

Proceedings of Meetings on Acoustics

Volume 4, 2008

<http://asa.aip.org>

**155th Meeting
Acoustical Society of America
Paris, France
29 June - 4 July 2008
Session 5aSCc: Speech Communication**

5aSCc14. Breathy and whispery voicing in White Hmong

Sean A. Fulop* and Chris Golston

***Corresponding author's address: Linguistics, California State University Fresno, 5245 N. Backer Ave., Fresno, California 93740, sfulop@csufresno.edu**

The White dialect of Hmong uses breathy voice as a tonal feature, and also a distinctive whispery voice as a stop consonant feature. In this paper, acoustic measurements are shown to validate the apparent differences between these two similar phonation types. In particular, relative harmonic intensity and harmonicity were found to be, in general, three ways distinct among Hmong modal, breathy, and whispery phonation. The discovery of distinctly pronounced breathy and whispery phonation in a single language has implications for the representational theory which is used to specify the phonetic grammar.

Published by the Acoustical Society of America through the American Institute of Physics

1. Introduction

Hmong is a language of the Hmong-Mien family (formerly called Miao-Yao) spoken in Laos, Thailand, and China, and more recently the United States. The dialect of White Hmong includes among its sounds a four-way distinction of alveolar plosives [t, t^h, d, d^h], where [d^h] is the symbol we have selected for a voiced stop with whispery voiced release. White Hmong is also a tone language, and includes a particular breathy voiced tone that sounds quite different from the whispery voiced plosive release.

The phonation types of breathy and whispery voice have often been regarded as merely varieties of breathy voicing; however they are thought to involve articulatorily and acoustically distinct manners of phonation which could potentially be exploited in a linguistic sound system. No language has been found to contrast these two phonation types phonologically; that is still technically true of White Hmong, however the difference in phonetic implementation is important to prevent near-homophony between certain syllables. In this paper we report on our successful acoustic measurement of the apparent phonetic difference between breathy and whispery voice in White Hmong. In order to perform the measurements, two possible metrics of the distinction were tried, since there is little or no extant literature which demonstrates how the two phonation types could be acoustically distinguished. Once appropriate measures are utilized, however, the acoustic distinction is maximally robust, to the extent that statistical testing of the distinction is rendered a formality.

2. Background to the study

2.1. White Hmong phonology

White Hmong was once thought to be a Sinitic language,¹ and although modern classifications place it into the small family of Hmong-Mien, it nevertheless displays a Sinitic type of morpho-syllabic structure hewing closely to the “one word equals one morpheme equals one syllable” outline. Each syllable is usually analyzed phonologically into a single onset “segment” and a rhyme. Some of the putative segments occurring as onsets are quite complex, however, leading to the proliferation of consonant phonemes shown in Table I using International Phonetic Alphabet symbols. The vowel system shown in Table II is considerably more manageable. Every syllable must contain at least one of the vowels, however a number of diphthongs combining these are also possible.

White Hmong is a seven-tone language, contrasting five modal-voice tones with distinctive pitch contours plus two others in which the phonation qualities (breathy versus creaky) seem to be the most salient feature.

Hmong tones: p̀ p̄ p̂ p̃ p̄ p̄ p̄

In this study our focus is on the breathy tone compared with the falling modal tone, which is judged to have the most similar pitch contour.

TABLE I. Consonant phonemes

labial	alveolar	retroflex	palatal	velar	uvular	glottal
p p ^h p ^l p ^{lh}	t t ^h d d ^h	ʈ ʈ ^h	c c ^h	k k ^h	q q ^h	ʔ
^m p ^m p ^h ^m p ^l ^m p ^{lh}	ⁿ t ⁿ t ^h	^ɳ t ^ɳ t ^h	^ɲ c ^ɲ c ^h	^ŋ k ^ŋ k ^h	^ɴ q ^ɴ q ^h	
	ts ts ^h	ʈʂ ʈʂ ^h				
	ⁿ ts ⁿ ts ^h	^ɳ tʂ ^ɳ tʂ ^h				
f v	s	ʂ ʐ	ç ʝ			h
m m̥ m ^l m ^{l̥}	n n̥		ɲ ɲ̥			
	l l̥					

TABLE II. Vowel phonemes

i	ɨ	u
e		
		ɔ
	a	

2.2. Breathy versus whispery voice

It is common for languages to employ phonation types other than modal phonation as phonologically contrastive attributes of certain sounds. *Breathy* phonation slackens the vocal cords so that they do not close completely during glottal cycles, with the result that excess air flows through at all times.² *Whispery* phonation manipulates the arytenoids so that the vocal cords do vibrate modally along part of their length, but excess air flows continually through a small posterior opening between the cords.^{2,3} No language contrasts breathy and whispery phonation; the two settings are usually regarded as variant types that may be found in one language versus another, or they may even be found in free variation in the same language. White Hmong presents the unique scenario in which the “breathy tone” syllable nuclei seem to be produced in the canonical way with slack vocal cords, but the voiced aspirated plosive [d^h] uses whispery voice during its release phase. These descriptions, at this stage, are impressionistic only, but the acoustic distinction between them will be validated below.

Numerous distinctly pronounced alveolar onsets are possible for a given nuclear segment in White Hmong, as well as breathy tone. Possible syllables involving alveolar plosives are: [tɔ, t^hɔ, dɔ, d^hɔ] with all tones except breathy, and also [tɔ̥, dɔ̥]. It is significant that the combinations [t^hɔ̥, d^hɔ̥] are not allowed in the language. In other words, no kind of

aspirated plosive is allowed as the onset to a breathy tone. The spectrograms in Figure 2.2 show typical examples of a three-way distinction involving voiced plosive onsets, in which it is expected that the portion of the vowel following the plosive release will possess a distinct *phonetic feature* value in each case (the notion of phonetic feature is discussed below). The plosive [d^h] has been described as employing whispery voice during the closure, with aspiration following release;⁴ this is never a correct description of our data.

3. Acoustic measures of breathy and whispery voice

Numerous past studies identify the relative harmonic intensities $H_1 - H_2$ and $H_1 - H_3$ as strong correlates of breathy vs. modal voice,⁶⁻⁸ although these studies are not specific as to whether they are dealing with phonetically breathy or whispery voicing. Another correlate of breathiness is the energy ratio of harmonics over noise (harmonicity or Harmonics-to-Noise Ratio).^{7,8} Since studies of breathy voice have uniformly denied the linguistic importance of distinguishing breathy (with slack vocal cords) from whispery phonation, there is no published investigation showing how to distinguish these.

Measurements were undertaken which pertain to two hypotheses, one concerning the harmonic relative intensity metric, and another concerning the harmonicity. We supposed that relative harmonic intensities $H_1 - H_2$ correlate to breathy voicing chiefly because of the slack vocal cords; this led to the hypothesis that the breathy tone of Hmong might have the highest relative H_1 intensity, and modal phonation the least relative H_1 intensity. We supposed that whispery voiced [d^h] release might have a relative H_1 intensity somewhere between the two extremes. Since Hmong whispery voicing differs audibly from breathy voicing chiefly in having more airflow noise through the glottal opening, we hypothesized that whispery phonation would have the lowest harmonicity, meaning the greatest proportion of energy in the noise. We supposed that modal phonation would have the highest harmonicity, and that breathy voicing would fall somewhere between the two extremes.

4. Measurement procedure

One male and one female native White Hmong speaker were recorded saying isolated words [dɔ̃], [d^hɔ̃], [dɔ̃]. Each word was repeated 18–20 times. In breathy voiced plosives such as those found in Hindi, studies have shown that this phonation type is crucial following the release,^{9,10} but is generally not produced during the closure. This finding cannot be assumed to apply to Hmong, but could be tested using a particular measurement procedure. Accordingly, to compare the three kinds of syllables, intensities (in dB) of harmonics H_1 , H_2 , H_3 were measured from Fourier power spectra over the plosive closures, and again for the first 50–80 ms interval following plosive release. Two relative intensity measures $H_1 - H_2$ and $H_1 - H_3$ were computed from the raw data. Harmonicity (in dB) was calculated for the same portions of the closures and vowels using Praat,¹¹ which implements Boersma's superior algorithm.¹² This gave a total of three metrics to test for their ability

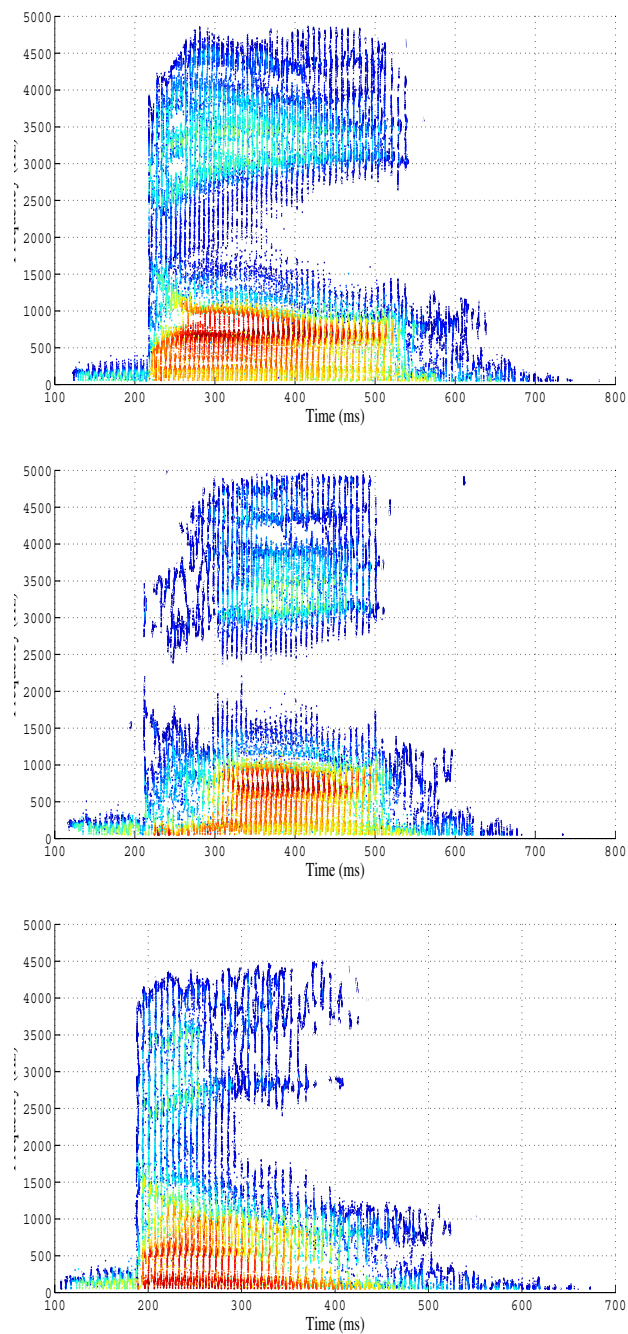


FIG. 1. Reassigned spectrograms⁵ (red = loud, blue = quiet) showing three syllables as spoken by the male Hmong speaker. Top: [dɔ̌] (falling tone); Middle: [dʰɔ̌] (falling); Bottom: [dɔ̌] (breathy).

TABLE III. Speaker 1 consonant release

	modal	whispery	breathy
Mean H1 – H2 (dB)	1.8	15.7	9.2
H1 – H3 (dB)	1.9	28.1	8.8
Harmonicity (dB)	12.9	0.5	12.2

TABLE IV. Speaker 2 consonant release

	modal	whispery	breathy
Mean H1 – H2 (dB)	–3.8	19.6	0.8
H1 – H3 (dB)	–5.1	17.9	11.9
Harmonicity (dB)	11.9	5.8	16.5

to discriminate the three syllables considered.

5. Results

The three syllables were able to be clearly and completely distinguished using most metrics with the release phases. Figure 2 shows typical power spectra from the release phases of the three different Hmong words. Table III gives the mean values of the release phase measurements for the male speaker. All three measures are three ways distinct (t-tests $p < 0.0005$) for this speaker, with the exception that the harmonicity does not differ between modal and breathy releases.

Table IV gives the mean values of the three release phase measurements for the female speaker. All three measures are once again three ways distinct (t-tests $p < 0.0005$). It might be noted that when statistically significant, the differences are also of a sizable dB magnitude.

Turning attention to the closure phases, most of our metrics failed to show any significant differences between the closure spectra of the three syllables analyzed, with one exception: $H_1 - H_2$ was significantly different between the closures of [d̚] (16.4 dB mean) and [d̚̚] (23.8 dB mean) for the female speaker only ($p < 0.0005$).

6. Analysis

Our hypotheses were generally correct that the three kinds of Hmong syllable [d̚, d^h̚, d̚̚] are three ways distinguished in their plosive release by harmonic intensity and harmonicity. Harmonicity alone could not distinguish [d̚, d̚̚] for the male speaker. Since

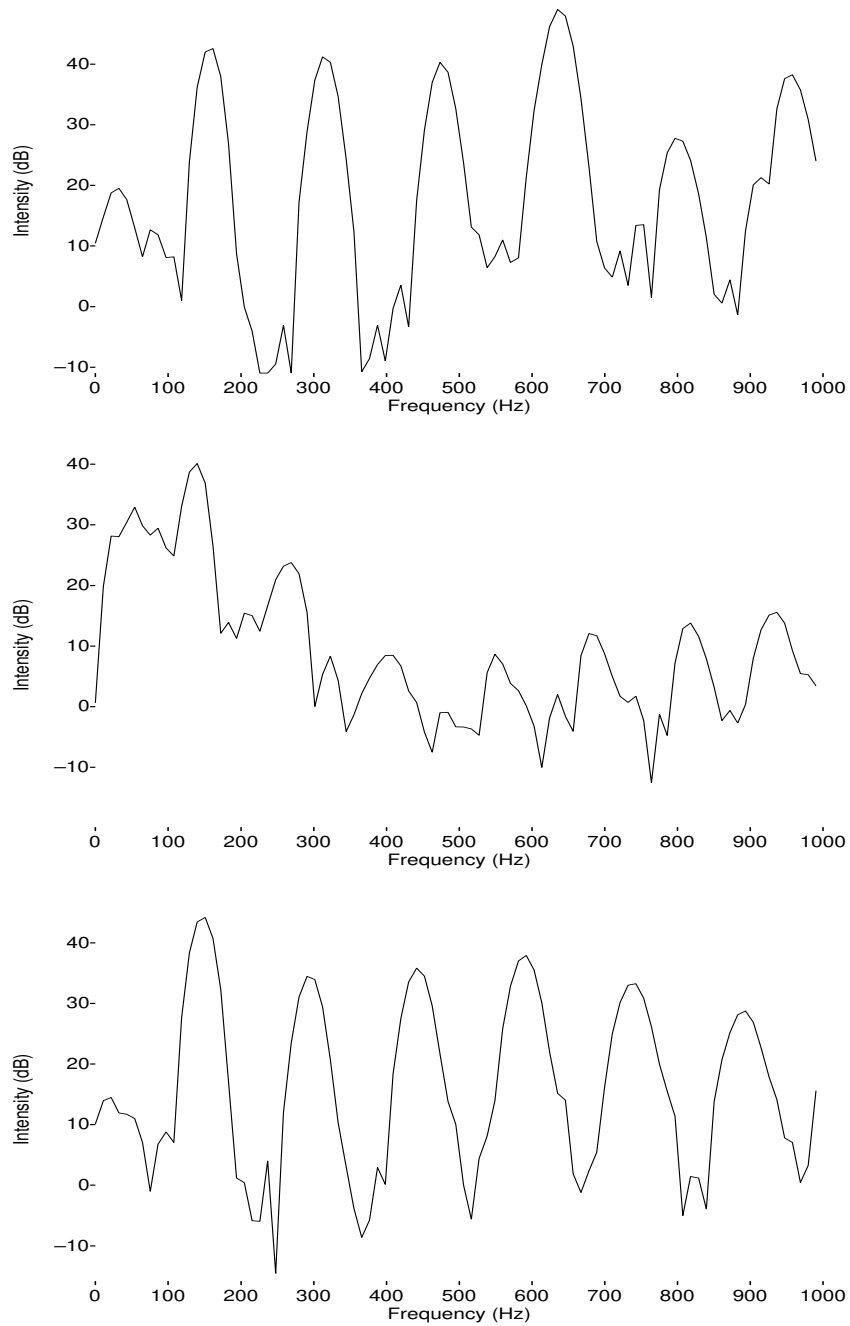


FIG. 2. Power spectrum of the first part of the vowel (~ 50 ms) in three syllables from the male Hmong speaker. Top: [dɔ̌]; Middle: [dʰɔ̌]; Bottom: [dɔ̌].

[d̀, d^h̀] are distinguished from [d̥] in the closure for one speaker using a harmonic intensity metric, it seems that breathy tone may spread to the onset, but not necessarily. Whispery voicing was never measured in the closure, replicating previous findings in the literature about similar plosives in Indic languages. The syllable [d^h̀] displayed the lowest harmonicity, and [d̀] the highest, confirming our hypothesis. [d^h̀] also displayed the highest relative H_1 intensity however, and this negates our hypothesis that it would be [d̥].

In any case, the three different phonation types that were impressionistically noted are easily measured as being completely different acoustically. In addition, the particular values of these metrics lends support to the general description of the phonation types as breathy (with slack vocal cords) and whispery. In particular, the mechanism posited for whispery voicing involves continuous airflow through a small glottal opening while also voicing, and this is expected to produce a higher noise intensity relative to the harmonics than the slack vocal cord mechanism for breathy voicing. The harmonicity results do bear this out, and this is as close as one can get to an acoustic verification of an articulatory hypothesis.

7. Discussion

Some authors such as Laver² and Catford³ have been careful to describe breathy and whispery phonation as involving different articulatory mechanisms. Others such as Ladefoged^{13,14} have described a simplified continuum of phonation types ranging from voiceless to breathy to creaky; the single parameter of variation is the degree to which the vocal cords are closed. This simplified view has been upheld in recent linguistic studies,⁸ but it admittedly excludes or ignores the phonatory mechanism of whisper. It has usually been presumed that whispery voice is an alternate implementation of breathy voice, and that for linguistic purposes any difference can be disregarded.

The phonetic distinctness of the whispery voicing in Hmong [d^h] from the breathy voiced tone is readily apparent and easily measured using the metrics explored here. The distinction helps to keep the syllables [d^h̀] and [d̥] from being so nearly homophonous in pronunciation. As such, the two phonation types should probably be viewed as different *phonetic features* or phonetic feature values.

In phonology/phonetics grammar models, each abstract phonological feature should be given a (context-sensitive) phonetic specification in the language at hand, as part of the phonetic component of the grammar.¹⁵ Chomsky and Halle advocated the need for different phonetic features or values whenever two sounds differ appreciably, whether or not they were contrastive phonologically. In the more recent model of Keating,¹⁶ it is low-level phonological rules which provide context-sensitive phonetic feature values to the segments. The specification of segments using phonetic feature values is called the **categorical phonetic representation**. These feature values are then implemented physically by language-specific phonetic instructions, but these are implicitly context-free or dependent upon speech physiology.

Given the phonetic situation in White Hmong, whispery voice and breathy voice must each be available to the grammar as different phonetic features or values at the categorical phonetic level. There is no other way of giving them distinct implementations in a model like Keating's. Simple and commonplace phonological feature combinations like [spread glottis] plus [voiced] will not be sufficient to specify whispery versus breathy voicing, so something is required at the level of phonetic grammar. Allowing this much, one is led to the conclusion that the simplified continuum of phonation types from voiceless to breathy to modal to creaky cannot be adopted as the complete range of possible values for a phonetic feature of [phonation type]; whispery must also be a possible value.

References

- ¹M. K. Huffman, "Measures of phonation type in Hmong," *Journal of the Acoustical Society of America* **81**(2), 495–504 (1987).
- ²J. Laver, *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge, 1980).
- ³J. C. Catford, *Fundamental Problems in Phonetics* (Edinburgh University Press, Edinburgh, 1977).
- ⁴N. Jarkey, "An investigation of two alveolar stop consonants in White Hmong," *Linguistics of the Tibeto-Burman Area* **10**(2), 57–70 (1987).
- ⁵S. A. Fulop, "Phonetic applications of the time-corrected instantaneous frequency spectrogram," *Phonetica* **64**, 237–262 (2007).
- ⁶P. Ladefoged, I. Maddieson, and M. Jackson, "Investigating phonation types in different languages," in *Vocal Physiology: Voice Production, Mechanisms, and Functions*, edited by O. Fujimura (Raven Press, New York, 1988).
- ⁷R. Wayland and A. Jongman, "Acoustic correlates of breathy and clear vowels: the case of Khmer," *Journal of Phonetics* **31**, 181–201 (2003).
- ⁸P. A. Keating and C. Esposito, "Linguistic voice quality," *UCLA Working Papers in Phonetics*, 85–91 (2006).
- ⁹R. P. Dixit, "Glottal gestures in Hindi plosives," *Journal of Phonetics* **17**, 213–237 (1989).
- ¹⁰A.-P. Benguerel and T. K. Bhatia, "Hindi stop consonants: an acoustic and fiberoptic study," *Phonetica* **37**, 134–148 (1980).
- ¹¹P. Boersma and D. Weenink, *Praat: Doing phonetics by computer*, Computer program, 2007, version 4.6.06.

- ¹²P. Boersma, “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound,” *Proceedings of the Institute of Phonetic Sciences, University of Amsterdam* **17**, 97–110 (1993).
- ¹³P. Ladefoged, “Linguistic phonetic features,” in *The Handbook of Phonetic Sciences*, edited by W. J. Hardcastle and J. Laver (Blackwell, London, 1997).
- ¹⁴M. Gordon and P. Ladefoged, “Phonation type: a cross-linguistic overview,” *Journal of Phonetics* **29**, 383–406 (2001).
- ¹⁵N. Chomsky and M. Halle, *The Sound Pattern of English* (Harper & Row, New York, 1968).
- ¹⁶P. A. Keating, “Phonetic representations in a generative grammar,” *Journal of Phonetics* **18**, 321–334 (1990).