Journal of Phonetics 40 (2012) 466–476

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

An acoustic and electroglottographic study of White Hmong tone and phonation

Christina M. Esposito*

Macalester College, St. Paul, MN 55105-1801, USA

ARTICLE INFO

Article history:
Received 26 May 2010
Received in revised form
17 February 2012
Accepted 23 February 2012
Available online 29 March 2012

ABSTRACT

This study examines tone and phonation in White Hmong, a language with seven tones (traditionally described as: high, mid, low, high-falling, mid-rising, low-falling, and mid-low) and three phonations (low-falling tone is creaky, mid-low tone is breathy and the remaining tones are modal). Thirty-two speakers were recorded producing words with all seven tones; audio and electroglottographic recordings were made. Acoustic measures were: cepstral peak prominence (CPP), H1*, H2*, H1*−H2*, H1*−A1*, H1*−A2*, H1*−A3*, and H2*−H4*. Electroglottographic (EGG) measures were: closed quotient and derivative-EGG closure peak amplitude (DECPA). F0 and duration were measured. Results showed that the traditional tonal descriptions are accurate except for the high-level tone which is better described as rising and the mid-low tone, which is falling. Furthermore, the rising and low-falling tones are shorter than the other five tones. In terms of acoustic and electroglottographic measures, none of the measures tested distinguished all three phonation types at a given time point. Several measures, H1*, H1*−H2*, CQ, CPP, and DECPA, distinguished two phonation categories, suggesting that phonation contrasts are realized across several phonetic dimensions. Additional results showed that many of the acoustic and EGG measures were correlated with F0 and that closed quotient and DECPA were most strongly correlated with H1*−H2*.

1. Introduction

The Hmong (Hmong-Mien) languages have a complex tonal system that includes changes in pitch as well as phonation. For example, White Hmong has seven phonemic tones, two of which are associated with non-modal phonation. The traditional tonal description[1] is: high, mid, low, high-falling, mid-rising, low-falling creaky and mid-low breathy. (There is also an eighth tone which is a syntactic variant of the low-falling creaky tone.) However, accounts of the tones and phonations differ slightly across sources. For example, both Smalley (1976) and Ratliff (1992) describe the mid-low breathy tone as a falling tone. Table 1 presents the traditional tonal description, the tonal description from Smalley (1976) and Ratliff (1992) (which are the same unless otherwise noted), the Romanized Popular orthographic symbol for that tone (which is indicated with one of seven consonants at the end of a syllable) and a sample word in White Hmong.

A number of studies have examined (or observed) various acoustic properties associated with these tones/phonations. Huffman (1987) examined four tones (low (22), falling (42), low-breathy (32), and checked (31), a modal vowel followed by a glottal stop) in White and Blue Hmong. (Dialectal differences between these two groups do not involve the tones/phonations; results were collapsed across dialects.) Duration, fundamental frequency, spectral tilt, the amplitude of the first harmonic minus the amplitude of the second harmonic (H1−H2), and two measures made by inverse filtering (glottal pulse symmetry and relative closed-phase duration) were made for each token. Results showed that the checked tone was significantly shorter than the other three tones. Results of the F0 measurements indicated that the low and low-breathy tones were more accurately described as falling. Finally, spectral and glottal flow measures yielded mixed results. While measures of spectral tilt and glottal pulse symmetry were not reliable measures of phonation, spectral balance (specifically H1−H2) and closed-phase duration successfully distinguished modal from breathy phonation.

Andruski and Ratliff (2000) examined three of the seven tones of Green Mong, the breathy (*-g*) tone, the creaky (*-m*) tone, and the low-falling modal (*-s*) tone. The *-s* tone was chosen for comparison because its F0 is similar to the F0 of the breathy

* Tel.: +1 651 696 6678; fax: +1 651 333 4754.
E-mail address: esposito@macalester.edu

[1] The traditional description of Hmong tones is based on Lyman (1979) for Green Mong. White Hmong and Green Mong are mutually intelligible. The tones/phonation system of Green Mong is identical to that of White Hmong. Green Mong does not contain the phoneme /m̤/. Therefore, in Green Mong, the word 'Hmong' is pronounced [mɔ̃ 55] and thus written without the ⟨h⟩ in the Romanized orthography.

0095-4470/$ - see front matter
doi:10.1016/j.wocn.2012.02.007

Table 1
White Hmong tones. The traditional description, Smalley (1976) and Ratliff's (1992) descriptions (which are the same unless otherwise noted), the orthographic convention and a sample word are given. A 5 indicates the highest F0 and a 1 the lowest.

Traditional description	Smalley (1976)/ Ratliff (1992)	Orthographic tone symbol	Example in White Hmong orthography
High	High level (55)	**b**	*pob* 'ball-like'
Mid	Mid level (33)	**ø**	*po* 'spleen'
Low	Low level (22)	**s**	*pos* 'thorn'
High-falling	High-falling (52)	**j**	*poj* 'female'
Mid-rising	Mid-rising (24)	**v**	*pov* 'to throw'
Low-falling creaky	Falling creaky (21)	**m**	*pom* 'to see'
Mid-low breathy	Falling breathy (42)	**g**	*pog* 'paternal grandmother'

and creaky tones. Results showed that H1−H2 values were greater for breathy phonation than they were for non-breathy phonations. However, when comparing across three vowel qualities ([a], [i] and [u]) results indicated that the breathy tone was less breathy for both [i] and [u] than it was for [a]. Furthermore, H1−H2 was significantly different for all three tones on the vowels [i] and [u], but not on the vowel [a]. For [a], H1−H2 differences only distinguished the breathy from the non-breathy tones. Creaky phonation, which had the shortest duration of the three tones, also had a slightly higher vowel height than either breathy or modal phonation. Breathy and modal phonation had the same vowel height, though breathy phonation was more centralized. Additional results showed that there was no significant main effect of gender.

Ratliff (1992) studied tonal morphology in White Hmong, but also reported on a number of features of the tones and phonations. For example, there was an observed gender difference in regards to the *-g* (falling breathy) tone. For male speakers, the *-g* tone was more accurately described as 31, while for female speakers, the *-g* tone was 53. Consistent with the aforementioned studies, the *-m* tone (21, creaky) was perceived to be shorter than the other tones/phonations. Ratliff also reported a difference between the *-s* (22) and *-ø* (33) tone: "The difference between the *-s* tone (22) and the *-ø* tone (33) seems to involve something other than pitch. It may be duration: there is a certain 'chanted' quality to the mid level tone, which has no perceivable fall at the end, while there is a more natural tapering at the end of the low level tone" (p. 11).

The current study examines several properties of White Hmong tones/phonations, and expands on the aforementioned works in several ways: by including all seven tones, by taking a wider range of spectral measures, and by introducing electroglottographic measures. Questions and observations raised in previous studies (regarding the F0 of the tones, the acoustic correlates of the phonation differences, the duration of the tones/phonations, etc.) will also be explored. More specifically, the following questions about the tones/phonations will be addressed:

(1) What is the most accurate description of the seven tones?
(2) Are there durational differences between the tones?
(3) What are the acoustic and electroglottographic correlates of the phonations?
(4) How does the phonation co-vary with tone (F0)?
(5) How does phonation vary with vowel quality?

In addition to investigating the tones and phonations of White Hmong, this study will also examine the relationship between spectral and electroglottographic measures.

2. Background

2.1. Review of the production of phonation

Ladefoged (1971) proposed a continuum of glottal states, from most open (spread as for a voiceless sound) to most closed (a glottal stop). The three main types of phonations that occur cross-linguistically, modal, breathy, and creaky, each represent a point on the continuum.

Breathy phonation is produced by vocal folds with minimal adductive tension and little longitudinal tension that vibrate without much contact. Modal phonation is produced by vocal folds that have normal adductive and longitudinal tension. Vocal fold aperture is smaller for modal phonation than it is for breathy phonation, but greater than for creaky phonation. Creaky phonation is produced by vocal folds that have the greatest degree of closure, high adductive tension, and little longitudinal tension. For additional details on these and other types of phonations, see Laver (1980) and Gordon and Ladefoged (2001).

More recently, Edmondson and Esling (2006) proposed that phonation is produced by manipulating six valves. The six valves are (1) glottal vocal fold adduction and abduction, (2) ventricular incursion, (3) sphincteric compression of the aryepiglottic fold constriction, (4) epiglotto-pharyngeal constriction, (5) laryngeal raising, and (6) pharyngeal narrowing. In this system, creaky voice is produced by vocal folds that vibrate slowly, with sphincteric compression of the aryepiglottic fold and little or no ventricular incursion, while breathy voice is produced with a partial closure of the glottal vocal fold which leaves a large aperture between the arytenoid cartilages.

The majority of studies on phonation have focused on defining and describing breathy, modal, and/or creaky phonation in various languages. Examples include: Fischer-Jørgensen (1967) on Gujarati, Huffman (1987) on Hmong, Thongkum (1988) on Nyah Kur and Kui, Kirk, Ladefoged, and Ladefoged (1993) on Jalapa Mazatec, and, more recently, Andruski and Ratliff (2000) on Green Mong, Wayland and Jongman (2003) on Khmer, Abramson, Thongkum, and Nye (2004) on Suai, DiCanio (2009) on Takhian Thong Chong, and Esposito (2010) on Santa Ana del Valle Zapotec. These and other similar studies have demonstrated a number of cross-linguistic similarities in the production of phonation. For example, H1−H2 is a successful measure for distinguishing phonations in a variety of languages (Hmong (Huffman, 1987), Mazatec (Blankenship, 2002), and Green Mong (Andruski & Ratliff, 2000)) suggesting that speakers of these languages produce phonation contrasts in the same ways. However, a number of differences across languages have also been revealed, emphasizing the multidimensional nature of phonation. For example, in Santa Ana del Valle Zapotec, for male speakers, phonations are distinguished by H1−A3 (the amplitude of the third formant peak) and not by H1−H2 differences (Esposito, 2010).

Furthermore, even within a language, there can be differences in the production of phonation. While male Santa Ana del Valle Zapotec speakers distinguish phonation categories along H1−A3, female speakers use a different acoustic dimension, H1−H2 (Esposito, 2010). And, in !Xóõ, breathy vowels can be distinguished from modal ones by H1−H2, spectral tilt, and/or noise, depending on the speaker (Ladefoged & Antoñanzas-Barroso, 1985).

A number of studies have also revealed a strong correlation between F0 and phonation. For example, breathy voiced stops in Hindi (Ohala, 1973) and breathy and creaky vowels in Santa Ana del Valle Zapotec (Esposito, 2010) are produced with a lower F0 than their modal counterparts. Tense phonation is associated with a higher F0 than its non-tense counterparts in Jingpho, Lahu, and Yi (Madiesson & Hess, 1987) and Chong (DiCanio, 2009; Thongkum, 1991).

Additional studies have shown that non-modal phonation is often localized to a portion of a vowel, especially in languages with contrastive phonation and contrastive tone. For example, in Jalapa Mazatec, non-modal phonation is localized to the beginning of the vowel (Blankenship, 2002; Silverman, 1997), while in Santa Ana del Valle Zapotec, non-modal phonation is confined to the end of the vowel (Esposito, 2003).

2.2. Acoustic measures of phonation

The difference between phonation types can be quantified through a variety of measurements, each reflecting a different aspect of production. The most common acoustic measures have been spectral measures. Primarily, the amplitude of the first harmonic minus the amplitude of the second harmonic (H1 – H2) has been used to measure phonation (e.g. Gujarati (Bickley, 1982; Fischer-Jørgensen (1967)), Mazatec (Blankenship, 2002), Green Mong (Andruski & Ratliff, 2000), Takhian Thong Chong (DiCanio, 2009), Santa Ana del Valle Zapotec (Esposito, 2010) and for a small set of data from Chong, Fuzhou, Green Mong, White Hmong, San Lucas Quiaviní Zapotec, Tlacolula Zapotec, and !Xóõ (Esposito, 2006)). Another low frequency measure, the amplitude of the second harmonic minus the amplitude of the fourth harmonic (H4) has been used to measure phonation in pathologically-disordered voices (Kreiman, Gerratt & Antoñanzas-Barroso, 2006). Other studies have made use of the relationship between H1 and harmonics exciting formants (for example, A1, A2, A3, and A4). These include: H1 – A3 (English (Stevens & Hanson, 1995), Chong (Blankenship, 2002), Takhian Thong Chong (DiCanio, 2009), for a small set of data from Chong, Green Mong, White Hmong, Mon, San Lucas Quiaviní Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ (Esposito, 2006), and Santa Ana del Valle Zapotec (Esposito, 2010)); H1 – A1 (!Xóõ (Ladefoged, 1983), for a small set of data from Chong, Fuzhou, San Lucas Quiaviní Zapotec, Santa Ana del Valle Zapotec, and Tlacolula Zapotec (Esposito, 2006); H1 – A2 (for a small set of data from Chong, Fuzhou, Mon, San Lucas Quiaviní Zapotec, Santa Ana del Valle Zapotec, Tlacolula Zapotec, and Tamang (Esposito, 2006) and the average of H1 – H2 compared to A1 (English (Stevens, 1988)). Other studies have used the relationship of higher formants to lower ones, such as A2 – A3 (English (Klatt & Klatt, 1990)).

The various spectral measures have been associated with physiological characteristics. Holmberg, Hillman, Perkell, Guiod, and Goldman (1995) showed that the difference between H1 and H2 correlated with the portion of time the glottis was open during each glottal cycle (the open quotient). When the vibration of the vocal folds has a large open quotient, the amplitude of the first harmonic is greater than that of the second harmonic. Thus, for breathy phonation, which has a large open quotient, the spectrum is dominated by H1; the value (in dB) of H1 – H2 is largely positive for breathy phonation. In addition, Stevens (1977) suggested that measures of spectral slope correlated with the abruptness of vocal fold closure. When the vocal folds come together abruptly, they excite the higher frequencies of a vowel. For breathy phonation, which is characterized by vocal folds that close slowly, the higher harmonics are lower in amplitude than H1; measures of the spectral slope are largely positive for breathy phonation. There has yet to be a study with direct articulatory observations that support the theory that spectral slope is related to the speed of vocal fold closure. One of the goals of the current study will be to provide support for the theory by comparing acoustic measures of the spectral slope with electroglottographic measures of the speed of vocal fold closure (specifically, DECPA).

Because they reflect different aspects of phonation, the different measures of spectral tilt do not always distinguish phonation types, even within a single language. For example, Blankenship (2002) found that in Mpi, H1 – H2 was a more reliable measure of phonation on vowels with a high tone than on vowels with a mid or low tone. Also, Esposito (2010) found that H1 – H2 was the best measure of phonation for female speakers of Santa Ana del Valle Zapotec in that it successfully distinguished breathy, modal, and creaky phonation in the expected directions, though this measure did not distinguish the three phonation types for the male speakers.

Studies have also demonstrated a relationship between spectral measures and F0. For example, Kreiman, Gerratt, and Barroso (2007) examined 78 different spectral measures on normal and pathological productions of the vowel [a]. Results showed that F0 was positively correlated with H1* – H2* (where * indicates that the measure was corrected for the effects of formant frequency and bandwidth). In addition, Iseli, Shue, and Alwan (2007) found that H1* – H2* was positively correlated with F0 for low-pitched talkers (where $F0 \leq 175$ Hz), while for high-pitched talkers, H1* – H2* was dependent on vowel height.

In addition to spectral measures, measures of the periodicity of the signal have also been used to quantify phonation differences. A common measure of periodicity has been cepstral peak prominence (CPP) (English (Hillenbrand, Cleveland, & Erickson, 1994); Chong, Mazatec, Mpi, Tagalog (Blankenship, 2002); for a small set of data from Chong, Fuzhou, Green Mong, White Hmong, Mon, Santa Ana del Valle Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ (Esposito, 2006).

2.3. Electroglottographic measures of phonation

An electroglottograph (EGG) indexes the degree of contact between the vocal folds and has been a useful tool in measuring and describing phonations (e.g. Maa (Guion, Post, & Payne, 2004), Vietnamese (Michaud, 2004), Santa Ana del Valle Zapotec (Esposito, 2005), Tamang (Mazaudon & Michaud, 2006) and Takhain Thong Chong (DiCanio, 2009)). Childers and Lee (1991) used EGG measures to extract source-related features for modal, vocal fry, falsetto and breathy voice as produced by subjects with and without vocal disorders. Four factors were found to be the most important in distinguishing the four phonation types: glottal pulse width, pulse skewing, the abruptness of glottal closure, and turbulent noise. Results for glottal pulse width showed that modal phonation was produced with medium glottal width; vocal fry, a short glottal width, and falsetto and breathy voice, a long glottal width. For pulse skewing, modal phonation was characterized by medium skewing; vocal fry, high skewing, and falsetto and breathy voice, low skewing. In terms of abruptness of closure, modal voice and vocal fry were characterized by abrupt closure, while falsetto and breathy phonations were characterized by progressive closure. Finally, breathy phonation was the only phonation produced with high turbulent noise.

The most common EGG measure is closed quotient, also referred to as contact quotient or CQ, which reflects the portion of time the vocal folds are closed during each glottal cycle. CQ is the inverse of the open quotient measure (OQ). DiCanio (2009) compared EGG and spectral measures (specifically, H1 – H2 and H1 – A3) for four contrastive registers in Takhian Thong Chong and found that open quotient was more closely correlated with H1 – H2 than with H1 – A3. Assuming a model of phonation based on glottal opening, phonations with greater opening have lower CQ values than phonations with greater vocal fold contact.

Recently, another EGG measure, Derivative-EGG Closure Peak Amplitude or DECPA (e.g. Michaud, 2004 for Mandarin, Naxi and Vietnamese; Tuan, d'Alessandro, & Michaud, 2005 on Vietnamese) has been used to measure the speed of glottal closure. Specifically, DECPA is a measure of peak closure amplitude from the derivative of the EGG signal. Phonations produced with faster glottal closure have greater DECPA values than phonations produced with slower glottal closure. (For more information on EGG measures see: Childers and Krishnamurthy (1985), Baken and Orlikoff (2000) and Heinrich, D'Alessandro, Doval, and Castellengo (2004).)

3. Current study

3.1. Methods

3.1.1. Speakers

Thirty-two native speakers of White Hmong were recorded. Subjects were recruited at the Hmong American Partnership (St. Paul, Minnesota) and the Immanuel Hmong Lutheran Church (St. Paul, Minnesota). Speakers ranged 18–48 years of age and were born in either Laos, Thailand or the US, but currently reside in Minneapolis/St. Paul, Minnesota. All the speakers spoke English in addition to White Hmong; the reported age of English onset ranged 5–35 years of age. All speakers reported that they used White Hmong daily. Table 2 summarizes the background information on the speakers; speaker gender, age, birthplace, number of years in the United States, and the reported age of English onset are given.

3.1.2. Speech materials

Speakers produced words with all seven tones for a total of 70 words. The word list is given in Appendix A.

All words were monosyllabic (CV) and contained one of six White Hmong vowels [i, e, ɨ, a, u, ɔ]. All seven tones were recorded. A total of thirty-eight words had modal phonation. (Five tokens had a *-b* tone, five *-ø* tone, fifteen *-s* tone, seven *-j* tone, and six *-v* tone.) Seventeen words had creaky phonation (*-m*) and fifteen had breathy phonation (*-g*). Throughout the paper, the phonation associated with the *-g* tone will be referred to as 'breathy'. The phonation on the *-m* tone will be labeled 'creaky'. The onset consonant varied, but was limited to voiceless or voiced oral obstruents. Aspirated consonants and nasalized vowels were avoided. An attempt was made to record minimal sets when possible. All words were uttered in the frame *rov hais* ____ *dua* [ɾɔ24 hai22 ____duə33] 'Say ____ again'. Fig. 1 is a set of waveforms, pitchtracks, and spectrograms for a minimal set of White Hmong tones as produced by a female speaker. From the top down, the words are: *pob* [pɔ 55] 'ball-like', *po* [pɔ 33] 'spleen', *pos* [pɔ 22] 'thorn', *poj* [pɔ 52] 'female', *pov* [pɔ 24] 'to throw', *pom* [pɔ̰ 21] 'to see', *pog* [pɔ̤ 42] 'paternal grandmother'.

Twenty-nine speakers were literate in White Hmong and read the wordlist. The remaining three speakers were less familiar with the White Hmong orthography. For these speakers, a research assistant, a native speaker of White Hmong, prompted the speakers with a spoken model of the word. These productions were double-checked to ensure that the subject was not simply mimicking the research assistant.

Table 2
Speakers, gender, approximate age, birthplace, number of years in the US, and reported age of English onset.

Speaker	Gender	Approximate age	Birthplace	Number of years in the US	Age of English onset
1	F	21	US	21	5
2	M	21	US	21	5
3	F	20	Thailand	13	6
4	M	37	Laos	20	12
5	M	33	Laos	22	11
6	M	27	Thailand	15	14
7	M	48	Laos	28	18
8	M	46	Laos	19	24
9	M	41	Laos	20	18
10	M	23	Thailand	4	19
11	F	23	Thailand	4	19
12	F	28	Laos	3	25
13	M	20	Thailand	3	17
14	F	24	Thailand	3	20
15	M	43	Laos	4	5
16	M	32	Laos	3	5
17	M	25	Thailand	4	21
18	M	44	Laos	3	40
19	M	28	Laos	4	25
20	F	18	Thailand	16	5
21	M	19	Thailand	14	5
22	M	26	Thailand	3	23
23	M	22	Laos	3	12
24	M	38	Laos	30	12
25	M	39	Laos	2	5
26	F	26	Thailand	3	23
27	F	19	US	19	5
28	M	31	Laos	3	29
29	F	26	US	26	5
30	F	28	Laos	4	25
31	M	28	Laos	4	25
32	M	21	Thailand	3	19

3.1.3. Procedure

For twenty-one speakers, audio recordings were made. For eleven speakers, simultaneous audio and electroglottographic recordings were made using a Glottal Enterprises two-channel electroglottograph. All words were recorded in a quiet room into a head-mounted microphone attached to an XAudioBox (an audio codec manufactured by Scicon R&D) attached to a laptop. Fig. 2 is audio and electroglottographic signals for three White Hmong words, each with a different phonation: [pɔ 55] 'ball-like', *pog* [pɔ̤ 42] 'paternal grandmother', and *pom* [pɔ̰ 21] 'to see'.

Acoustic and EGG measures were taken automatically using VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011) and EGGWorks (Tehrani, 2009), respectively. Fundamental frequency (F0), eight spectral measures and two EGG measures were made for each vowel. Of the eight spectral measures, one measure, cepstral peak prominence (CPP), was a measure of periodicity. (The cepstral peak prominence measurement was based on Hillenbrand, Cleveland, & Erickson, 1994.) The other seven were spectral measures, including:

1. amplitude of the first harmonic (H1*);
2. amplitude of the second harmonic (H2*);
3. amplitude of the first harmonic minus the amplitude of the second harmonic (H1*−H2*);
4. amplitude of the first harmonic minus the amplitude of the first formant peak (H1*−A1*);
5. amplitude of the first harmonic minus the amplitude of the second formant peak (H1*−A2*);
6. amplitude of the first harmonic minus the amplitude of the third formant peak (H1*−A3*);
7. amplitude of the second harmonic minus the amplitude of the fourth harmonic (H2*−H4*).

Corrections were done automatically in VoiceSauce. VoiceSauce corrects harmonic amplitudes by using measured formant frequencies and bandwidths estimated by those frequencies. Corrections are based on Iseli et al. (2007), which is an extension of Hanson (1995). For more details on VoiceSauce see Shue et al. (2011). The two EGG measures were closed quotient (CQ) and Derivative-EGG Closure Peak Amplitude (DECPA). Closed quotient was measured automatically by EggWorks (Tehrani, 2009) using the CQ by Hybrid method (Herbst & Ternström, 2006; Howard, 1995; Orlikoff, 1991; Rothenberg & Mahshie, 1988). With the

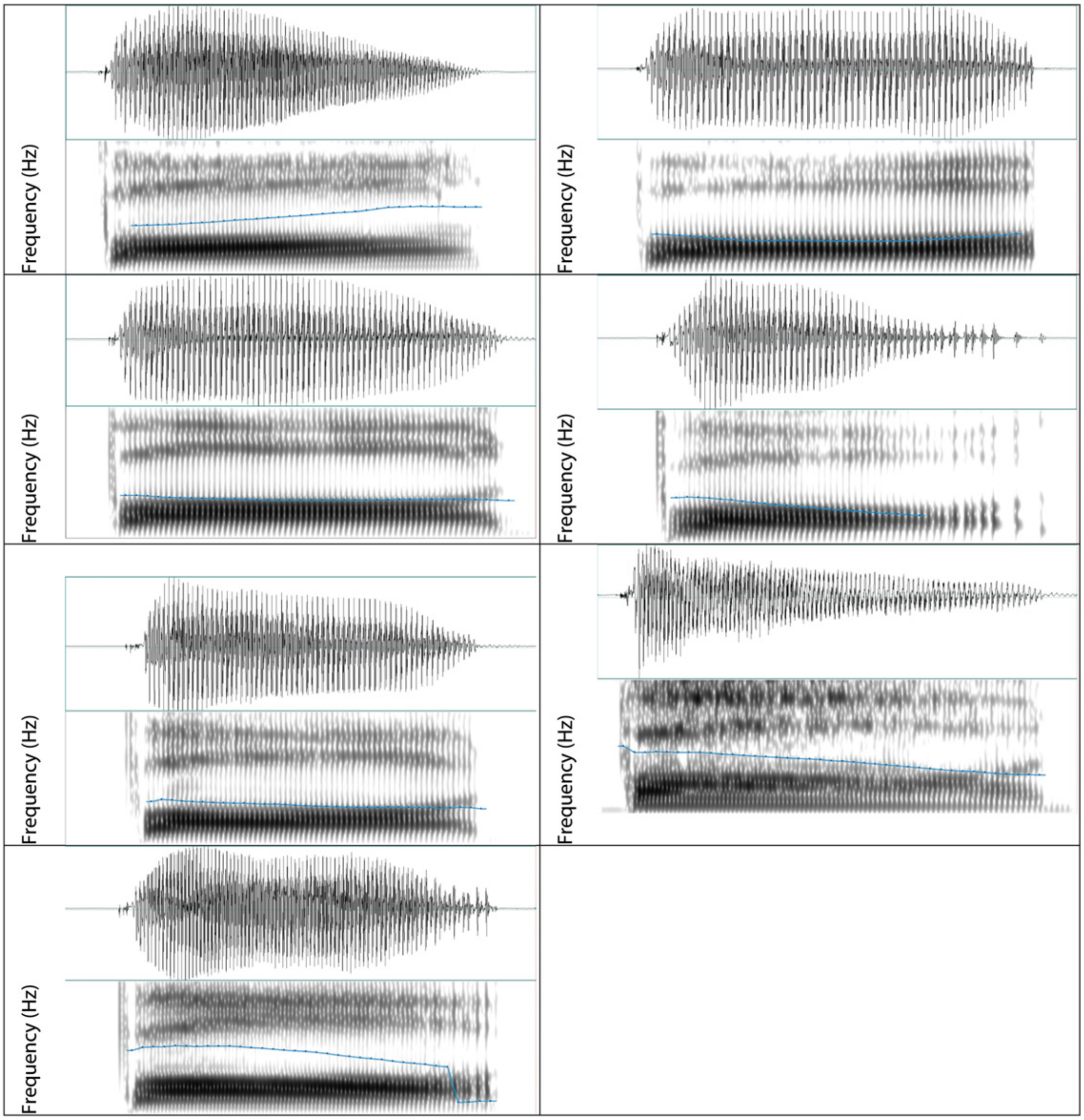

Fig. 1. Waveforms, pitchtracks, and spectrograms for a minimal set of White Hmong tone as produced by a female speaker. From the top-left down, the words are *pob* [pɔ 55] 'ball-like' *po* [pɔ 33] 'spleen' *pos* [pɔ 22] 'thorn' *pov* [pɔ 24] 'to throw' *pom* [pɔ̰ 21] 'to see' *pog* [pɔ̤ 42] 'paternal grandmother'. (Pitch is indicated as a blue line in the spectrogram.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

hybrid method for calculating closed quotient, the beginning of the contact phase is defined as the positive peak in the first derivative of the EGG signal, while the end of the contact phase is based on a fixed threshold, in this case 25% (Orlikoff, 1991) of the difference between the minimum and maximum amplitude values in each cycle of the EGG signal. Using a fixed threshold is advantageous because the negative peak in the EGG derivative is not always well-defined.

The second measure, DECPA, is the positive peak value from the derivative of the EGG signal (dEGG). The positive peak in the dEGG signal corresponds to the maximum speed of vocal fold closure and is a measure of the speed of the vocal folds at the moment of closure.

All measures were made automatically using VoiceSauce and EGGWorks. Measurements were made at multiple time points depending on the nature of the question being asked. To determine the best acoustic and electroglottographic measure of phonation, each vowel was divided into nine parts with equal duration. Measurements were made at all nine parts by averaging the value (for a given measure) of that part. To determine the F0 of the seven tones, all nine points were examined. To determine the acoustic and electroglottographic correlates of phonation, only points 1, 5, and 9 were examined (essentially, the beginning, middle, and end of the vowel). An additional set of spectral, EGG and F0 measurements was made at 35 points to examine how the acoustic and EGG measures might correlate with F0. And, finally, to determine the correlation between the acoustic and EGG measures, these measurements were made at every millisecond.

4. Results

4.1. *What is the most accurate description of the seven tones?*

The F0 of the seven tones was measured to determine the most accurate description for the tonal categories. Fig. 3 is a set of graphs of the F0 values over nine time points for males and females.

The results showed that the traditional description is accurate for the *-ø*, *-s*, *-j*, *-v*, and *-m* tones (the *-ø* tone is mid, *-s* is low,

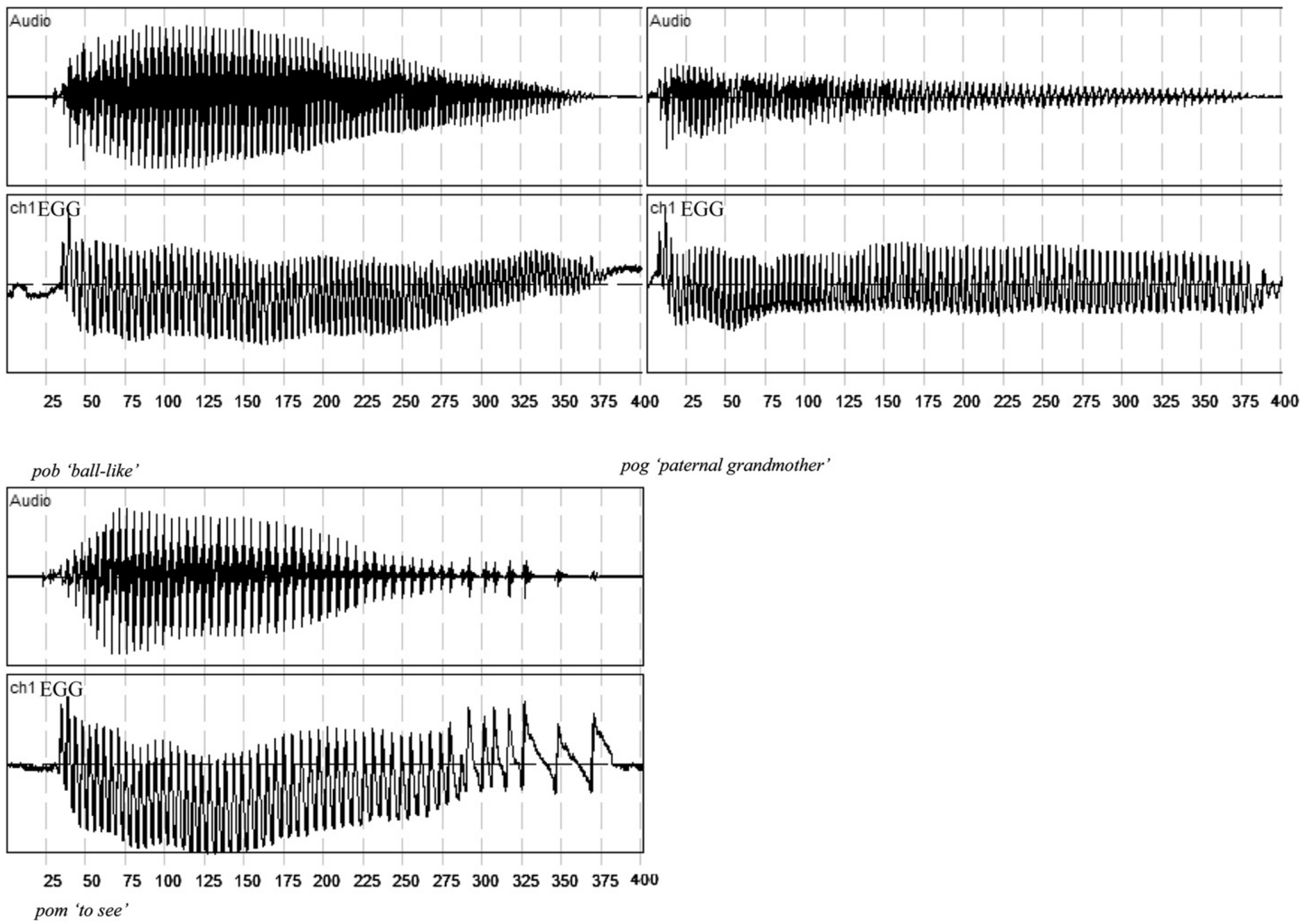

Fig. 2. Audio and electroglottographic signals for three White Hmong words: *pob* [pɔ 55] 'ball-like', *pog* [pɔ̤ 42] 'paternal grandmother', and *pom* [pɔ̰ 21] 'to see'.

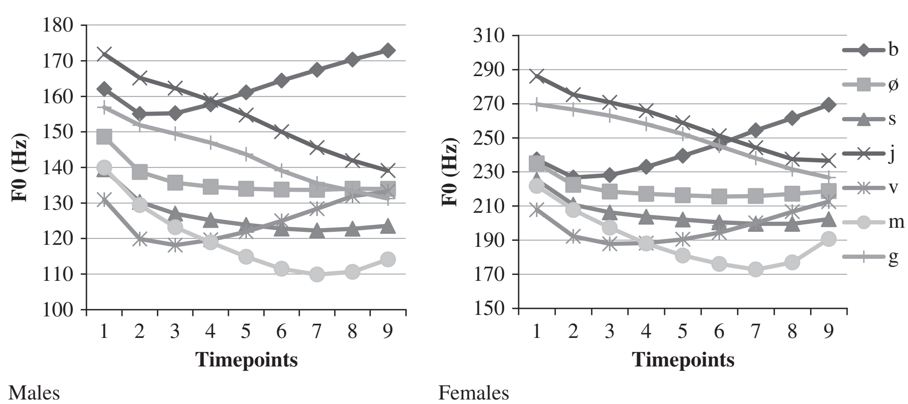

Fig. 3. Average F0 values measured at nine time points, which are proportional to vowel duration, for the seven tones for male (left) and female (right) speakers of White Hmong.

-*j* is high-falling, -*v* is mid-rising, and -*m* is low-falling). However, the traditional description of the -*g* tone as a mid-low tone is inaccurate; the -*g* tone is a falling tone as noted by Smalley (1976), Huffman (1987) and Ratliff (1992). Previous accounts have described the -*b* tone as a high level tone; the data here suggests that the -*b* tone is more accurately described as rising.

Furthermore, there are some apparent gender-based differences in the tones. For the female speakers, the -*j* and -*g* tones appear to be more similar than they were for the males. There is an average of an 8.37 Hz difference between the -*j* and -*g* tones for the females; there is an average of an 11.26 Hz difference between these tones for the male speakers. A paired *t*-test indicated there is not a significant difference between these two tones for females ($t(8)=7.5$, $p>.05$), while there is a significant difference for the males ($t(8)=14$, $p<.05$).

4.2. Are there durational differences between the tones?

Fig. 4 is a graph of the average duration (ms) of the seven tones.

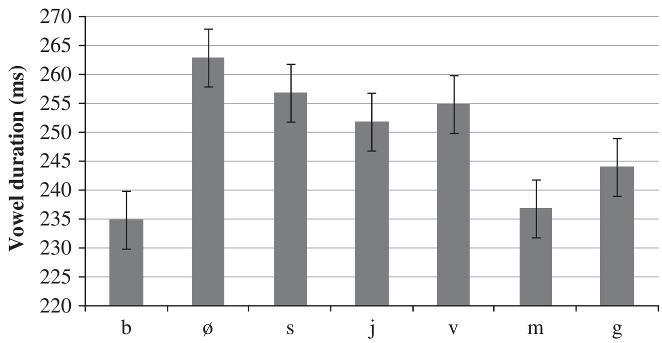

Fig. 4. Average duration (ms) for the seven tones. Error bars indicate standard error.

Table 3
Summary of the measures that successfully distinguish the three phonation types (breathy, modal, and creak) at three time points within a vowel (1, 5, 9). Each pair-wise comparison is shown (M=modal, B=breathy, C=creaky). A check mark indicates that there was a significant difference between the phonation categories. All results are significant at a $p<.001$ level unless otherwise noted. Results for males and females are the same unless otherwise noted. F=females only.

Pair-wise comparisons	Timepoints								
	1			5			9		
	MB	MC	CB	MB	MC	CB	MB	MC	CB
CPP				✓F		✓			
H1*	✓		✓	✓		✓		✓	✓
H2*									
H1*–H2*	✓		✓	✓		✓			
CQ			✓	✓		✓	✓		✓
DECPA	✓			✓		✓		✓	✓

A one-way repeated measures analysis with tone as a factor showed a significant main effect for duration ($F(7,30)=47.4$). Post-hoc tests revealed that the -*b* and -*m* tones are significantly shorter than the other five tones ($p<.05$). However, there is not a significant difference in the duration of the -*b* and -*m* tones ($p>.05$). Previous studies reported that the creaky tone (-*m*) was the shortest tone, but made no special mention of the duration of the -*b* tone.

4.3. *What are the acoustic and electroglottographic correlates of the phonations?*

To determine the acoustic and electroglottographic correlates of phonation, separate repeated measures ANOVAs for males and females with each tone as an independent variable for each measure at each time point were performed. All seven measures were entered into each of the ANOVAs. Results showed that there was a significant difference amongst the groups test on the measures H1* ($F(7,21)=83.1$), H1*−H2* ($F(7,21)=43.2$), CQ ($F(7,20)=57.6$), DECPA ($F(7,6)=63.8$), and CPP ($F(7,6)=48.2$) for males and the measures H1* ($F(7,9)=71.1$), H1*−H2* ($F(7,10)=51.1$), CQ ($F(7,10)=59.4$), DECPA ($F(7,5)=64.2$), and CPP ($F(7,5)=51.9$) for females. Post-hoc pair-wise comparisons were used to determine which tones significantly differed from each other on these measures. These results are presented in Table 3 and Fig. 5. All results are significant at a $p<.001$ level unless otherwise noted. None of the measures tested distinguished any of the modally-phonated tones (-*b*, -*ø*, -*s*, -*j*, - and -*v*). (There was no phonation difference between the -*s* tone and the -*ø* tone, which Ratliff suggested involved something other than pitch.) Therefore, the five modally-phonated tones are all collapsed into 'modal' in Table 3 and Fig. 5 to simplify the presentation.

For both male and female speakers, at the beginning of a vowel (point 1), H1*, H1*−H2* and CQ distinguished breathy from non-breathy (modal and creaky) tones. At point 1, none of the measures tested distinguished modal from creaky tones. At the middle of the vowel (point 5), H1*, H1*−H2*, DECPA and CQ distinguished breathy from modal tones while CPP, H1*, H1*−H2*, CQ, and DECPA distinguished breathy from creaky tones. In addition, for female speakers, CPP was also successful at distinguishing modal from breathy tones at point 5. At the end of the vowel (point 9), CQ distinguished breathy from modal tones, H1*, CQ and DECPA distinguished breathy from creaky tones, and H1*, DECPA distinguished creaky from modal tones. H2*, H1*−A1*, H1*−A2*, H1*−A3*, and H2*−H4* did not distinguish any of the phonation types. Table 3 summarizes these findings.

Separate graphs of the average value of the successful measures at each time point are presented below. A higher value (in dB) for the spectral measures indicates a breathier phonation. A higher value (in unspecified units) for CQ indicates greater vocal fold closure (e.g. a creakier phonation). A higher value (in unspecified units) for DECPA indicates a phonation with greater speed of vocal fold closure (e.g. a creakier phonation).

H1* and DECPA are the best measures of phonation in White Hmong, in that they distinguish all three phonation types, though not at the same time point. At points 1 and 5, H1* distinguishes breathy from non-breathy tones, but does not distinguish breathy and creaky tones. However, at point 9, H1* does distinguish breathy from creaky tones, as well as creaky from modal tones. At point 5, DECPA distinguishes breathy from non-breathy tones and at point 9 it distinguishes breathy from creaky tones, as well as creaky from modal tones.

There was a gender difference in phonation in that CPP distinguished modal from breathy tones at point 5 for the females. This was the only gender-based difference in phonation. (A two-way repeated measures ANOVA with gender as a between-subject factor and phonation/tone as a within-subject factor showed that there was a significant interaction between gender and phonation/tone at time point 5, with breathy versus modal phonation being distinguished by CPP only for females ($p<.05$). No significant main effect of gender was found.)

One unusual finding is in regards to DECPA, which is higher for breathy phonation than it is for modal and creaky phonation. This is unexpected, considering that previous research on phonation has suggested that breathy phonation is produced with a less abrupt glottal closure (see: Childers & Lee, 1991; Klatt and Klatt, 1990; Huffman, 1987 in particular for Hmong.). Visual inspection of the EGG signal for each individual speaker indicated that breathy phonation did have a faster vocal fold closure than modal and creaky phonation.

4.4. *How does the phonation co-vary with tone (F0)?*

To determine how phonation might co-vary with F0, measurements were made at 35 time points for each vowel. Only measures that successfully distinguish the phonations/tones of White Hmong will be examined here. Correlations between F0 and the acoustic/EGG measures are presented in Table 4. The correlations reflect an alpha-level correction, adjusting for multiple comparisons. A significant correlation ($p\leq.05$) is marked with an asterisk. For the correlations, the genders were examined separately, because the results of the current study showed that there was a gender difference in F0 and because Iseli et al. (2007) showed a difference in correlation between spectral measures and F0 due to gender.

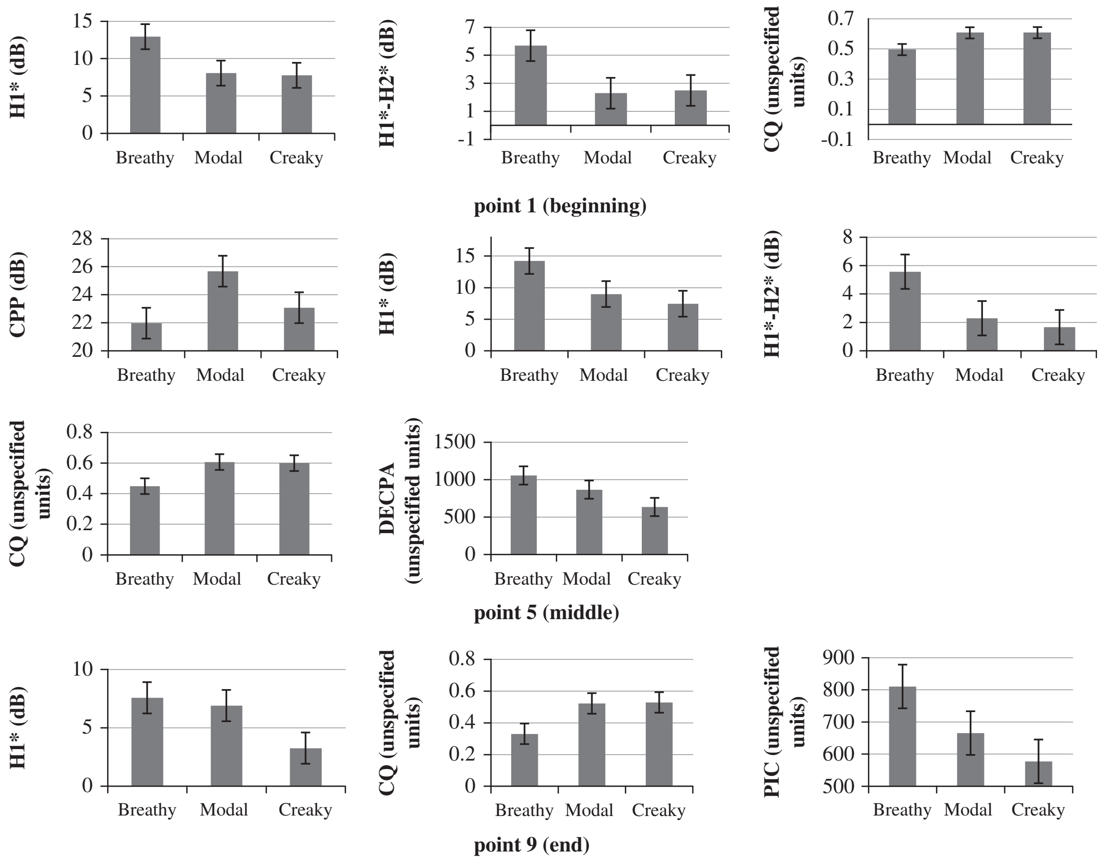

Fig. 5. Graphs of the average values for measures that successfully distinguished the three phonation types (breathy, modal, and creaky) at three time points within a vowel (1, 5, 9). Error bars indicate standard error.

Table 4
r-values for correlations between F0 and CPP, H1*, H1*−H2*, CQ and DECPA for both males (M) and females (F). Significance is marked with an asterisk.

	CPP		H1*		H1*−H2*		CQ		DECPA	
	M	F	M	F	M	F	M	F	M	F
-b	.20	−.32	−.45*	−.45*	−.68*	−.65*	−.83*	−.75*	−.92*	−.90*
-ø	−.85*	−.87*	.56*	.44*	.56*	.63*	−.54*	−.53*	−.50*	.49*
-s	−.60*	−.68*	.64*	.02	−.57*	−.85*	.65*	.67*	.10	.08
-j	.85*	−.02	.85*	.48*	−.85*	−.86*	.85*	.83*	.26	.24
-v	−.73*	−.75*	−.09	.53*	.74*	.76*	.83*	.81*	.81*	.84*
-m	−.32*	−.35*	.74*	.50*	−.46	.84*	.46*	.42*	.83*	.87*
-g	−.32	−.31	.90*	.77*	.88*	.82*	−.47*	−.45*	.78*	.80*

For the *-ø* tone, every measure tested was significantly correlated with F0. For the remaining tones, the majority of measures were correlated with F0. For the *-b* and *-g* tones all measures were significantly correlated except for CPP. For the *-s* tone, all measures were significantly correlated except DECPA; for the *-j* tone, all measures were significantly correlated except for DECPA for both males and females and CPP for the female speakers only. For the *-v* tone, all measures were significantly correlated except for H1* for the male speakers. And, for the *-m* tone all measures were significantly correlated except for H1*−H2* for the male speakers.

There is a great deal of variation in this data. The correlations between measurements and F0 do not always go in the same direction. For example, on the *-g* tone, there is a negative correlation between F0 and CQ, but there is a positive one between F0 and the other measures tested. Even for tones with the same phonation, the correlations do not always go in the same direction. The *-b*, - *ø*, *-s*, *-j*, - and *-v* tones all have modal phonation, yet there is a negative correlation between the *-b*, and *-ø* tones and CQ, while there is a positive correlation between CQ and the other modally-phonated tones (*-s*, *-j*, - and *-v*).

Table 5
Correlation between measures. Significance is marked with an asterisk.

	CQ	DECPA
Correlation between acoustic and EGG measures		
H1*	−.46*	.15*
H2*	.22	.18
H1*−H2*	−.60*	.35*
H1*−A1*	−.05*	−.16*
H1*−A2*	−.04*	−.14*
H1*−A3*	−.07*	−.26*
H2*−H4*	.06	.12
Correlation between EGG measures		
CQ and DECPA	−.56*	

Furthermore, even when there is a correlation between F0 and a given measure for both genders, the correlations do not always go in the same direction. For example, there is a negative correlation between DECPA and the *-ø* tone for males, but a positive one for female speakers.

4.5. Correlation between acoustic and EGG measures

To determine the relationship between the measures, measurements were made every millisecond. Results are presented in Table 5. A significant correlation ($p \leq .05$) is marked with an asterisk. The correlations reflect an alpha-level correction, adjusting for multiple comparisons.

Closed quotient was significantly inversely correlated with H1*, H1*−H2*, H1*−A1*, H1*−A2*, and H1*−A3*. There was not a significant correlation between CQ and H2* or CQ and H2*−H4*. Of these measures, the strongest correlations were between CQ and H1* and CQ and H1*−H2*; this is expected because closed quotient, H1*, and H1*−H2* are all measures of glottal aperture, whether directly or indirectly. Consistent with DiCanio (2009), the weaker correlations were between CQ and the measures H1*−A1*, H1*−A2*, and H1*−A3*, the three measures which reflect the speed of vocal fold closure rather than degree of glottal aperture.

DECPA was significantly correlated with H1*, H1*−H2*, H1*−A1*, H1*−A2*, and H1*−A3*. There was not a significant correlation between DECPA and H2* or DECPA and H2*−H4*. DECPA was inversely correlated with H1*−A1*, H1*−A2*, and H1*−A3*, suggesting that these acoustic measures might not reflect the speed of vocal fold closure. This finding is in line with Michaud (2004), which showed that DECPA was not a measure of spectral slope. In the current study, DECPA was positively correlated with H1* and H1*−H2*, but inversely correlated with Closed Quotient, suggesting that there is a relationship between DECPA and the degree of vocal fold opening.

4.6. How does phonation vary with vowel quality?

Separate two-way repeated measures ANOVAs for males and females with vowel and tone as factors (i.e. vowel, tone, and vowel X tone) for each corrected measure at each time point for each vowel ([i, a, ɔ, u, a͡ɨ, a͡u]) were performed. The diphthong [u͡ə] was not looked at because there was only one token with this vowel. Results showed that there was not a significant difference in the phonation types due to vowel quality.

5. Discussion and conclusion

The goal of this study was to answer the following questions about the tones/phonations of White Hmong:

(1) What is the most accurate description of the seven tones?
(2) Are there durational differences between the tones?
(3) What are the acoustic and electroglottographic correlates of the phonations?
(4) How does the phonation co-vary with tone (F0)?
(5) How does phonation vary with vowel quality?

In terms of tone, sources (Lyman, 1979; Ratliff, 1992; Smalley, 1976) agreed on the tonal description for the *-b*, *-ø*, *-s*, *-j*, *-v*, and *-m* tones (the *-b* is a high level tone, *-ø* is mid, *-s* is low, *-j* is high-falling, *-v* is mid-rising, and *-m* is low-falling). Results of the current study indicate that these descriptions are correct for all but the *–b* tone. While the aforementioned studies have described the *-b* tone as a high level tone, it is more accurately described as a rising tone. Sources have disagreed on the description of the *-g* tone. While Lyman (1979) describes the *-g* tone as a mid-low tone, Smalley (1976) and Ratliff (1992) describe the *-g* tone as a falling tone. Results of the current study confirm that *-g* is a falling tone. An additional finding was that female speakers are merging the F0 of the *-j* and *-g* tones. While the *-j* and *-g* are both high falling tones for males and females, the F0 differences between these tones was only significantly different for male speakers. Perhaps, these results indicate that a sound change is taking place in White Hmong led by the female speakers; studies have shown that females often lead sound change (Eckert, 1989; Labov, 1990). In the current study, the age range for females speakers is 18–28. In order to confirm this hypothesis, data will have to be collected from a wider range of ages. In addition, it would be interesting to see if males begin merging these tones in the future.

The duration of the tones/phonations was also measured. Results showed that the *-b* and *-m* tones were significantly shorter than the other five tones. Previous studies reported that the *-m* tone was the shortest tone, but made no mention of the duration of the *-b* tone. One unusual aspect of this finding is that the *-b* and *-m* tones are both contour tones (the *-b* tone is rising and the *-m* tone is falling), but have the shortest duration of the seven tones. Cross-linguistically, long vowels, not short ones, are more likely to carry contour tones (see Gordon, 2001 for a typological study on the restriction of contour tones).

The results of the acoustic and EGG measures showed that none of the acoustic measures tested distinguished all three phonation types at a given time point. However, several measures, H1*, H1*−H2*, CQ, CPP, and DECPA distinguished at least two phonation categories, suggesting that phonation contrasts are realized across several phonetic dimensions in White Hmong. The measures H2*, H1*−A1*, H1*−A2*, H1*−A3*, and H2*−H4* did not distinguish any of the phonation types.

Of the measures tested, the most successful were H1* and DECPA, in that they distinguished all three phonation types, though not at the same time point. On points 1 and 5, H1* distinguished breathy from non-breathy tones, and at point 9, it distinguished breathy from creaky tones. At point 5, DECPA distinguished breathy from creaky tones and at point 9, it distinguished creaky from non-creaky tones. None of the measures tested distinguished the three phonations at the same time point, emphasizing the importance of measuring phonation at various points within a vowel. In addition, H1* is a novel measure of phonation. Previous studies have measured H1(*) but always subtracted the value from that of another harmonic; no previous study has made used of H1* to measure phonation. It will be interesting to see if H1* is a successful measure of phonation in other languages.

Furthermore, many of the acoustic and EGG measures were correlated with F0. For the *-ø* tone, every measure tested was significantly correlated with F0. For the remaining tones, the majority of measures were correlated with F0. This suggests that there is a great deal of redundant information that listeners can use to determine relative F0 in White Hmong. A perception study would be needed to confirm which of these cues is the most important to listeners.

Finally, results showed that there was not a significant difference in the phonation types due to vowel quality.

One unusual finding is in regards to DECPA, which is higher for breathy phonation than it is for creaky and modal phonation. Previous research suggests that breathy phonation is produced with vocal folds that vibrate slowly (see Childers and Lee (1991), Klatt and Klatt (1990), and Huffman (1987) in particular for Hmong) and therefore has less abrupt vocal fold closure. Creaky phonation is produced with vocal folds that vibrate rapidly and have an abrupt closure. One possible explanation for the current finding is that creaky phonation is produced with a slower rate of vibration, which would suggest a slower rate of vocal fold closure.

There is some evidence to support this: (1) in Edmondson and Esling's (2006) laryngoscopic study of phonation, they describe creaky phonation as being produced by slow vocal fold vibrations and (2) higher DECPA values for breathy phonation, when compared to modal, were also found in Gujarati (Khan, 2009).

A second explanation has to do with degree of glottal opening. In breathy phonation, the vocal folds are further apart; perhaps they must move more quickly to get back together. Creaky phonation, on the other hand, is produced by vocal folds that are close together, and therefore do not need to move as quickly to reach a state of closure (Patricia Keating, p.c.). This explanation is supported by the negative correlation between DECPA and CQ ($r=-.56^*$). The more contact between the vocal folds, the higher the CQ value, but the lower the DECPA value.

However, if creaky phonation does indeed have a slower vocal fold closure, then this should also affect the spectral measures H1*−A1*, H1*−A2*, H1−A3*. These measures work under the principle that faster vocal fold closure excites the higher frequencies of a vowel, making A1*/A2*/A3 greater than H1*. Though the measures H1*−A1*, H1*−A2* and H1*−A3* were not significantly different for the three phonation types, they were in the expected direction (with breathy phonation having a higher value than creaky phonation). Previous studies have reported lower H1*−A1*/A2*/A3* values for creaky than for breathy phonation (Blankenship, 2002; DiCanio, 2009; Esposito, 2010). These results indicate that DECPA and H1*−A1*/A2*/A3* are not reflecting the same aspect of speech production. Indeed, there was a negative correlation between DECPA and these spectral measures. If the vocal folds are vibrating more slowly during creaky phonation and have a slower rate of vocal fold closure, then what is causing the value of A1*/A2*/A3* to be greater than that of H1*? More research, in particular into DECPA, what aspect(s) of phonation it is measuring, and its relationship to spectral slope, speed of glottal closure and glottal aperture, is needed to fully understand these findings.

Acknowledgments

This article derives from research supported by a National Science Foundation grant entitled *Production and perception of linguistic voice quality* with Abeer Alwan, Patricia Keating and Jody Kreiman. I would like to thank Joseph Ptacek, Sherrie Yang, the staff and clients at the Hmong American Partnership and the members of the Immanuel Hmong Lutheran Church for their assistance with finding and recording subjects. I would also like to thank Patricia Keating, Marc Garellek, and Sameer Khan for their helpful comments.

Appendix A. White Hmong Wordlist

Organized by tone. White Hmong words are written in the International Phonetic Alphabet and the Romanized Popular orthography. Reflecting the results of the current study, the *-b* tone has been transcribed as 45 and not the traditional 55.

IPA	Romanized popular orthography	Translation
ca 45	cab	'haul or tow'
ka 45	kab	'insect'
pɔ 45	pob	'ball-like'
ʈa͡u 45	taub	'pumpkin'
tɔ 45	tob	'very deep'
ca 33	ca	'log'
pɔ 33	po	'spleen'
ʈa͡u 33	rau	'six'
ʈa͡u 33	tau	'to be able'
tɔ 33	to	'puncture'
ca 22	cas	'why, how'
ka 22	kas	'maggots'
ki 22	kis	'to infect'
kɔ 22	kos	'stem'
pa 22	pas	'stick'
pɔ 22	pos	'thorn; cover'
pu 22	pus	'escape with all or part of the trap/arrow still attached'
qa 22	qas	'disgust, sicken'
ʈa͡u 22	raus	'to be full (of food)'
ta 22	tas	'finished, end, done'
ʈa͡ɨ 22	taws	'to emit light'
ti 22	tis	'wing'
tɔ 22	tos	'to wait'
tu 22	tus	'open-minded, stable, peaceful'
ʐa 22	zas	'color, paint'
ca 52	caj	'ridge'
ka 52	kaj	'dawn, bright'
pa 52	paj	'flower'
pɔ 52	poj	'female'
ʈa͡u 52	rauj	'hammer'
ʈa͡u 52	tauj	'species of grass'
tɔ 52	toj (nyob saum toj)	'ahead'
ca 24	cav	'argue, disagree'
ka 24	kav	'stem, stalk'
pɔ 24	pov	'throw'
ʈa͡u 24	rauv	'to light a fire'
ʈa͡u 24	tauv	'to dam up (water)'
tɔ 24	tov	'to mix'
ca̰ 21	cam	'to stick; to argue'
cḭ 21	cim	'mark'
ka̰ 21	kam	'allow'
kḭ 21	kim	'expensive'
kṵ 21	Kum	a proper name (male)
pa̰ 21	pam	'bridge; blanket'
pḭ 21	pim	'vagina'
pɔ̰ 21	pom	'to see'
pṵ 21	pum	'to see'
ʈa͡ṵ 21	raum	'kidney'
ta̰ 21	tam	'represent; sharpen (as in a knife)'
ʈa͡ṵ 21	taum	'bean'
ʈa͡ɨ̰ 21	tawm	'to go out'
tḭ 21	tim	'because of it'
tɔ̰ 21	tom	'to bite'
tṵ 21	tum	'to stack'
ʐa̰ 21	zam	'avoid'
ca̤ 42	cag	'root; origin'
ci̤ 42	cig	'light up'
cṳ 42	cug	'collect'
kṳ 42	kug	'very poor'
pɔ̤ 42	pog	'paternal grandmother'
qa̤ 42	qag	'axle; swivel'
ʈa͡ṳ 42	raug	'to suffer (something)'
ta̤ 42	tag	'done'
ʈa͡ṳ 42	taug	'to follow'
ʈa͡ɨ̤ 42	tawg	'explode'

t i̤ 42	tig	‘turn; steer’
tɔ̤ 42	tog	‘chair’
ʈ ʂ u͡ə̤ 42	tsuag	‘numb’
tṳ 42	tug	‘whose; properly’
ʐ a̤ 42	zag	‘sentence’

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.wocn.2012.02.007.

References

Abramson, A., Thongkum, T., & Nye, P. W. (2004). Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. *Phonetica*, *61*, 147–171.

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

Baken, R., & Orlikoff, R. (2000). *Clinical measurement of speech and voice*. San Diego: Singular homson Learning.

Bickley, C. (1982). Acoustic analysis and perception of breathy vowels. *Speech communication group working papers*, *Research Laboratory of Electronics* (pp. 73–93). Boston: MIT Press.

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

Childers, D., & Krishnamurthy, A. (1985). A critical review of electroglottography. *CRC Critical Reviews in Biomedical Engineering*, *12*(*2*), 131–161.

Childers, D., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, *90*(*5*), 2394–2410.

DiCanio, C. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*(2), 162–188.

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*, 245–267.

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

Esposito, C. M. (2003). *Santa Ana del Valle Zapotec phonation*. M.A. Thesis, University of California, Los Angeles.

Esposito, C. M. (2005, January). 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*.

Esposito, C. M. (2006). *The effects of linguistic experience on the perception of phonatio*. Ph.D. Dissertation, UCLA.

Esposito, C. M. (2010). Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association*

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

Gordon, M. (2001). A typology of contour tone restrictions. *Studies in Language*, *25*, 405–444.

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

Guion, S. G., Post, M. W., & Payne, D. L. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics*, *32*, 517–542.

Hanson, Helen M. (1995). *Glottal characteristics of female speakers*. Ph.D. Dissertation. Harvard University.

Heinrich, N., D'Alessandro, C., Doval, B., & Castellengo, M. (2004). On the use of the derivative of electroglottographic signals for characterization of non-pathological phonation. *Journal of the Acoustical Society of America*, *115*(*3*), 1321–1332.

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

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

Holmberg, E., Hillman, R., Perkell, J., Guiod, P., & Goldman, S. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, *38*, 1212–1223.

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

Huffman, M. (1987). Measures of phonation type in Hmong. *Journal of the Acoustical Society of America*, *81*, 495–504.

Iseli, M., Shue, Y., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustical measures related to the voice source. *Journal of the Acoustic Society of America*, *121*(4), 2283–2295.

Khan, S. (2009, October). An acoustic and electroglottographic study of breathy phonation in Gujarati. *Poster presented at the 158th meeting of the Acoustical Society of America*, Baltimore.

Kirk, P., Ladefoged, J., & Ladefoged, P. (1993). Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec. In: A. Mattina, & T. Montler (Eds.), *American Indian linguistics and ethnography in honor of Laurence C. Thompson* (pp. 435–450). Missoula, MT: Universtiy of Montana Press.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820–857.

Kreiman, Jody, Gerratt, Bruce, R., & Antoñanazas-Barros, Norma (2006). Analysis and synthesis of pathological voice quality'. Available from: 〈http://www.surgery.medsch.ucla.edu/glottalaffairs/files/GASoftwareManual2006.pdf〉.

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

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, *2*, 205–254.

Ladefoged, P., & Antoñanzas-Barroso, N. (1985). Computer measured of breathy voice quality. UCLA working papers, Vol. 61. pp. 79–86).

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago: University of Chicago.

Ladefoged, P. (1983). The linguistic use of different phonation types. In: D. Bless, & J. Abbs (Eds.), *Vocal fold physiology: Comtemporary research and clinical issues* (pp. 351–360). San Diego: College-Hill Press.

Laver, J. (1980). *Phonetic description of voice quality*. Cambridge: University of Cambridge Press.

Lyman, T. (1979). *Grammar of Mong Njua (Green Miao): A descriptive descriptive study*. California: Blue Oak Press.

Madiesson, I. & Hess, S. (1987). The effect of F0 on linguistics use of phonation types. *UCLA working papers in phonetics*, Vol. 67. pp. 112–118.

Mazaudon, M., & Michaud, A. (2006). Pitch and voice quality characteristics of the lexical word-tones of Tamang, as compared with level tones (Naxi data) and pitch-plus-voice-quality tones (Vietnamese data). *Speech Prosody 2006*, Dresden.

Michaud, A. (2004). Final consonants and glottalization: new perspectives from Hanoi Vietnamese. *Phonetica*, *61*(2–3), 119–146.

Ohala, J. (1973). The physiology of tone. *Southern California Occasional Papers in Linguistics*, *1*, 1–14.

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.

Ratliff, M. (1992). *Meaningful tone: A study of tonal morphology in compounds, form classes, and expressive phrases in White Hmong*. Center for Southeast Asian Studies Monograph series.

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

Shue, Y., Keating, P.,Vicenik, C., & Yu K. (2011) VoiceSauce: A program for voice analysis. In *Proceedings of the XVIIth international congress of phonetic sciences (ICPhS)*. Hong Kong.

Silverman, D. (1997). Laryngeal complexity on Otomanguean vowels. *Phonology*, *14*, 235–262.

Smalley, W. (1976). The problem of consonants and tone: Hmong (Meo, Miao). In: W. Smalley (Ed.), Phonemes and Orthography: *Languages Planning in Ten Minority Languages in Thailand*. *Pacific Linguistics Series C no. 43*. Canberra: Australian National University.

Stevens, K. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34*, 264–279.

Stevens, K. (1988). Modes of vocal fold vibration based on a two-section model. In: O. Fujimura (Ed.), *Vocal physiology: Voice production, mechanisms and functions* (pp. 297–317). New York: Raven Press.

Stevens, K., & Hanson, H. (1995). Classification of glottal vibration from acoustic measurements. In: O. Fujimura, & H. Minoru (Eds.), *Vocal fold physiology: Voice quality control*. San Diego, CA: Singular Publishing Group.

Tehrani, Henry (2009). 'EGGWorks: a program for automated analysis of EGG signals'. 〈http://www.linguistics.ucla.edu/faciliti/facilities/physiology/EggWorksSetup.exe〉.

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

Thongkum, T. (1991). An instrumental study of Chong registers. In: Jeremy Davidson (Ed.), *Essays in Mon-Khmer linguistics in honor of H. L. Shorto* (pp. 141–160). London: School of Oriental and AfricanStudies.

Tuan, V. N., d'Alessandro, C., Michaud, A. (2005). Using open quotient for the characterization of Vietnamese glottalized tones. In *Proceedings of the international conference. Eurospeech-Interspeech 2005: 9th European Conference on Speech Communication and Technology*, Lisbon, pp. 2885–2889.

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