Acoustic and Electroglottographic Analyses of Nonpathological, Nonmodal Phonation

Heriberto Avelino, *Ontario, Canada

Summary: Languages where phonation type and tone are contrastive make use of extremely fine and controlled actions of laryngeal structures; hence, there is little opportunity to variation in either phonation or pitch. Nonetheless, many American Indian languages have contrastive nonmodal phonation, which, moreover, is subject to a great deal of variation. There are a few studies addressing the phonetics of nonmodal phonation in American Indian languages, and little is known about the phonetics/phonology interface of laryngeal features within the sound patterns of these languages. This article aims to contribute to the knowledge of nonmodal phonation through the detailed study of the phenomenon in Yalálag Zapotec (YZ) and American Indian language. A series of spectral and electrophysiological analyses contribute to the description of YZ nonmodal phonation and its variability across gender. It is argued that the temporal patterns in realization of laryngealization are a property of YZ speaker's grammar.

Key Words: Laryngealization–Electroglottography–Nonmodal phonation–Creakiness–Acoustics–American Indian languages.

INTRODUCTION

Peter Ladefoged pointed out some years ago that "one person's voice disorder might be another person's phoneme,"[1,2p351] referring to the fact that, in some languages, nonmodal phonation is part of the phonological system, whereas in others, it is a manifestation of pathological voice quality. Phonetic research accomplished in recent years has increased our knowledge about the complex phenomena of human phonation.[2–4] Likewise, the techniques to study human voice have been greatly diversified and make it accessible to more researchers, especially in the field of voice pathology.[5–7] Nevertheless, there is, still, a hiatus in the basic phonetic description of many languages for which nonmodal phonation is an underlying feature of their pattern of sounds, and little is known about the patterns of phonetic variability of voice quality in these languages. Hence, the goal of this article is to provide an instrumental phonetic account of the phonemic contrast between modal and nonmodal phonation and its variability in the Yalálag Zapotec (YZ) language by using electroglottographic and acoustic analyses.

BACKGROUND

One of the most remarkable features in the Otomanguean family is the use of contrastive laryngeal activity in vowels.[8–10] Descriptive grammatical studies of many Zapotecan languages[11–13] use a variety of labels to describe different vowel types, such as "rearticulated," "glottalized," and "aspirated," for instance. This usage, although not intended to be phonetically technical, reflects the impression of very skilled field linguists, in a way that strongly suggests that the languages of this family might have between two or three contrastive voice types, including modal, breathy, and creaky voices. In addition, Otomanguean languages use contrast in pitch as a form of signaling semantic and morphological differences. The use of contrastive phonation type and tone in a single language requires extremely fine and controlled laryngeal maneuvers, as both dimensions are controlled and implemented by the same anatomical structures, namely the larynx. Previous research of the physiological correlates of these features has revealed that the activity of the arytenoid cartilages is associated primarily with the implementation of phonation, whereas the activity of the thyroid cartilages is associated in first instance with the realization of tone.[3,14–18] YZ is one of the Otomanguean languages in which the contrast of tone and phonation is orthogonal. The YZ language is spoken in Villa Hidalgo, in the Municipality of Villa Alta, Oaxaca, Mexico. According to the Mexican census for the year 2000, there were 2115 people residing in Yalálag.[19] YZ has three tones: high, low, and falling. A number of factors indicate the phonemic status of the three tones, and especially that of the contour tone. First, they enter in contrastive relations, so that there are minimal pairs of tone; second, none can be derived from one of the other two in the lexicon; and third, the three tones are the maximum of tone heights found in the tone system. High and low tones are essentially realized as level, although slight variations (either rising or falling) can occur toward the end. However, these variations are nonphonemic. Falling tone is characterized by a prominent slope that occupies the whole range of the tonal space—that is, it starts at frequencies closer to those of high tone and then falls until reaching the ranges of low tone. Figure 1 illustrates the F_0 contours of a representative triple contrast between high, low, and falling tones in modal syllables. The high and low tones illustrated by /já/ "temazcal" (traditional sweathouse) and /jà/ "bell" have a fairly steady frequency, in contrast with the significant falling trajectory observable in /jâ/ "cane."

Accepted for publication October 2, 2008.

Late Peter Ladefoged supervised the research of my dissertation, (Topics in Yalálag Zapotec with particular reference to its phonetic structures [Ph.D. thesis]. UCLA; 2004), which led to this article. Although physically he is not here anymore, I would like to express my wholehearted gratitude to Peter for his mentorship and guidance. Thanks are also due to Matt Gordon and Pam Munro, who provided invaluable advice and suggestions. I am grateful to Barbara Blankenship, Helen Hanson, Ian Maddieson, John Ohala, Dan Silverman, and Janet Slifka for comments and criticisms to earlier versions of this paper. Thanks to David Freedman for advice on statistical analyses. Thanks to Tim Arbisi-Kelm for copyediting assistance. Thanks to two anonymous reviewers for helpful comments and suggestions. Finally, I would like to thank the speakers who graciously participated in this study, particularly Daria Allende, Ana Daisy Alonso, Jose Bollo Estela Canseco, Agustin Limeta, Margarita Manuel, and Mario Molina. I am solely responsible for the errors and interpretations contained in this article.

From the Department of Linguistics, University of Toronto, Toronto, Ontario, Canada.

Address correspondence and reprint requests to Heriberto Avelino, Department of Linguistics, University of Toronto, 130 St. George Street, Room 6076, Toronto, Ontario M5S 3H1, Canada. E-mail: avelino@berkeley.edu

Journal of Voice, Vol. 24, No. 3, pp. 270-280

0892-1997/$36.00

doi:10.1016/j.jvoice.2008.10.002

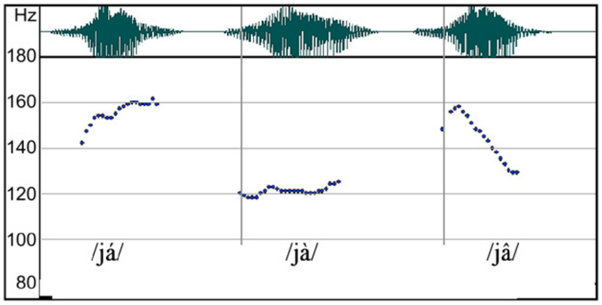

FIGURE 1. Pitch tracks of the words /já/ “sweathouse,” /já/ “bell,” and /já/ “cane” illustrating the contrast between high, low, and falling tone.

In addition, the phonology of YZ makes use of a contrast between laryngealized and modal phonation. The list of words in Table 1. shows canonical examples of the contrast.[a] Figure 2 shows the waveforms and spectrograms of a representative pair of words, /bà/ “tomb” and /bá̰/ “animal,” illustrating the contrast between modal and laryngealized vowels in YZ with words of the same pitch. Overall, the duration of the two vowels is fairly similar (in these particular tokens, 190 and 195 ms for modal and laryngealized, respectively). The spectrograms display the major differences between the two types of vowels. The modal vowel has a constant pulse interval and uniform high distribution of the acoustic energy throughout. In contrast, the laryngealized vowel exhibits a modal spectrum in its first half, followed by an abrupt transition signaling a different glottal activity, which is characterized by irregular, widespread glottal pulses with an overall low amplitude. The two sections of the creaky vowel are indicated in the spectrogram by boundary lines.

TABLE 1.
Contrast Between Modal and Laryngealized Phonation Types in Yalálag Zapotec

Tone	Phonation Type			
	Modal		Laryngealized	
High	/zé/	‘each’	/zḛ́/	‘wall’
	/jín/	‘a chest of clothes’	/jḭ́n/	‘chili’
Low	/bà/	‘tomb’	/bà̰/	‘animal’
	/gà/	‘nine’	/gà̰/	‘basket’
Falling	/bê/	‘echoe’	/bḛ̂/	‘in the morning’
	/zî/	‘far’	/zḭ̂/	‘heavy’

YALÁLAG ZAPOTEC PHONATION CONTINUUM

All spoken languages make use of the laryngeal activity to produce a source of acoustic energy, which is subsequently filtered by actions and configurations of the vocal apparatus.[20,21] The notion of phonation commonly accepted refers to the function of the laryngeal system to transform the airstream into audible sound.[1,3,22–24] Despite this, in principle, simple notion, there are common confusions and misinterpretations originated in part by the multiple and, occasionally, inconsistent terms coming from different research fields (phonetics, speech pathology, engineering, and others) used to describe the mode of vibration of the vocal folds when producing speech sounds.[2,3,23] Thus, for example, labels such as creaky voice, vocal fry, laryngealization, glottalization, and irregular voicing, among others, have been used to refer to the sound produced when the vocal folds are vibrating anteriorly but with the arytenoid cartilages pressed together. Clearly, such a multiplicity of nomenclatures is a major problem for a comprehensive study of human voice. Ladefoged[1,25] has suggested that, to describe the diversity of phonation types across languages, it might be sensible to consider the different phonation types as a continuum whose common denominator is the degree of aperture between the arytenoid cartilages. The continuum ranges from voiceless sounds, where the cartilages are completely separated, through breathy and then modal voice, until reaching the laryngeal setting where the cartilages are completely occluded. A phonemic contrast in the continuum of phonation of YZ is made only between modal and laryngealized phonations; however, the implementation of the underlying laryngeal specification is open to a wide range of allophonic variation, both across speakers and differences based on sex. The spectrograms and waveforms in Figure 3 show representative examples of the various ways of implementing phonemic laryngealized vowels in YZ. The examples illustrate the vowel of the word /ná̰:/ “now” produced by four female speakers. Each panel shows sections of the waveform to the right of the spectrograms that present a detailed view of individual pulses. The figure reveals three different laryngeal settings produced during the course of the vowel. The interval between individual pulses is indicated by arrows. The waveforms are taken from the points indicated

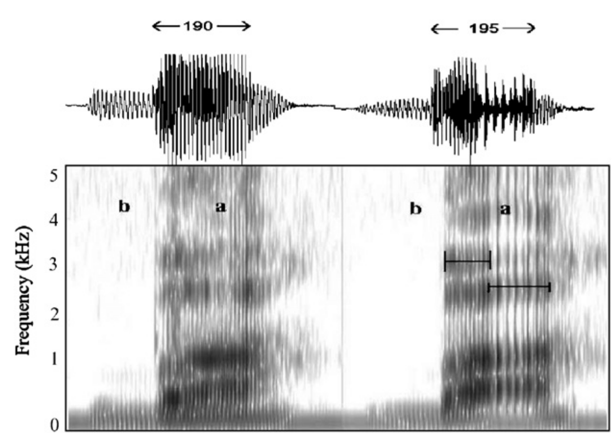

FIGURE 2. Waveforms and spectrograms illustrating the contrast between modal and laryngealized phonation. The laryngealized vowel is divided into two sections.

[a]Throughout the article, I will observe the conventions accepted for the International Phonetic Alphabet (IPA). Accordingly, I will use the standard conventions of transcribing narrow phonetic representations between squares [ ], whereas phonemic representations will be enclosed in diagonals //.

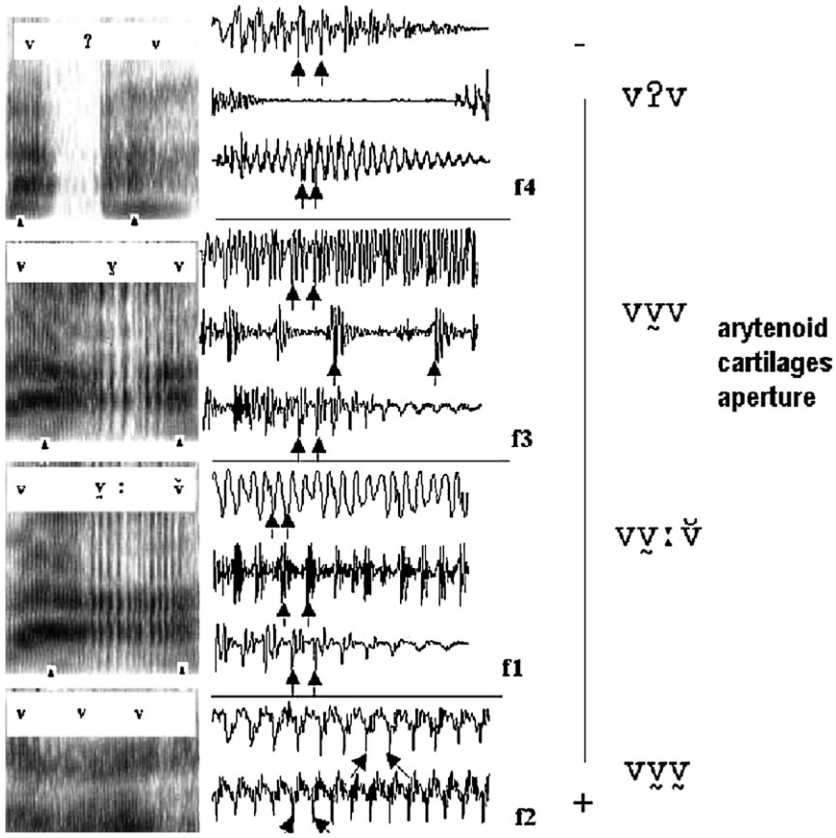

FIGURE 3. Intraspeaker variability of laryngealized voice.

by small arrows in the spectrograms. As is evident in the first panel, corresponding to speaker f4, there are three gestures in what is considered a single phonemic unit: a modal vowel followed by a complete closure of the glottis followed by a vowel of reduced amplitude. This figure represents the archetype of a "rearticulated" vowel, which has been often described in previous grammatical descriptions. The second speaker (f3) shows also three laryngeal settings through the vowel. The first part illustrates a typical modal vowel followed by four aperiodic, long cycles, and then a new segment of sustained periodicity and decreased amplitude pulses. This panel contrasts with the next one, of speaker f1, which exemplifies a modal vowel followed by a longer fragment where the waveform is quite irregular. This section is then followed by a shorter segment of more periodic pulses, indicated as an extra-short vowel in the spectrogram. Finally, the last panel (f2) shows a vowel with three sections of quasi-periodic pulses, which nevertheless give the auditory impression of a single vowel with vocal fry throughout. These examples illustrate one of the most interesting problems found in languages, such as YZ, which make use of phonemic contrasts in voice quality—the great deal of variability in the implementation of a feature that is part of the underlying pattern of sounds.

Previous research has noticed that underlyingly laryngeal segments do not present the typical acoustic properties associated with this phonation. In her analysis of Mazatec and Mpi, Blankenship notices that underlying nonmodal vowels "do not consistently have an audible creak nor display irregular pulses on a spectrogram. Creakiness appears to be an occasional side effect of the laryngealization rather than its goal."[8] For this reason, she uses the term "laryngealized," instead of the more specific "creaky." Based on similar considerations, I will make the same distinction in this article, that is, I reserve the use of the term "creaky" or "creakiness" to refer to the actual phonetic implementation of the phonological category "laryngealized." Thus, the implicational relationship between the two categories entails that, even though all creaky vowels are laryngealized, not all laryngealized vowels are creaky. For instance, all the vowels in Figure 3 are laryngeal, but only the last portions of the vowels for subjects f3 and f1 could be properly labeled as creaky.[b]

In the rest of the article, I present a series of phonetic analyses of modal and nonmodal voice and show the range of variability displayed by these phonation types. First, I present a section dedicated to acoustic data and analyses, and in a subsequent section, I present the electroglottogaphic analyses associated with voice quality in YZ. The article finishes with a discussion of the main findings and concluding remarks.

ACOUSTIC PROPERTIES OF PHONATION IN YALÁLAG ZAPOTEC

Among the phonetic properties distinguishing different phonation types across languages,[3,9,21,26–28] the parameter of spectral

[b]IPA symbols: [ʔ] glottal stop, [:] long segment, [∼] creakiness, [˘] short segment. V is not an IPA symbol; it refers to any vowel.

tilt, “the degree to which intensity drops off as frequency increases,”[9] has been proved to be a reliable indicator of the degree of abruptness or gradualness of vocal fold closure.[21,27] A measure of the relative amplitude of the first two harmonics, H1–H2 has been used as an indicator of the ratio of the time that the folds are open to the duration of a complete cycle of vibration (or open quotient).[21,29,30] Other measures reported in the literature, in addition to H1–H2, include the amplitude difference between the first harmonic and the highest amplitude in the vicinity of the first, second, and third formants; specifically, H1–A1 has been associated with degree of glottal opening, whereas H1–A2 and H1–A3 have been associated with the skewness of the glottal pulse and the ratio of the closing phase. In particular, previous studies have found that the acoustic effects of a configuration of the compressed vocal folds will produce, on one hand, greater amplitudes of the spectrum at high frequencies compared with that of modal voice, and, on the other, smaller amplitudes at low frequencies.[4,21] Hence, it is expected that in creaky phonation, the energy of the second harmonic will be higher than that of the first harmonic compared with the spectrum of modal phonation, where the magnitude of the first harmonic is higher than that of the second.

METHOD

Data acquisition and speakers

Six adults (three female, three male) in their mid-30s and 50s were recorded to obtain acoustic data. The consultants are bilingual Zapotec-Spanish; however, their native, first and dominant language is YZ. All the consultants are bilingual Zapotec-Spanish. The consultants were born and raised in Yalálag, and use their native language on a daily basis. Consultants were instructed to read and repeat the list of four monosyllabic words shown in Table 2 that illustrated the contrast among the two phonation types, modal and laryngealized in an open vowel [a] and [a0]. Five repetitions of each word were recorded to avoid prosodic effects, and the first and last words were not considered in the analyses. Utterances were recorded on an analog audiotape. To obtain as much of the acoustic signal as possible from the glottal source, a microphone (Shure SP19L, cardioid, Shure Inc., Niles, IL) was set at a distance of 10 cm in front of the speaker. Words were pronounced in the carrier sentence *cho wne __ gaye* “let us say __ five times”. The sentence does not include nonmodal phonation that may affect the target word. Tokens were digitized at a sampling rate of 22 050 Hz. Analyses were made with *PcQuirer* software.[31] A fast Fourier transform (FFT) was calculated over a 26-ms window at three points within a vowel: 25%, 50%, and 75% from the beginning of the vowel.

TABLE 2.
Word-List Illustrating the Contrast Between Modal and Laryngealized Phonation

Modal		Laryngealized	
gà	‘nine’	gà̰	‘mat’
nà	‘and’	ŋà̰	‘now’

Data analysis

The spectral parameters considered in the present study are illustrated in Figure 4. Identification of H1, H2, A1, and A3 was made by locating the frequencies of interest on an linear predictive coding (LPC) overlaid on the FFT plot. Occasionally, the measures yielded conspicuous outliers. These cases were inspected individually, remeasured manually, and corrected or excluded if the tokens had acoustic properties, which made them unmeasurable. The measures used to investigate the acoustic correlates of phonation type in YZ were the relative magnitudes obtained from the difference between H1–H2, H1–A1, and H1–A3. As discussed earlier, these measures have been used in previous analyses of languages where nonmodal phonation is contrastive.[8–10,22,32]

Results

Figure 5 shows the results of the H1–H2, H1–A1, and H1–A3 measures for the two types of phonation. The mean values of all three time points during all three iterations for all six speakers of the three measures are reported. The results show a clear difference between laryngealized and modal vowels. The difference between H1–H2 laryngealized vowels (1.34 dB) contrasts with that shown by modal vowels, which have a more steeply positive difference of 4.89 dB. This pattern is consistent with previous findings, demonstrating that the magnitude of creaky phonation is small or negative, whereas both modal and breathy phonation show a greater difference magnitude.[27] The results of H1–A1 indicate an increasing in the amplitude of A1: 0.49 dB for modal vowels, and -5.46 dB for laryngealized ones. Finally, the spectral tilt (positive) found in the H1–A3 measure shows the expected greater magnitude for modal phonation (10.28 dB) than that of laryngealized vowels (3.39 dB), as the gradual adduction of the vocal folds would excite frequencies close to F_0. The results confirm the pattern anticipated by Stevens: as the vocal folds vibrate abruptly when they are compressed together, they produce excitation of high frequencies. The difference between phonation types was highly significant with $P < 0.0001$ in all the three measures.

Further inspection of the data showed important intraspeaker differences in phonation type. The results of the three measures for each subject are summarized in the corresponding panels in Figure 6. As conspicuous from the figure, the speech of females and males in the three measures observed is, in general, well differentiated. The main trend is a steeper positive spectral slope for female voices, whereas men show the opposite trend toward a steeper negative spectral tilt. These results suggest that, in YZ, female subjects trend toward producing more modal voice values than do male subjects. Overall, the results describe a continuum with canonical values of modal phonation at one end and creaky phonation at the other end. The results for each subject fit evenly along the continuum, with subject f1 at the modal end of the continuum and subject m1 at the opposite extreme of creaky voice. These findings are consistent with

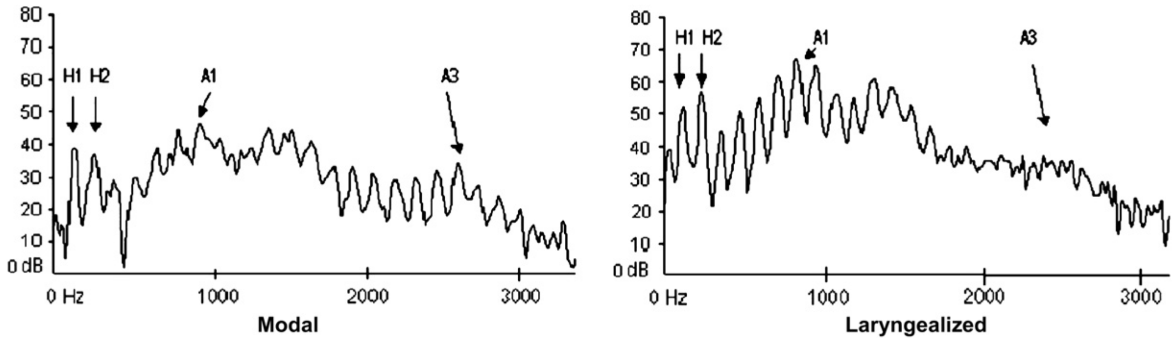

FIGURE 4. Source spectrum of modal and laryngealized vowels. The acoustic parameters indicated are the amplitudes of the first harmonic (H1), second harmonic (H2), first formant (A1), and third formant (A3).

previous studies showing that phonation is subject to a great deal of intraspeaker variation,[27,33] and might be comparable with other reports in which women's voices tend to be more breathy.[34] It is noteworthy to mention that the results obtained in the present study are comparable to the report of another Zapotec language by Gordon and Ladefoged,[9] in which, females had breathier vowels than male speakers.

The results of H1–H2 shown in Figure 6 indicate that females and males implement laryngealized phonation differently. In an opposite direction to females, the males consistently had a prominent peak amplitude in H2 than H1; hence, the values were negative. A number of inferences about the configuration of the subject's vocal folds can be drawn based on the studies of Hanson et al and Stevens and Hanson[27,28] looking at the correlation of spectral tilt measurements with the extent of glottal opening: According to these studies, presumably, the results of 12 and 8.7 dB found for females and males in the modal voice, respectively, correspond roughly to 50% and 35% of the opening phase—that is, the open phase of modal vowels in females is relatively greater than that of males. The results of H1–A1 indicate that there was a prominent first formant peak for most speakers, so that the magnitude of the differences was, in general, negative. These results replicate the findings of Gordon and Ladefoged,[9] who report that, for two speakers of a different Zapotec language, A1 is greater than H1. The results concerning the H1–A3 measurement show a wide variation of spectral tilt among speakers, with greater ranges for females than males in modal voicing, in contrast with the greatest ranges for males in laryngealized vowels. For this measure as well, modal vowels are reliably associated across speakers with a greater positive slope. The results for modal vowels confirm the same relation frequently mentioned in the literature: in modal phonation, the energy is concentrated in low frequencies; hence, H1 obtains greater amplitudes than the peaks of higher frequencies. Although the results indicate a trend toward a reliable differentiation of phonation types, the statistical tests showed that not all the measures are equally suited to predict a difference. The results of a series of analysis of variance (ANOVA) in Table 3 show that, for males, the difference in phonation was significant at all the three measures, whereas for females, the only significant measure was H1–A3.

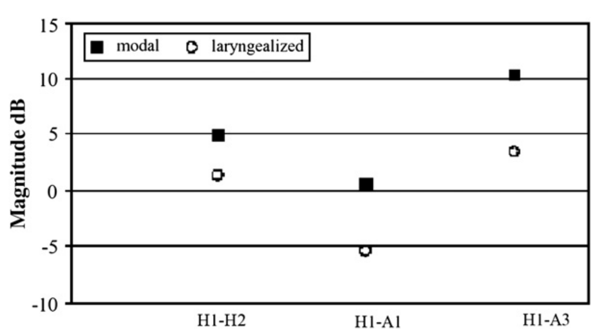

FIGURE 5. Results of the measures H1–H2, H1–A1, and H1–A3, for modal and laryngealized vowels.

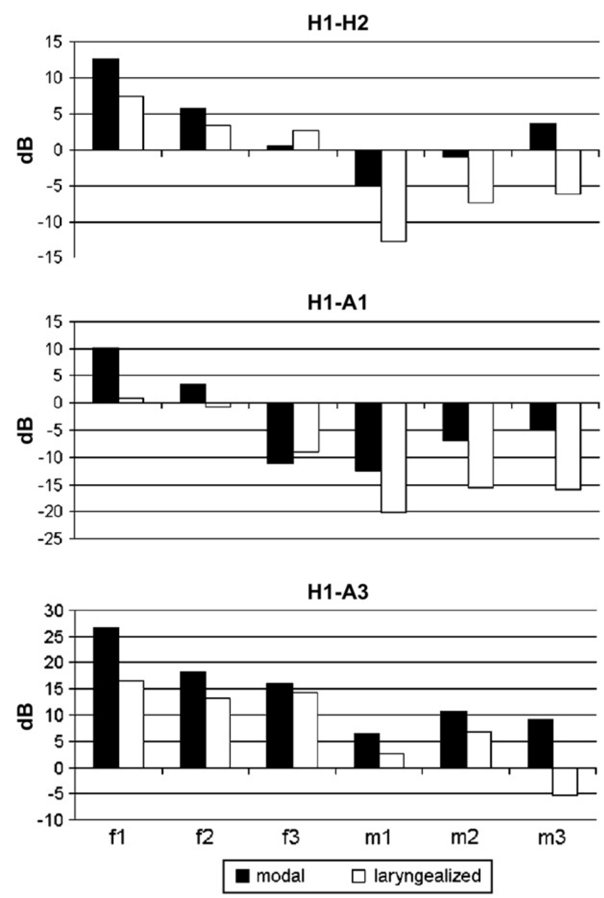

FIGURE 6. Interspeaker variability of modal and laryngealized phonation according to H1–H2, H1–A1, and H1–A3 parameters.

TABLE 3.
Differences Among Measures H1–H2, H1–A1, and H1–A3 as Indicators of Phonation Type in Female and Male Speakers

Measure	Females: *P* value	Males: *P* value
H1–H2	Nonsignificant	0.0001
H1–A1	Nonsignificant	0.0001
H1–A3	0.0001	0.0001

Other patterns of languages where nonmodal phonation is contrastive in vowels have been found in recent research. For example, the laryngeal setting correlated with nonmodal phonation lasts longer than in languages where nonmodal phonation is nonphonemic, and has acoustic properties that clearly differentiate them from modal voice. This pattern has been reported for some languages of the same family to which YZ belongs, notably in Blankenship.[8] Based on these findings, the spectral parameters already discussed were measured at three points in the vowel (25%, 50%, and 75% from the beginning) to see whether the phonation properties are sustained through the entire duration of the vowel or just confined to certain portions of it. Figure 7 summarizes the results of the three measures. The results show that at the beginning of the vowel, both phonation types are nearly identical because of the fact that laryngealized vowels start with a modal onset configuration. From this point onward, modal vowels follow a constant increase in amplitude throughout the duration of the vowel, whereas laryngealized vowels show a consistent dip in the middle portion of the vowel. Finally, both vowel types have an increasing magnitude trajectory toward the end. These patterns are consistent for all three measurements. These results were tested with ANOVAs summarized in Table 4.

Interim summary

Overall, the results showed that modal and laryngealized vowels are well differentiated by the acoustic measurements investigated. The results are consistent with previous research reporting a relative dominance of the low frequencies in the spectra of modal vowels as compared with laryngealized ones.[21,26,34] The results showed a tendency for all the subjects to make important differences in the production of both phonation types, despite the fact that not all the measurements were

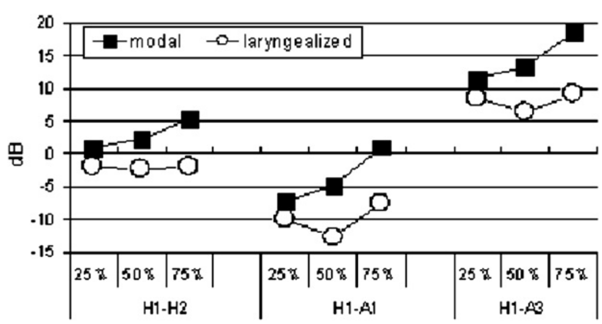

FIGURE 7. Intraspeaker variablility of modal and laryngealized phonation according to H1–H2, H1–A1 and H1–A3 parameters at three vowel points.

statistically significant for all female speakers. A parameter that reliably distinguished both phonation types for females and males was H1–A3—a result that is consistent with the suggestion of Hanson et al.[27] The H1–A3 measurement has been associated with the ratio of the duration of the closed phase to the duration of a complete glottal cycle; hence, it provides a close characterization of the prominent adduction of the vocal folds entailed by creaky voice.[3,35] The results showed an important variation across genders. Although both types of phonation were distinguished, females had a trend to more modal phonation than males. This result is consistent with previous research showing the differences in phonation between females and males.[27,33] Regarding the time course of phonation along the vowel, the results showed that modal and laryngealized vowels are different only in the middle and end portions of the vowel. In particular, the two vowels revealed opposite trajectories in the middle of the vowel. This might be interpreted as a gesture to maximally distinguish the modal and laryngeal phonation.

ELECTROGLOTTOGRAPHIC ANALYSIS

The results of the acoustic analyses obtained in the previous sections allowed us to make a number of remarks about the glottal settings of modal and laryngealized vowels in YZ. This section deals with data obtained with electroglottography (EGG). EGG is a noninvasive electrophysiological technique that allows observation of the properties of the vocal folds in vibration by measuring the electrical resistance/conductance between two electrodes positioned on the neck approximately over the thyroid cartilages. Some studies that offer a comprehensive account of EGG technique and methodology include,[4,30,35–38] among others. As the accuracy of the correlation between the output glottal waveform and the vocal fold activity has been confirmed by independent techniques that permit direct visual examination of the larynx,[6,38–40] reliable information about the degree of vocal fold contact area in YZ phonation can be drawn from EGG data.

METHOD

Data acquisition and speakers

EGG data were recorded from two subjects, one female and one male, using a portable electro-laryngograph processor (Laryngograph Ltd., London, UK) connected to a transducer box of the PCQuirer X16 multi-channel data acquisition system.[31] The EGG signal and the simultaneous acoustic recordings were sampled at 11 kHz. The EGG signal and the simultaneous recording of the acoustic signal were recorded to separate channels. The acoustic recording setting was similar to that described in the previous section. The distance from the microphone to the subject's lips was set at 10 cm. The gold-plated electrodes of the laryngograph were held to either side of the subject's thyroid cartilage and held stable by holding a velcro strap around the subject's neck. A test of the signal was obtained until it was confirmed that the location of the electrodes was adequate. Throughout the recording session, the electrodes were relocated when the signal did not appear to be reliable on the system; in a few occasions, the

TABLE 4.
Mean and Probability Values from ANOVA Tests at Three Points in the Vowel

Measure	25%		50%		75%	
	Modal	Laryngealized	Modal	Laryngealized	Modal	Laryngealized
H1–H2	0.86	1.92	2.33	2.42	5.03	1.97
	$P = 0.132$ (2.326)		$P = 0.016$ (6.139)		$P = 0.003$ (9.602)	
H1–A1	−7.25	9.97	−4.78	12.81	0.92	7.47
	$P = 0.209$ (1.609)		$P = 0.001$ (10.9556)		$P = 0.002$ (9.839)	
H1–A3	11.50	8.39	13.50	6.39	18.42	9.25
	$P = 0.120$ (2.474)		$P = 0.003$ (9.302)		$P = 0.0001$ (16.106)	

Mean values expressed in dB. *F* values are given in parentheses after the probability values. n = 72 for every comparison. DF is 1 for all the three groups.

electrodes were cleaned of sweat. The subjects repeated between four and eight times each of the words selected that illustrated the contrast between modal and creaky phonation (Table 5).

Data analysis

Figure 8 shows the glottal waveform and the landmarks that were used to analyze the data. The relevant sections and points (adapted from Henrich et al;[41] see also Niimi and Miyaji[42]) are: T_0, the pitch period; t_1, beginning of the adduction of the vocal folds and offset of the airflow; t_2, instant in time of maximum vocal fold contact and concomitant minimum glottal flow through the glottis; t_3, moment of glottal opening and moment of the minimum change of glottal flow—some parts of the vocal folds might still make contact so that the speed of the opening movement decreases; t_4, instant of complete glottal opening, that is, the glottis is fully open and total airflow occurs. There is a progressively decreasing airflow in the interval between t_1 and t_2, and an increasing airflow between t_3 and t_4.

There are two measures well known in the literature to be reliable indicators of vocal folds activity: open quotient (OQ) and speed quotient (SQ). The two parameters have been defined as early as the studies of Timcke,[43] in which, high-speed film recordings of the vocal folds supported a definition of OQ as the ratio of the duration of the open phase to the duration of a complete glottal cycle, and SQ as the open phase divided by the duration of the closing phase. However, recent studies have pointed out that the EGG signal is better at accounting for the closed phase than for the open phase.[4,44] Therefore, this study reports a close quotient (CQ) measure, defined as the ratio of the time in which the vocal folds are in contact during a complete glottal cycle. As Figure 8 illustrates, the current flow between the electrodes increases as a function of a greater vocal fold contact and decreases with lesser contact. Orlikoff[19] and Davies et al[45] give an almost equivalent definition of the end of the contact phase (and beginning of the opening phase) as the point where the glottal waveform reaches 3/7 and 25% of the amplitude for each glottal cycle, respectively. In this study, the threshold to calculate the vocal fold contact phase was defined at 25% of the EGG waveform amplitude. Measurements were made 10% before and after the beginning and at the end of whole vowel acoustic duration to avoid consonantal and word-final effects. In addition, separate measurements were made of two halves of the vowel. Figure 9 illustrates audio and glottal waveforms of a modal and a laryngealized (realized in fact as creaky) vowels in YZ with the words /ba/ “tomb” and /ba0/ “animal” (closing phase is shown downward).

A number of remarks can be made from these canonical examples. First, the pulses of the EGG modal waveform are very regular and its shape is almost sinusoidal, with a slight tendency

TABLE 5.
Words Illustrating the Contrast Between Modal and Laryngealized Phonation

Modal		Laryngeailzed	
la	‘hot’	la̰	‘smells’
ba	‘tomb’	ba̰	‘animal’
ga	‘nine’	ga̰	‘mat’
na	‘and’	/ná̰:/	‘now’

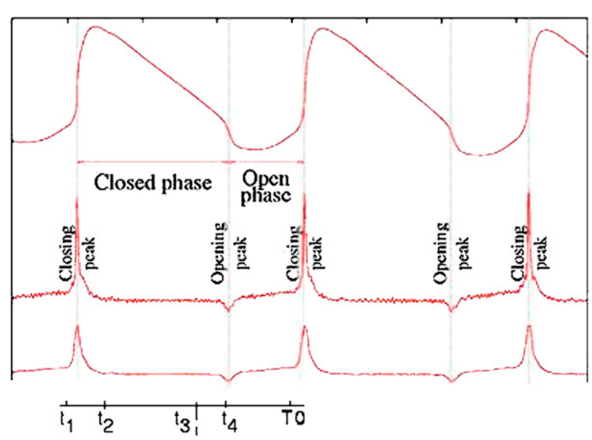

FIGURE 8. Schematic representation of the glottal waveform (adapted from Marasek[4], after Stevens and Hanson[28]).

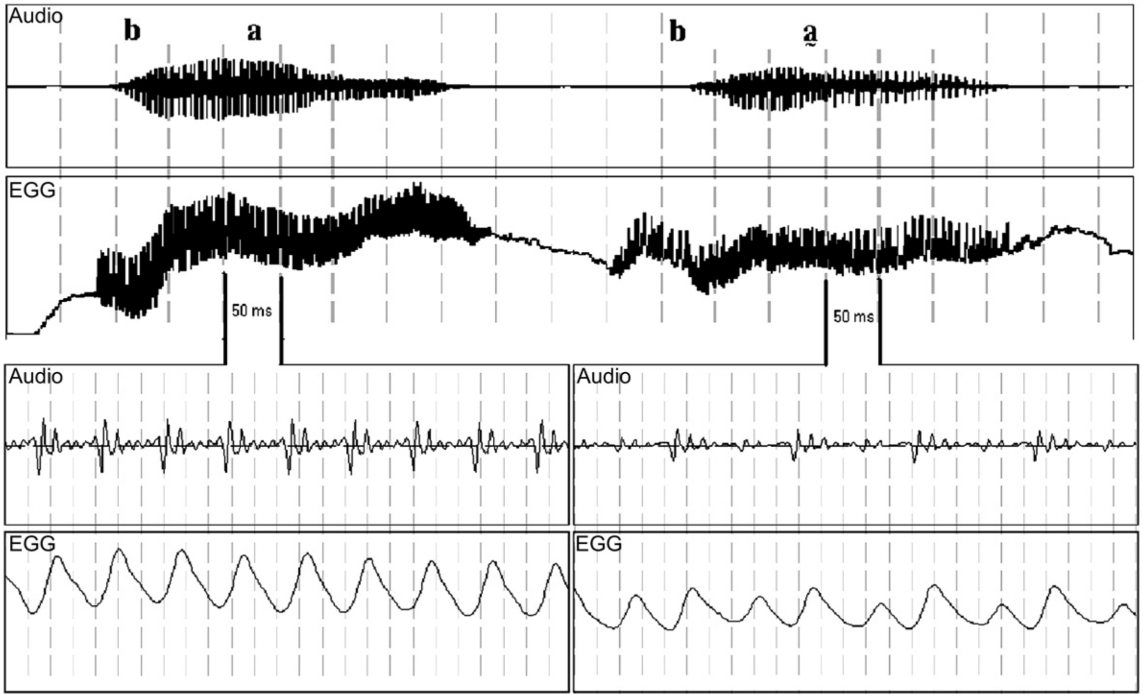

FIGURE 9. Audio and EGG waveforms of modal laryngealized vowels. An expansion of 50 ms from the middle of the two signals is shown in the lower panel. Female speaker: at the bottom shows an expansion of 50 ms.

to a skewness to the left, indicating a longer time in fall as compared with the rise. The closing section is moderately short, and the interval between maxima peak-to-peak amplitude is relatively high. The modal EGG waveform in the figure is a good instance of the commonly reported duty ratio of the OQ nearing 50%. In contrast, the EGG waveform of the creaky vowel has characteristic irregular pulses throughout. The figure illustrates the typical triangular shape with a rounded vertex, and a skewness to the left observed in other descriptions of creaky voice.[3] A cycle of the creaky waveform is almost double that of the modal (five pulses vs eight pulses in 50 ms). The beginning of the closed phase is very short and rises fast; hence, the slope increases in an angle close to 90° until reaching the maximum point of glottal closure. The complete closure is rather short, but the closed phase lasts for a great amount of the period. It is also noticeably longer than the overall opening, which, in turn, is very short, and has an additional partial closure peak, indicating that the abduction is incomplete. This instance is representative of the type of dicrotic voice characteristic of the laryngealized voice observed in YZ speakers. Not surprisingly, often the automatic pitch tracker (based on an correlation algorithm [see Talkin[46]]) rendered "halved" pitches for laryngealized vowels.

Results

Figure 10 displays the mean values obtained from the two sections, early and late, of the closed quotient parameter for the two types of phonation, modal and laryngealized. The results showed a clear differentiation between the two types of phonation. Overall, laryngealized vowels had a greater closed quotient than modal vowels, 64% versus 59%, respectively. A one-way ANOVA confirmed the reliability of these results ($F(1, 1794) = 114.365$, $P = 0.0001$). Similar results were obtained considering the differences of phonation type by sex. The mean value of the closed quotient for the female was 60.8% in modal vowels and 64.1% in laryngealized vowels, whereas for the male, the mean was 55.6% in modal vowels and 64.9% in laryngealized vowels. The results of a one-way ANOVA confirmed the consistency of the results: $F(1, 1238) = 37.524$, $P = 0.0001$ for females, and $F(1, 554) = 127.70$, $P = 0.0001$ for males.

Although the results showed consistent differences in the phonation of the two speakers, a significant difference between male and female vowels was only found in modal voice ($F(1, 960) = 73.304$, $P = 0.0001$). Nonetheless, it is interesting to note that the male subject showed more extreme pronunciations

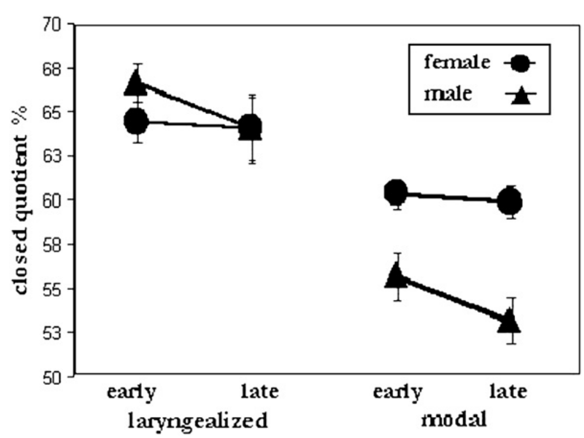

FIGURE 10. Closed quotient percentage of modal and creaky vowels divided into two sections and classified by sex.

of the two types of vowels, that is, his laryngealized vowels had a greater ratio of closing than those of the female, and his modal vowels had a lesser proportion of closing that the female's modal vowels.

The figure shows an asymmetrical relationship between the two vowel sections, in which, the earlier one had a greater closed quotient than the later one. This tendency was consistent for the two subjects in both types of phonation.

DISCUSSION AND CONCLUDING REMARKS

The acoustic and electroglottographic analyses presented in this study have shown reliable phonetic correlates of modal and nonmodal phonation in YZ, a language in which such a contrast is exploited to signal differences in the meaning of words. Overall, the analyses revealed consistent cues in the spectra and in the properties of the glottal waveform. The data supports the observations made in previous research on pathological voice, which have reported a great and abrupt adductive tension together with medial compression in creaky phonation, as inferred from the glottal wave.[47–50] In principle, the properties of pathological voice and the corresponding EGG waveform can be caused by quite different physiological conditions.[51] For this reason, a direct comparison of creaky voice with the irregular phonations encountered in pathological voice may be misleading.[c] With this caveat in mind, the data analyzed here show a fundamental difference between YZ nonmodal phonation and the general properties of pathological creaky phonation. YZ speakers implement laryngealized vowels in a three-way phasing: an initial modal section followed by a section of creakiness and a final modal section. In contrast, creaky voice in pathological speech is uncontrolled and often interspersed with random alternations of modal and nonmodal voice. This indicates that the pattern of nonmodal phonation in YZ is part of the sound system of the language and could not be attributed to chance. Moreover, these findings are consistent with recent literature addressing the issue of the timing of nonmodal phonation. Blankenship[8] and Gordon and Ladefoged[9] have observed that nonmodal properties often do not persist through the entire vowel. Avelino et al[52] present data from Yucatec Maya that follows the same trend. They describe three patterns in the implementation of what constitutes a single laryngealized vowel: (1) a full glottal stop between two portions of modal phonation; (2) a modal vowel interrupted by a period of creaky voice; and (3) an initial modal voice portion that shifts to creaky voice toward the end of the vowel. According to these studies, the separation of modal and nonmodal phonation may reflect conflicting perceptual demands. Along these lines, Silverman et al[53] have suggested that the sequential realization of phonological features, which could obscure each other in perception if implemented simultaneously, is a strategy that ensures the perceptual recoverability of the features in question. It appears that the timing of nonmodal phonation in YZ is organized to guarantee the production and perception of multiple phonemic features that could otherwise contradict each other in actual implementation: phonation and tone. Creakiness is characterized by a low frequency and irregular vibration, properties that conflict with a stable pitch, especially at high frequencies. It is frequently found in the literature that contrastive voice qualities are associated with different tone patterns;[54–55] nonetheless, prior investigations have shown that control of tone and phonation are independent.[24] In fact, although YZ laryngealized phonation often is associated with low tone, the co-occurrence of laryngealized vowels with high tone is also part of the lexical possibilities of the language, thus demonstrating that, in YZ, phonation and pitch are independently controllable variables, a possibility that is not frequently documented.

The results indicate that there are reliable differences of phonation types based on sex of the speakers, with males producing creaky phonation more than females. This strongly suggests that, regardless of the contrastive use of creakiness, extralinguistic considerations (mainly anatomical, eg, Harrison[56]) play a role in the actual production in both modal and nonmodal phonation, although they might render the same perceptual voice quality.[2,57,58] Previous studies have found consistent differences of phonation based on the speaker's sex in languages where nonmodal phonation is noncontrastive.[59,60] Some studies report higher rates of glottalization in females than males in languages, such as English and Swedish.[61,62] However, others[33] found divided evidence showing that, in English, female professional speakers glottalize more often than males, but for a group of nonprofessional speakers, the males glottalized more than the females did. The relevance of the variation of voice quality found across gender in a language in which the contrast between modal and nonmodal phonation is phonemic, as in YZ, is that, because the implementation of creakiness is part of the speaker's knowledge of the grammar, it is, thus, consciously controlled.

The findings about the YZ patterns of phonation can inform us about the nature and manner of the mechanisms used in the articulation of voice. These patterns of variation can be compared with the variation of voice found in patients of languages where nonmodal phonation does not signal phonological contrasts. The results reported in this study can be summarized in a paraphrasis of Ladefoged's maxim by saying that, although the phonetic patterns of pathological and normal laryngealization are dissimilar, one person's voice disorder might be cured by looking at another person's phoneme. One hopes that the knowledge gained from the patterns found in languages, such as YZ, might contribute not only to our theoretical knowledge and the typology of nonmodal phonation across languages, but also may shed light on the differences in the production of pathological voice.

[c]Thanks to an anonymous reviewer for pointing this out.

REFERENCES

1. Ladefoged P. The linguistic use of different phonation types. In: Bless DM, Abbs JH, eds. *Vocal Fold Physiology: Contemporary Research and Clinical Issues*. San Diego: College-Hill Press; 1983:351-360.
2. Gerratt B, Kreiman J. Toward a taxonomy of non-modal phonation. *J Phon*. 2001;29:365-381.

3. Laver J. *The Phonetic Description of Voice Quality*. In: *Series Cambridge Studies in Linguistics*, Vol. 31. Cambridge: Cambridge University Press; 1980.
4. Marasek K. Electroglottographic description of voice quality. [Habilitationsschrift. Institute für Maschinelle Sprachverarbeitung]. Doctoral dissertation. University of Stuttgart; 1997.
5. Hess M, Ludwigs M. Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration. *J Voice*. 2000;14:255-271.
6. Fukuda H, Kawaida M, Kohno N. New visualization technique with a three-dimensional video-assisted stereoendoscopic system: application of the BVHIS display method during endolaryngeal surgery. *J Voice*. 2002;16:105-116.
7. Motta G, Cesari U, Iengo M, Motta G Jr. Clinical application of electroglottography. *Folia Phoniatr.* 1990;42:111-117.
8. Blankenship B. The timing of nonmodal phonation in vowels. *J Phon*. 2002;30:163-191.
9. Gordon M, Ladefoged P. Phonation types: a cross-linguistic overview. *J Phonet*. 2001;29:383-406.
10. Kirk PL, Ladefoged J, Ladefoged P. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In: Mattina A, Montler T. eds. American Indian Linguistics and Ethnography in Honor of Laurence C. Thompson, Missoula, MT: University of Montana Press.
11. Long RC, Butler IH. Gramática zapoteca. In: Long RC, Cruz SM, eds. *Diccionario Zapoteco de San Bartolomé Zoogocho, Oaxaca*. México, D.F.: Instituto Lingüístico de Verano; 1999.
12. Lyman L, Lyman R. Choapan Zapotec phonology. In: Merrifield WR, ed. *Studies in Otomanguean Phonology.* Summer Institute of Linguistics/University of Texas at Arlington; 1977:137-161.
13. Newberg, R. Yalálag Zapotec Phonology. Manuscript; 1983. 20pp.
14. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr.* 1970;28:1-20.
15. Hirano M, Vennard W, Ohala JJ. Regulation of register, pitch and intensity of voice. *Folia Phoniatr.* 1976;28:89-94.
16. Hirose H. Investigating the physiology of laryngeal structures. In: Hardcastle WJ, Laver J, eds. *The Handbook of Phonetic Sciences*. Cambridge: Blackwell Publishing; 1997:116-136.
17. Ohala JJ. The production of tone. In: Fromkin VA, ed. *Tone: A Linguistic Survey.* New York: Academic Press; 1978:5-39.
18. Orlikoff RF. Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. *J Speech Hear Res*. 1991;34:1066-1072.
19. Instituto Nacional de Estadística. Geografía e Informática (Mexico) Censo *General de Población y Vivienda* 2000. México; 2001.
20. Fant G. *Acoustic Theory of Speech Production*. The Hague, Netherlands: Mouton; 1960.
21. Stevens KN. *Acoustic Phonetics: Current Studies in Linguistics*. Cambridge, Massachusets/London, England: The MIT Press; 1998.
22. Ladefoged P, Maddieson I, Jackson M. Investigating phonation types in different languages. In: Fujimura O, ed. *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven; 1988:297-317.
23. Ladefoged P, Maddieson I. *The Sounds of the World's Languages*. Blackwell; 1996.
24. Ní Chasaide A, Gobl C. Voice source variation. In: Hardcastle WJ, Laver J, eds. *The Handbook of Phonetic Sciences*. Cambridge: Blackwell Publishing; 1997:427-461.
25. Ladefoged P. *Preliminaries to Linguistic Phonetics*. Chicago, IL: University of Chicago Press; 1971.
26. Ananthapadmanabha TV. Acoustic factors determining perceived voice quality. In: Fujimura O, Hirano M, eds. *Vocal Fold Physiology. Voice Quality Control*. San Diego: Singular Publishing Group. Inc; 1995:113-126.
27. Hanson HM, Stevens KN, Je HK, Chen MY, Slifka J. Towards models of phonation. *J Phon*. 2001;29:451-480.
28. Stevens KN, Hanson HM. Classification of glottal vibration from acoustic measurements. In: Fujimura O, Hirano M, eds. *Vocal Fold Physiology. Voice Quality Control*. San Diego: Singular Publishing Group. Inc; 1995.
29. Holmberg EB, Hillman RE, Perkell JS, Guiod PC, Goldman SL. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *J Speech Hear Res*. 1995;38:1212-1223.
30. Scherer RC, Druker DG, Titze IR. Electroglottography and direct measurement of vocal fold contact area. In: Fujimura O, ed. *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York: Raven; 1988: 279-291.
31. Tehrani H. PcQuirerX. Scicon R&D Inc., 2007.
32. 1985 Maddieson I, Ladefoged P. Tense" and "lax" in four minority languages of China. *J Phon*. 1985;13:433-454.
33. Redi L, Shattuck-Hufnagel S. Variation in realization of glottalization in normal speakers. *J Phon*. 2001;29:407-429.
34. Hanson HM. Glottal characteristics of female speakers: acoustic correlates. *J Acoust Soc Am*. 1997;30:267-274.
35. Childers DG, Hicks DM, Moore GP, Eskenazi L, Lalwani AL. Electroglottography and vocal fold physiology. *J. Speech Hear Res*. 1990;33: 245-254.
36. Fourcin AJ. Electrolaryngographic assessment of phonatory function. *J Phon*. 1987;14:435-442.
37. Hong KH. Electroglottography and laryngeal articulation in speech. *Folia Phoniatr.* 1997;49:225-233.
38. Baken RJ, Orlikoff RF. Correlates of vocal fold motion: electroglottography. In: Baken RJ, ed. *Clinical Measurement of Speech and Voice*. New York: Singular, Thompson Learning; 2000:413-427.
39. Baer T, Löfqvist A, McGarr NS. Laryngeal vibrations: a comparison between highspeed filming and glottographic techniques. *J Acoust Soc Am*. 1983;73:1304-1308.
40. Hertegard S, Gauffin J. Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography. *J Speech Hear Res.*. 1995;38:85-100.
41. Henrich N, d'Alessandro C, Doval B, Castellengo M. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am*. 2004;115:1312-1332.
42. Niimi S, Miyaji M. Vocal fold vibration and voice quality. *Folia Phoniatr.* 2000;52:1-3.
43. Timcke R, Von Leden H, Moore P. Laryngeal vibrations: measurements of the glottic wave. I. The normal vibratory cycle. *Arch Otolaryngol*. 1958;68: 1-19.
44. Alaska YA. Approximations of open quotient and speed quotient from glottal airflow and egg waveforms: effects of measurement criteria and sound pressure level. *J Voice*. 1998;12:1.
45. Davies P, Lindsey GA, Fuller H, Fourcin AJ. Variation in glottal open and closed phase for speakers of English. *Proc. Inst Acoust*. 1986;8: 539-546.
46. Talkin D. A robust algorithm for pitch tracking (RAPT). In: Kleijn WB, Paliwal KK, eds. *From Speech Coding and Synthesis*. Elsevier: Amsterdam, the Netherlands; 1995:495-518.
47. Dejonckere PH, Lebacq J. Electroglottography and vocal nodules. An attempt to quantify the shape of the signal. *Folia Phoniatr.* 1985;37:195-200.
48. Guidet C, Chevrie-Muller C. Computer analysis of prosodic and electroglottographic parameters in diagnosis of pathologic voice. In: Winkler P, ed. *Investigations of the Speech Process, Quantitative Linguistics*. Bochum: Studienverlag Dr. N. Brockmeyer; 1979:19.
49. Hanson D, Gerratt BR, Watd PH. Glottographic measurement of vocal dysfunction. A preliminary report. *Ann Otol Rhinol Laryngol*. 1983;92:413-420.
50. Hanson DG, Gerratt BR, Karin RR, Berke GS. Glottographic measures of vocal fold vibration: an examination of laryngeal paralysis. *Laryngoscope*. 1988;98:541-549.
51. Abberton E, Fourcin A. Electrolaryngography. In: Code C, Ball M, eds. *Experimental Clinical Phonetics*. London: Croom Helm; 1984:62-78.
52. Avelino H, Tilsen S, Kim E. The phonetics of laryngealization in Yucatec Maya. In: Paper presented at Linguistic Society of America. Anaheim. Annual Meeting, Jan 4–7, 2007. Anaheim, CA.
53. Silverman D, Blankenship B, Kirk P, Ladefoged P. Phonetic structures in Jalapa Mazatec. *Anthropol Linguist*. 1995;37:70-88.
54. Huffman MK. Measures of phonation type in Hmong. *J Acoust Soc Am*. 1987;81:495-504.
55. Maddieson I, Hess S. The effect on F0 of the linguistic use of phonation type. *UCLA Working Papers Phon*. 1987;67:112-118.
56. Harrison DFN. *The Anatomy and Physiology of the Mammalian Larynx*. Cambridge University Press; 1995.

57. Blomgren M, Chen Y, Ng ML, Gilbert HR. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J Acoust Soc Am*. 1998;103:2649-2658.
58. Gelfer ME, Bultemeyer DK. Evaluation of vocal fold vibratory patterns in normal voices. *J Voice*. 1990;4:335-345.
59. Sodersten M, Lindestad PA. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *J Speech Hear Res*. 1990;33:1-11.
60. Sodersten M, Hertegard S, Hammarberg B. Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *J Voice*. 1995;9:182-197.
61. Byrd D. Relations of sex and dialect to reduction. *Speech Commun*. 1994;15:39-54.
62. Huber D. Aspects of the communicative function of voice in text intonation [Ph.D. thesis]. Chalmers University, Gokteborg/Lund; 1988.