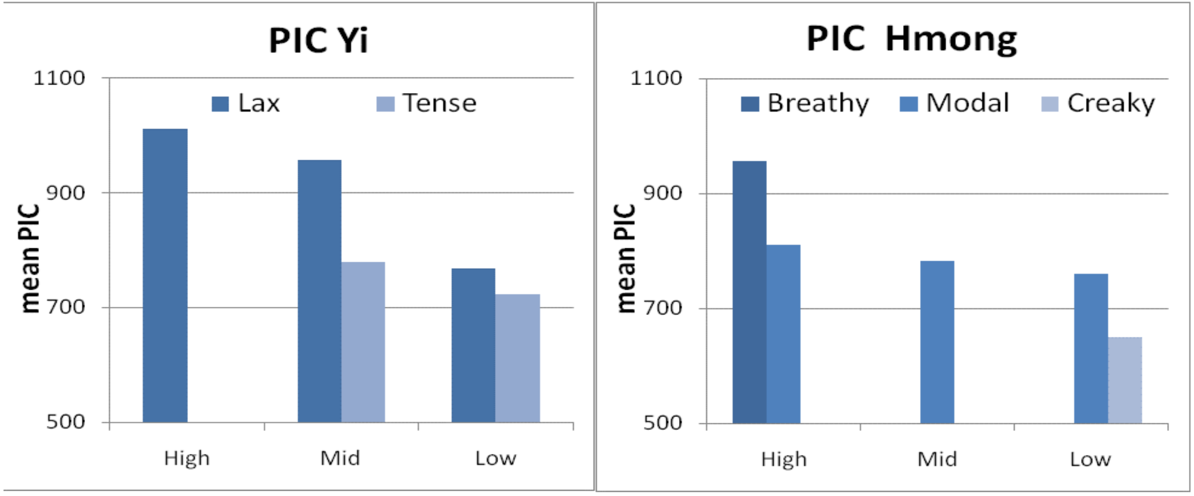
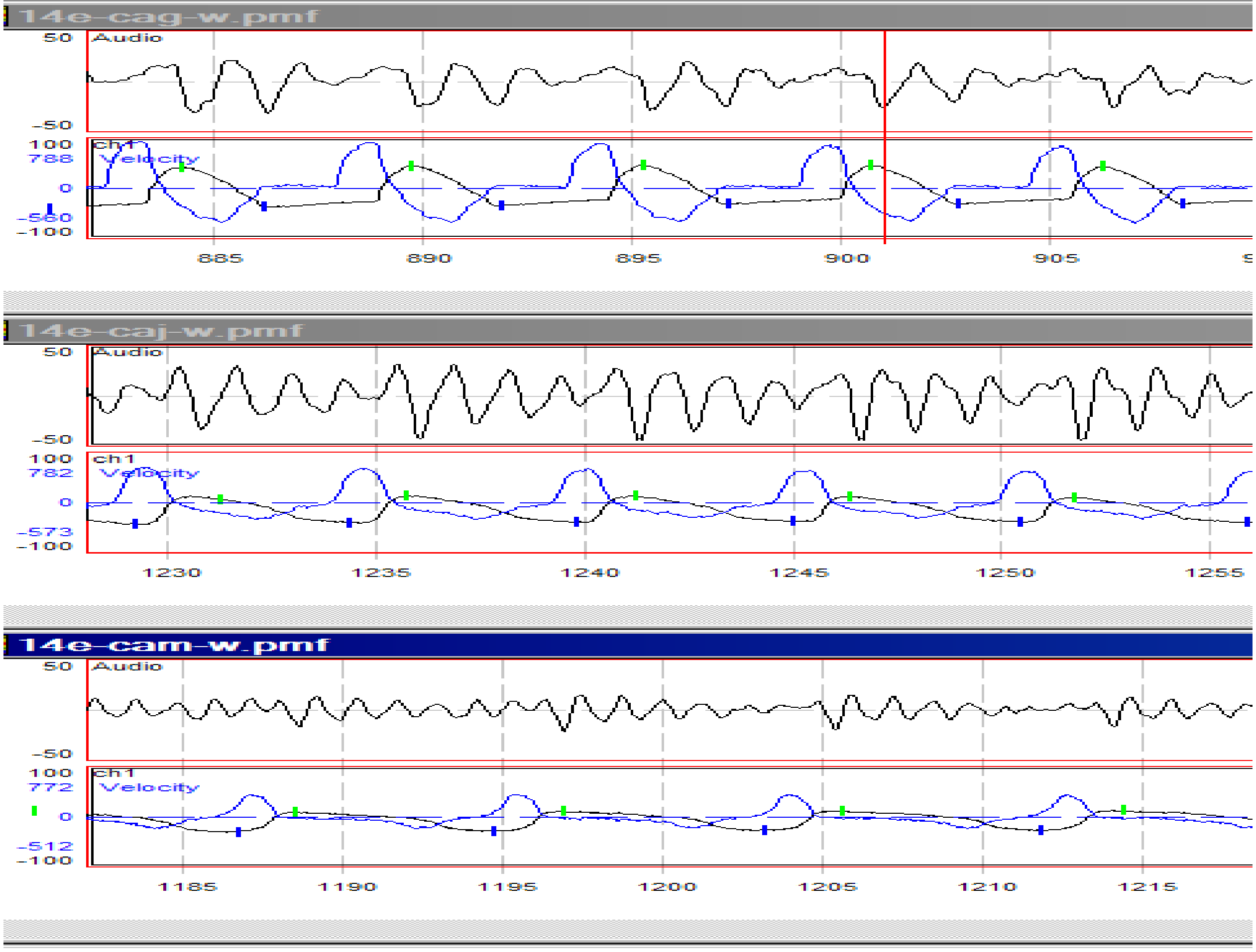


Figure 4. *Sample EGG signals from Hmong, breathy (top) vs. modal (middle) vs. creaky (bottom). Under the audio signal in each panel is the EGG signal with its first derivative (labeled "Velocity"). On the EGG waveform, maximum and minimum values are marked by green and blue tick marks, respectively, and the differences in amplitude among the three phonations can be seen by comparing these marks. The positive peaks in the derivatives differ correspondingly.*



Although we have not carried out statistical correlation analyses, a few observations can be made about qualitative relations between EGG and acoustic measures. First, all three languages for which we have EGG data distinguish their phonations by both the CQ measure of the EGG signal and H1*-H2* from the audio signal. (This is true in Hmong for creaky vs. modal phonations when just the end of the vowel is considered.) It is not surprising that these two measures should pattern together, given previous literature on the relation of H1-H2 to CQ (or its inverse, OQ), e.g. Holmberg et al. (1995) for English, DiCanio (2009) for Takhian Thong Chong, and Esposito (submitted) for Hmong. Second, the PIC measure of the EGG signal distinguished the phonations in the two languages with contrasting creaky or tense phonation (Hmong, Yi). Although speed of vocal fold closure is generally thought to be related to spectral tilt, there is no obvious relation in Table 2 between PIC and the spectral tilt measures H1*-An*[3].

[3] In Yi, however, PIC is slightly correlated with H1*-A3*.

However, these are the languages in which the acoustic measure CPP also distinguished the phonations. The connection between PIC and CPP should be explored further.

3.1.2. Timecourse effects

Table 2 shows results for measures over entire vowels (with the exception of Mazatec), but the languages in fact differ in what portion of a vowel most clearly shows phonation contrasts. Only in Yi is the contrast clear over entire vowels. In Gujarati, contrasts are clearest in the middle of vowels, though the patterns of results are no different when only these portions of vowels are compared. In Mazatec, contrasts are clearest at the beginnings of vowels, and only those data have been presented here. In Hmong, breathiness is concentrated early in vowels, but creakiness is strongest at the ends of vowels. When the phonation types of Hmong are compared at beginnings, middles, and ends of vowels, we find that modal is distinct from creaky only at vowel ends. Modal is distinct from breathy at all three timepoints, but the distinction is found in the most measures at mid-vowel; most notably, CPP is different only there. Creaky and breathy are distinct at all timepoints.

3.1.3. Gender effects

While many measures show main effects of speaker gender (that is, women have overall higher or lower values on a measure), there are no significant interactions of phonation with gender in any of the within-language analyses. Thus it appears that, within each language, men and women make the phonation contrasts in similar ways.

3.1.4. Tone effects

Three of the languages studied here contrast lexical tones as well as phonations. As the vocal folds adjust to vary their rate of vibration for tonal contrasts, voice quality could vary too. Some differences in voice measures across tones appear in Figure 1 (H1*-H2*) and in the EGG measures in Figures 2 and 3 for Yi and Mazatec. However, not all the apparent differences are statistically significant. In Yi, the tones do differ in their H1*-H2* values, and within the lax phonation, in their PIC values. In Mazatec, the only measure that differs with tone is CPP, with Mid tones having the highest CPP value, and Low tones having the lowest value. See Kuang (in preparation) and Garellek and Keating (this volume) for more detail on Yi and Mazatec, respectively.

3.2. Comparison of all language categories

We took each language-specific phonation category, total 10 (2 Gujarati + 3 Hmong + 3 Mazatec + 2 Yi), and compared them all together acoustically. How many of them are distinct? At the most conservative extreme, all 10 could be different from one another, but we might expect that the modal categories would not differ among themselves, and likewise the breathy and creaky categories respectively. The Yi “tense” and “lax” categories might not differ from categories in the other languages, though it would remain to be seen whether Yi lax was more like breathy or more like modal, etc. Thus, at the other extreme, the 10 categories could cluster into 3 groups. Of course, something between these two extremes is also possible: that the 10 categories would cluster into more than three but fewer than 10 groups.

The tokens included in this analysis were as in the previous analysis, but here also included words with aspirated onsets in Gujarati and Mazatec. These comparisons were made with Linear

Mixed Effects models, one for each of the acoustic measures, with speaker and item as random effects. The result was that every category differed from every other, i.e. the most conservative extreme. The modal, breathy, and lax phonations all differed on a large set of measures. The creaky and tense phonations differed on only three measures (H1*-A1*, CPP, and energy), but they were significantly different across all the languages. That is, in this corpus, the result was the perhaps unexpected extreme, that phonation categories with the same descriptive names (e.g. Gujarati breathy, Hmong breathy, and Mazatec breathy) nonetheless significantly differ in several acoustic dimensions. This suggests that speaker and language differences (including differences between our corpora for the different languages) are larger than phonation differences. This possibility is examined in another way in the next analysis.

3.3. Multi-dimensional scaling across languages

The multiple acoustic measures made in this study can be thought of as defining a multi-dimensional acoustic space within which different phonations can be located. However, many of these measures are inter-correlated, so the dimensions of the space are not necessarily independent. Multi-Dimensional Scaling (MDS) is one method for reducing many individual measures to a smaller number of independent dimensions. MDS uses measured distances between items to define a map in which those distances are preserved in a lower-dimensional space.

For the MDS, tokens were further controlled for vowel height, tone, and consonant aspiration, to increase cross-language comparability; only non-high vowel tokens with unaspirated consonants were used for all four languages, and for Mazatec and Yi, only the mid tones were selected. The Manhattan (or city-block) distances on the set of acoustic measurements were used as the basis for estimates of the physical distances between all pairs of tokens, and these distances were inputted into the MDS (performed in R using the isoMDS function). This yielded solutions with different numbers of dimensions, where more dimensions typically do better at preserving the original distances, but too many dimensions offer diminishing returns in data-fitting, and can be hard to interpret and visualize; here, the three-dimensional (3-D) solution seemed best. This solution is shown in Figure 5, in which each of the 10 within-language phonations is plotted in the 3-D space. Each language is plotted in a different color.

If languages dispersed their phonations within an overall phonation space, we would expect to see the same-colored (i.e. cross-phonation) bars spread well apart in the figure. Instead, they tend to cluster together on one or more of the dimensions of the space. That is, the cross-language differences appear greater than the cross-phonation differences. Dimension 1 (shown here as the vertical dimension) primarily distinguishes Gujarati, and to some extent Hmong, from the other languages. Within Gujarati, Hmong, and Yi, the breathy/lax phonation is higher on this dimension; but in Mazatec the modal phonation is highest. Dimension 2 (shown here as the front-to-back dimension) primarily separates the languages into two groups: Gujarati and Yi with positive values, Hmong and Mazatec with negative values. Within-language differences between phonations are small on this dimension.

Thus, these first two, most important, dimensions of the MDS solution serve mainly to put each language (with all its phonation categories) into its own region of the space. This could be because the languages genuinely have overall different voice quality settings; or it could be because of idiosyncratic speaker differences, different recording conditions, or other differences in the corpora.

Dimension 3 (shown here as the left-to-right dimension) seems more important in distinguishing the phonation categories within the languages. The Mazatec and Yi phonations are particularly clearly separated along this dimension. The Hmong modal and creaky are not so well separated, but this is expected since these data are averaged across entire vowels while Hmong creakiness is localized at the ends of vowels. The Gujarati phonations are not well separated on any one dimension, but Dimensions 1 and 3 together may be crucial. Overall, there is no apparent tendency for all instances of any one kind of phonation to cluster tightly together in any part of the space. However, the creaky/tense phonations have negative values on dimension 3, the modal phonations are between -1 and 0, and the breathy/lax phonations have zero or positive values. Thus this dimension of the MDS solution seems to provide a continuum of phonation types along the lines suggested by Blankenship (1997) and Gordon and Ladefoged (2001).

To better understand the basis for these distinctions, the dimensions can be tested for correlations with individual acoustic measures. The weighting of each acoustic measure on each dimension is shown in Figure 6. Dimension 1 is most strongly related to acoustic energy, and to a lesser extent to H2*-H4*, H1*-A3*, and CPP. Recall that while this dimension mostly separates the languages, it also contributes to distinguishing breathy/lax phonation from other phonations, apparently on the basis of spectral tilt and noise. Dimension 2 is most strongly related to energy and H1*-H2*. The strong relations of these first two dimensions to acoustic energy underscores that these dimensions are mainly characterizing language/speaker, not phonation, differences. Dimension 3, which does the most to distinguish the phonations, is most strongly related to H2*-H4* and H1*-H2*. The importance here of H2*-H4* is surprising, since unlike H1*-H2* it never distinguished the phonations within languages; it must be contributing to the between-language distinctions seen even on this dimension. Thus H1*-H2* seems the most important dimension for distinguishing the phonations, and it also contributes to the language differences seen on Dimension 2. Two spectral tilt measures, H1*-A1* and H1*-A2*, are not strongly related to any of the MDS dimensions.

4. Conclusions

In this study we asked what measures best distinguish phonation categories, both within and across languages. Comparing the different phonations within a language for each of our four test languages, we found, as expected, that several acoustic measures differentiate the categories within each language, but only one measure did so in all of the languages: H1*-H2*. One EGG measure did so in all three languages with EGG data: Contact Quotient (Hybrid method). Comparing each phonation category across languages, differences were found on several measures, though the creaky/tense phonations differed on fewer measures than the other categories did. These consistent cross-language differences were not expected, and suggest that language/speaker differences in voice quality are larger than phonation category differences. This suggestion finds support in an MDS analysis of the acoustic measures. A three dimensional space turns out to mostly distinguish the languages, less so the phonations. The phonation categories do not form clusters in this space across languages as might have been expected, but they do occupy separate regions along the third dimension of the space, a dimension correlated with H1*-H2* and H2*-H4*. While H1*-H2* distinguished the phonations in the within-language statistical comparisons, H2*-H4* did not, and therefore it is unlikely that this measure is contributing to distinguishing the phonations in this space. More likely, it differs across the languages/speakers. On this interpretation, H1*-H2* is again seen to be the most important measure of phonation contrasts across languages.

Figure 5. *3-D MDS solution for the ten phonation categories, rotated so that Dimension 3 is seen most clearly. Each language is shown in a different color. The heights of the colored bars show the values on Dimension 1.*

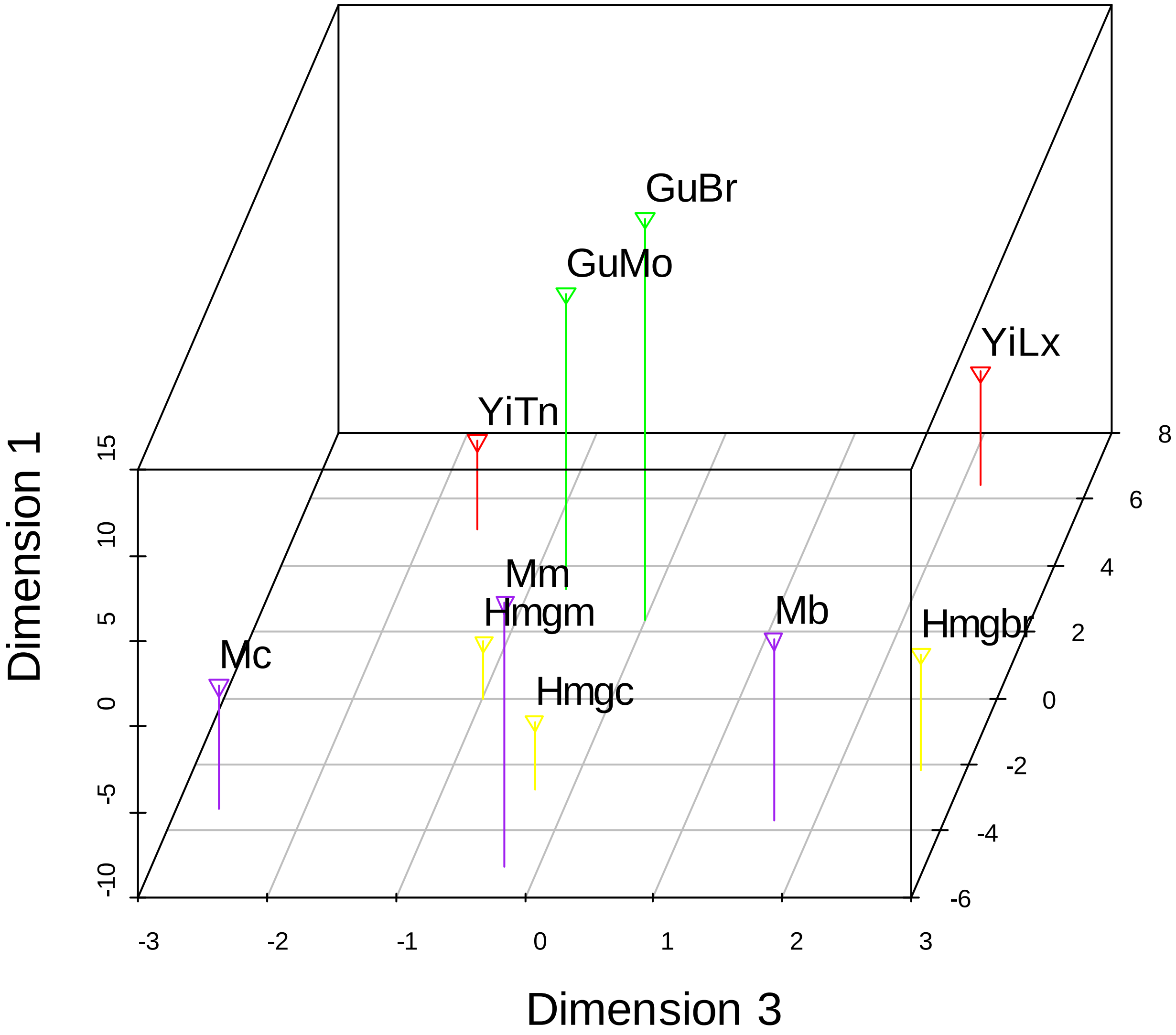
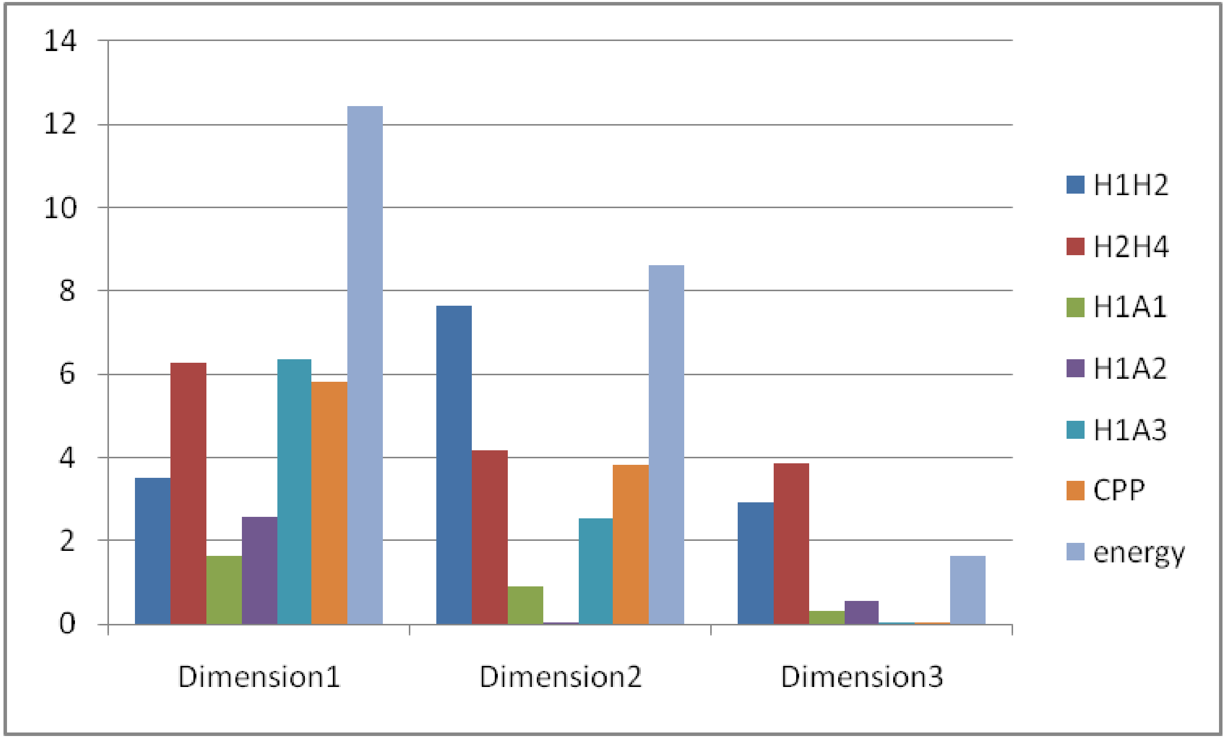


Figure 6. *Weights of seven acoustic measures on each dimension of the 3-D MDS solution.*



Acknowledgments

Thanks to NSF grant BCS-0720304; to Yen Shue for VoiceSauce and Henry Tehrani for EggWorks; and to Sherrie Yang, Joe Ptacek, Victoria Thatte, and the Hmong American Partnership for assistance.

References

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

Bishop, J. & P. Keating (2010) Perception of pitch location within a speaker's own range: fundamental frequency, voice quality and speaker sex, this volume.

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

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

DiCanio, C. (2009) The phonetics of register in Takhian Thong Chong, *Journal of the International Phonetic Association 39,* 162–188.

Esposito, C.M. (2010) The effects of linguistic experience on the perception of phonation, *Journal of Phonetics 38*, 303–316.

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

Garellek, M.& P. Keating (2010) The acoustic consequences of phonation and tone interactions in Jalapa Mazatec, this volume.

Gordon, M. & P. Ladefoged (2001) Phonation types: a cross-linguistic overview, *Journal of Phonetics 29,* 383–406.

Hanson, H. M. (1995) *Glottal characteristics of female speakers,* Ph.D. dissertation, Harvard.

Hanson, H. M., K. N. Stevens, H.-K. Kuo, M. Y. Chen & J. Slifka (2001) Towards models of phonation, *Journal of Phonetic*s 29, 451–480.

Herbst, C. & S. Ternström (2006) A comparison of different methods to measure the EGG contact quotient, *Logopedics Phoniatrics Vocology 31*, 126–138.

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

Holmberg, E. B., R. E. Hillman, J. S. Perkell, P. C. Guiod, & S. L. Goldman (1995) Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice, *Journal of Speech and Hearing Research 38,* 1212–23.

Howard, D. M (1995) Variation of electrolarynglographically derived closed quotient for trained and untrained adult female singers, *Journal of Voice 9,* 163–172.

Iseli, M., Y.-L. Shue & A. Alwan (2007) Age, sex, and vowel dependencies of acoustic measures related to the voice source, *Journal of the Acoustical Society of America 121,* 2283–2295.

Kawahara, H., A. de Cheveign, & R. D. Patterson (1998) An instantaneous-frequency-based pitch extraction method for high quality speech transformation: revised TEMPO in the STRAIGHT-suite, Proc. ICSLP'98, Sydney, Australia, December 1998.

Khan, S. D. (in preparation) Contrastive phonation in Gujarati, Brown University ms.

Kuang, J. (in preparation) Production and perception of the phonation contrast in Yi, Masters thesis, UCLA.

Michaud, A. (2004) A Measurement from Electroglottography: DECPA, and its Application in Prosody, Speech Prosody 2004, Nara, 633–636.

Orlikoff, R. F. (1991) Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects, *Journal of Speech and Hearing Research 34*, 1066–1072.

Rothenberg, M. and J. J. Mahshie (1988) Monitoring vocal fold abduction through vocal fold contact area, *Journal of Speech and Hearing Research 31*, 338–351.

Shue, Y.-L., P. Keating & C. Vicenik (2009) VoiceSauce: A program for voice analysis, *Journal of the Acoustical Society of America 126*, 2221 (abstract).

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