Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Phonetics 31 (2003) 181–201

Journal of Phonetics

www.elsevier.com/locate/phonetics

# Acoustic correlates of breathy and clear vowels: the case of Khmer

Ratree Wayland[a,*], Allard Jongman[b]

[a] Program in Linguistics, 4131 Turlington Hall, University of Florida, P.O. Box 115454, Gainesville, FL 32611-5454, USA
[b] Linguistics Department, Blake Hall, The University of Kansas, Lawrence, KS 66045, USA

Received 19 March 2001; received in revised form 23 October 2002; accepted 15 November 2002

## Abstract

This study investigates acoustic correlates of the putative breathy and clear phonation type contrast in a dialect of Khmer (Cambodian) spoken in Chanthaburi Province, Thailand. The goal is to determine whether this Khmer dialect still preserves this historical contrast. Out of seven acoustic parameters measured, four, namely $^*H_1 - {}^*H_2$, $^*H_1 - A_1$, $^*H_1 - {}^*A_3$, and vowel RMS amplitude successfully distinguished between breathy and clear vowels, with $^*H_1 - {}^*H_2$ measured at the beginning of the vowel being the most robust cue. However, the use of these cues varied from speaker to speaker. The $^*H_1 - {}^*H_2$ measurement obtained from male speakers' production suggested that the contrast being realized may be that of a tense versus lax voice rather than a breathy versus clear voice. It is concluded that the historical breathy and clear phonation distinction in Khmer is preserved among female speakers, but this distinction may be disappearing or have become a tense versus lax distinction among male speakers.

## 1. Introduction

The present phonetic investigation sets out to evaluate phonological claims about historical stages in the development of the Khmer vowel system. Khmer is the national language of Cambodia. It is also the common language of Khmer settlement areas which include the Mekong Delta region of the southern part of Vietnam and at least 12 provinces in the lower part of the northeast and east of Thailand bordering Cambodia (Premsrirat, 1995). The Khmer spoken in the lower part of northeast Thailand is referred to by Smalley as Northern Khmer and has been well studied by a few linguists (e.g., Smalley, 1964, 1976; Jenner, 1974c; Chantharupanth &

---

*Corresponding author. Fax: +1-352-392-8480.
*E-mail address:* ratree@ufl.edu (R. Wayland).

doi:10.1016/S0095-4470(02)00086-4

Phromjagarin, 1978; Phon-ngam, 1987; Thomas & Tienmee, 1990). In contrast, relatively little is known about the Khmer spoken in the east of Thailand.

One of the outstanding problems in the history of Khmer (Cambodian) phonology concerns the existence and development of 'register' or 'phonation types' (i.e., breathy versus clear voice) in the language. A few linguists (e.g., Maspéro, 1915; Jacob, 1960; Jenner, 1974a) have treated this question in the past and proposed reconstruction of stages of development of both Khmer vowels and consonants. However, since Khmer is the only member of the Khmeric branch of Austroasiatic (e.g., Thomas & Headley, 1970; Diffloth, 1974; Huffman, 1976a), traditional comparative reconstruction is not possible. Thus, reconstruction was "primarily based on orthography, and on examination of orthographic developments as revealed in texts dating back to the 7th century" (Huffman, 1978, p. 1). The only alternative method for evaluating the evidence of the writing system is internal reconstruction based on the vowel systems of existing modern Khmer dialects. However, none of the existing dialects investigated so far has kept the 'register' or 'phonation type' constrast, which would be essential evidence for the existence of this phenomenon in the history of Khmer.

One dialect of Khmer that is purported to have kept the original 'breathy' versus 'clear' distinction in voice quality is the one spoken in Chanthaburi Province, Thailand. According to Thongkum (1991), this 'Thung Kabin' Khmer dialect is very conservative and has preserved the lexical contrast between clear (normal, modal) voice and breathy voice that has been lost in other Khmer dialects. Diffloth (1994) has made a similar claim. However, no systematic instrumental phonetic analysis was available to substantiate the claim. Confirmation of the existence of the aforementioned phonation type contrast in this dialect of Khmer would provide crucial evidence for the reconstruction of the history of Khmer phonology.

## 2. History of Khmer phonology

Khmer is a language with a large vowel system: with researchers' estimates ranging from 29 to 33 vowel nuclei (e.g., Henderson, 1952; Headley, Chhor, Lim, Lim, & Chun, 1977; Huffman & Proum, 1977; Wayland, 1998; Wayland & Jongman, 2001). From the point of view of comparative Mon-Khmer, this is the consequence of a historical process of devoicing of consonants which has turned initial voiced stops /b, d, ɟ, g/ into voiceless ones /p, t, c, k/. The older consonantal distinction has been transferred to the following vowels, causing the vowel system to split into two sub-systems, which have variously been termed 'a series' and 'ò series' vowels (e.g., Maspéro, 1915), 'first register' and 'second register', or 'high' and 'low' register vowels (e.g., Huffman, 1967, 1977; Jenner, 1974b; Martin, 1975).

There exists indirect phonological evidence to support the devoicing of initial stops hypothesis in Khmer. Huffman (1976b) examined 15 Austroasiatic languages and found that some languages, especially the Bahnaric, are 'conservative', retaining the original voiced and voiceless stop contrast with little or no effect on the vowels. Some Katuic languages, on the other hand, are 'transitional', retaining a tense-lax distinction in the initials (/p′, t′, c′, k′/ versus /p, t, c, k/) with phonetic differentiation in the vowels as well. The third group of languages including the Monic and some Katuic languages are 'pure registered' languages, with a complete merger of the stops and a complete register (i.e., phonation type) dichotomy in the vowels.

The fourth group including Khmer is what Huffman called ‘restructured’ languages, in which the phonetic and phonological merger of initial stops is complete, with the vowel split reflected by a change in absolute articulatory position and/or diphthongization (Huffman, 1976b). Since the merger of the original voiced and voiceless stops is complete in Khmer, previous attempts to reconstruct the history of Khmer vowels have relied heavily on Khmer orthography and its development in texts dating back to the 7th century (Huffman, 1978). For example, Pinnow (1957) based his reconstruction on the modern writing system. Jacob, on the other hand, based her reconstruction on Pre-Angkor (8th century) and Middle Khmer (16th century) texts (Jacob, 1960, 1963, 1965, 1976a, b, c, 1977). Jenner attempted to date the development of Khmer vowels by examining the rhyme patterns of Middle Khmer texts (Jenner, 1974a, b, 1975, 1976a, b). However, there has been no phonetic evidence to support the existence of the ‘breathy’ and ‘clear’ voice contrast in the history of Khmer vowels. In fact, Huffman (1978) stated that Khmer may never have been a register language in which there is a dichotomy of phonation type throughout the entire vowel system. Therefore, the breathy and clear phonation contrast in Chanthaburi Khmer, if confirmed, would provide direct phonetic evidence for an intermediate stage in the historical reconstruction of the Khmer vowel phonology. In addition to shedding light on the historical development of the Khmer vowel system, acoustic analysis of the breathy vowels in this dialect of Khmer will also contribute to a better understanding of the phonetic implementation of breathy and clear phonation types across languages. Following a detailed survey of the acoustic and perceptual correlates of breathy and clear phonation in Section 3, acoustic measures of the vowels of Chanthaburi Khmer are presented in Section 4. Discussion of the results is presented in Section 5. Finally, implications of the findings for the historical development of Khmer are summarized in Section 6.

## 3. Acoustic and perceptual correlates of breathy and clear phonation

### 3.1. Amplitude of the first harmonic ($H_1$)

Breathiness is thought to be due to incomplete and nonsimultaneous glottal closure during the ‘closed’ phase of the phonatory cycle (Fairbanks, 1940; Zemlin, 1968; Klatt & Klatt, 1990; Hillenbrand, Metz, Colton, & Whitehead, 1990; Hillenbrand, Cleveland, & Erickson, 1994; Hillenbrand & Houde, 1996). Breathy glottal source signals obtained through inverse filtering typically show more symmetrical opening and closing phases with little or no complete closed phase (Fischer-Jorgensen, 1967; Bickley, 1982; Huffman, 1987). The round near-sinusoidal shape of breathy glottal waveforms is responsible for a relatively high amplitude of the first harmonic ($H_1$) and relatively weak upper harmonics (e.g., Bickley, 1982; Huffman, 1987; Klatt & Klatt, 1990; Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). However, as mentioned in Klatt & Klatt (1990), in order to assess whether there is an increase in the $H_1$ amplitude or not, $H_1$ amplitude must be compared with some reference that takes into account recording level such as: (a) RMS amplitude of the vowel; (b) amplitude of the second harmonic ($H_2$) (Bickley, 1982; Hillenbrand et al., 1994; Hillenbrand & Houde, 1996); or (c) amplitude of $F_1$ (Ladefoged, 1982; Ladefoged & Antañanzas-Barroso, 1985; Kirk, Ladefoged, & Ladefoged, 1993).

Enhanced $H_1$ amplitude in the spectra of breathy voice signals has been observed by a number of investigators (e.g., Fischer-Jorgensen, 1967; Bickley, 1982; Ladefoged, 1981, 1983; Huffman, 1987; Klatt & Klatt, 1990; Kirk et al., 1993). Fischer-Jorgensen (1967) used a variety of techniques to study the acoustic characteristics of Gujarati murmured vowels. She considered the high intensity of $H_1$ to be the most salient spectral feature of murmured vowels. However, despite her belief that the relative amplitude of $H_1$ was "the most obvious and constant feature" (pp. 133–134), Fischer-Jorgensen concluded that no single acoustic feature was sufficient to produce the perception of breathiness.

A spectral analysis of $H_1$ amplitude relative to that of $H_2$ for !Xóõ and Gujarati vowels by Bickley (1982) revealed that the amplitude of $H_1$ was higher than the amplitude of the adjacent $H_2$. Results from a perceptual study using synthetic Gujarati breathy and clear word pairs varying in aspiration noise and $H_1$ amplitude showed that identification of the stimuli as breathy or clear by native Gujarats was affected by $H_1$ amplitude only, with no effect of aspiration noise. However, as noted by Klatt & Klatt (1990), the 15 dB $H_1$ enhancement that was needed to affect the decisive shift from clear to breathy in Bickley's study greatly exceeded measured $H_1$ amplitude differences between naturally produced breathy and clear word pairs (6 dB for Gujarati and 9.7 dB for !Xóõ data). Thus, $H_1$ amplitude may not be the sole cue given that it had to be exaggerated to achieve consistent responses from listeners (Klatt & Klatt, 1990).

Ladefoged (1982) also reported enhanced $H_1$ amplitude for breathy vowels in !Xóõ. Similar to Bickley (1982), a follow-up study by Ladefoged & Antañanzas-Barroso (1985) found that the breathiness judgements of American listeners were more strongly correlated with $H_1$ amplitude than aspiration noise. Kirk et al. (1993) reported that in Jalapa Mazatec, the difference between the amplitude of $H_1$ and that of $F_1$ successfully identifies breathy voice: the value for breathy voice was higher than that for modal voice for all five speakers. Huffman (1987) used inverse filtering to derive glottal waveforms from samples of four phonation types used in Hmong. Breathy (murmured) samples showed stronger first harmonics than nonbreathy samples.

### 3.2. *Additive noise*

When a portion of the air stream from the lungs passes through a persistent and relatively narrow glottal chink during the production of breathy vowels, this results in the generation of noise (Klatt & Klatt, 1990; Hillenbrand et al., 1994; Stevens, 2000). The spectrum becomes dominated by dense aspiration noise, particularly at high frequencies where noise may actually replace harmonic excitation of the third and higher formants (Ladefoged & Antañanzas-Barroso, 1985; Klatt, 1986).

Based on spectrographic observations, Fischer-Jorgensen (1967) noticed small but inconsistent differences in spectral noise between murmured and clear vowels in Gujarati. Similarly, Bickley (1982) found no correlation between breathiness ratings and increases in additive noise in synthetic vowels. In contrast, Hillenbrand (1988) found a strong relationship between breathiness ratings and additive noise using synthetic stimuli.

To isolate and estimate the relative strength of noise components of [ha] samples, Klatt and Klatt (1990) used a bandpass filter centered at $F_3$. The signals were judged to be unsuitable for noise estimation due to the fact that the greater energy of the periodic component at low frequencies tends to dominate the visual impression. The degree of periodicity in the band-limited signals was judged by visual inspection of time-domain waveforms using a five-point rating scale. The noise rating accounted for approximately 60% of the variance in listener ratings of breathiness. A follow-up study using synthetic stimuli found that increases in spectral noise were the single most important cue to perceived breathiness.

Another method for calculating a spectral harmonics-to-noise ratio (HNR) in speech signals was proposed by de Krom (1993). This harmonics-to-noise ratio algorithm used a comb-filter defined in the cepstral domain to separate the harmonics from the noise. The sensitivity of the HNR to additive noise and jitter was tested with synthetic vowel-like signals, generated at 10 fundamental frequencies. Results showed a major effect of both noise and jitter on HNR: HNR decreased almost linearly with increasing noise levels or increasing jitter. Sensitivity of de Krom's (1993) HNR to both noise and jitter makes it a valid method for determining the amount of spectral noise.

De Krom's HNR algorithm reliably distinguished breathy from clear vowels in 31 minimal breathy/clear word pairs produced by two male and two female speakers of Javanese, with higher HNRs for clear than breathy tokens (Wayland, Gargash, & Jongman, 1994). A preliminary perception experiment, however, revealed that for English native speakers, degree-of-breathiness ratings only moderately correlated with HNRs.

Hillenbrand et al. (1994) evaluated the effectiveness of several acoustic measures in predicting breathiness ratings for sustained vowels spoken by healthy native English speakers who were asked to produce nonbreathy, moderately breathy, and very breathy phonation. Hillenbrand et al. (1994) reported that periodicity measures obtained from two different methods provided the most accurate predictions of perceived breathiness, accounting for approximately 80% of the variance in breathiness ratings. These methods were: (a) Cepstral Peak Prominence (CPP), a measure of peak amplitude normalized for overall amplitude; (b) Pearson $r$ at autocorrelation Peak (RPK), a measure of the degree of correlation in intensity between adjacent pitch pulses of the signal. The relative amplitude of $H_1$ correlated only moderately with breathiness ratings and two measures of spectral tilt correlated weakly with perceived breathiness. Hillenbrand & Houde (1996) found the same results in a study of dysphonic voice.

Besides the periodicity or additive noise measures discussed above, there are also measures of jitter (e.g., Lieberman, 1963) and shimmer (e.g., Horii, 1980). Accurate measurement of jitter is, however, difficult to achieve due to the influence of low-frequency signal components that tend to dominate waveform characteristics because of their relatively high energy (Klatt & Klatt, 1990), and the difficulty in reliably and automatically identifying the oscillographic landmarks used by these measures in noisy voice signals (Hillenbrand et al., 1994).

### 3.3. *Spectral tilt*

The relatively more symmetrical or near-sinusoidal shape of breathy glottal waveforms does not only boost the lower harmonics, it is also responsible for a decrease in the amplitude of the harmonics in the higher frequency region, or degree of spectral tilt. According to Ní Chasaide & Gobl (1997), the more symmetrical the glottal pulse, the steeper the spectral tilt. However, as pointed out by Hillenbrand & Houde (1996), while the harmonic (periodic) component of breathy signals "tends to be relatively weak in high-frequency energy, the presence of aspiration noise, which is stronger in the mid and high frequencies than in the lows, can result in a voice signal that is richer in high-frequency energy than nonbreathy signals" (p. 312). In other words, the overall spectral balance of breathy signals is affected by two separate phenomena acting in opposite directions, i.e., glottal rounding acts to increase energy at the low-frequency end of the spectrum, while aspiration noise tends to increase energy at the mid- and high-frequency regions. As a result, it is difficult to determine a priori the direction of the spectral tilt for breathy and nonbreathy signals when both the periodic and nonperiodic components are combined. Moreover, several spectral tilt measures have been used in the literature. While some of these measures may effectively capture both the glottal rounding and the aspiration effects (the spectral balance) on breathy signals, others may reflect only the effect of glottal rounding. There are yet others that cannot be easily interpreted in these terms.

Mixed results have been reported on the relationship between breathiness ratings and spectral tilt measures designed to capture the spectral balance of a breathy signal. Klich (1982), for example, reported a strong correlation between perceived degree of breathiness and several measures of spectral tilt calculated as the energy ratio of low-, mid-, and high-frequency bands. However, Hillenbrand et al. (1994) found that two measures of spectral tilt (i.e., breathiness index and ratio of high-to mid/low-frequency energy) correlated only weakly with breathiness ratings of sustained vowels. On the other hand, a strong correlation between breathiness rating of these two measures of spectral tilt were reported for sentences (the Rainbow Passage) in Hillenbrand & Houde (1996).

Klatt & Klatt (1990) found no significant correlation between breathiness ratings and two measures of spectral tilt: (a) amplitude of $F_1$ in dB, relative to amplitude of $F_2$, obtained at the beginning, middle, and end of the vowel; (b) amplitude of $F_3$, $F_4$, and $F_5$ (whichever is the greatest) in dB, relative to the amplitude of $F_2$, obtained at the beginning, middle, and end of the vowel. Furthermore, using synthetic stimuli, Hillenbrand (1988) found that breathiness ratings were affected only by the level of aspiration noise, with no effect of spectral tilt. In Stevens and Hanson (1994) and Hanson (1995), on the other hand, the amount of spectral tilt is defined as the difference in dB between the amplitude of $H_1$ and the amplitude of the most prominent harmonic in the $F_3$ region. An algorithm to correct potential effects of $F_1$ and $F_2$ on the spectrum amplitude of $F_3$ is required (see Hanson, 1995). The positive values (downward slope) of $H_1 - A_3$ obtained in these studies suggested that this measure reflected the effect of glottal rounding on the signal. A high correlation between this tilt measure and two methods of noise judgement using native speakers of American English was reported in Hanson (1995).

### 3.4. *Tracheal coupling*

Another potential cue to breathiness in the vowel spectrum in acoustic coupling with the trachea (Fant, Ishizaka, Lindqvist, & Sundberg, 1972; Klatt, 1986). The acoustic effects of tracheal coupling on the normal transfer function of the vocal tract for a vowel include: (a) possible addition of poles (formants) and zeros associated with the tracheal and lung systems below the glottis; (b) increased losses at the glottal termination, which primarily affect $F_1$ bandwidth (Klatt & Klatt, 1990).

Bandwidth is related to the rate of energy loss in the vocal tract (Stevens & Hanson, 1994). The resistance of the yielding walls of the vocal tract, and heat conduction and frictional losses at the walls are among several sources of energy loss in the frequency range of $F_1$. When there is airflow through the open glottis, the resistance of the glottis can contribute to energy loss, and add significantly to $F_1$ bandwidth (Stevens & Hanson, 1994). The bandwidth of $F_1$ of the transfer function determines the width of the resonance peak and the relative strength of the $F_1$ peak in the acoustic output. Thus, measurements of $F_1$ bandwidth can provide an indirect indication of degree of glottal opening (Stevens & Hanson, 1994). However, Klatt & Klatt (1990) reported no correlation between degree of breathiness ratings and two indirect measures of $F_1$ bandwidth: (a) amplitude in dB of $F_1$ relative to amplitude of $F_2$, as measured at the beginning, middle, and the end of the vowel; (b) a three-point scale estimate of degree of visibility of the first formant peak. An alternative estimate of the prominence of the $F_1$ peak is obtained by measuring the amplitude difference in dB between $H_1$ and the most prominent harmonic in the $F_1$ region (Stevens & Hanson, 1994).

### 3.5. *Fundamental frequency ($F_0$)*

During the production of breathy phonation, to allow the vocal folds to vibrate while they stay relatively far apart, the vocal folds have to be relatively less taut. Thus, the fundamental frequency of a breathy vowel is expected to be lower than that of a clear vowel. This expectation was borne out in Javanese (Wayland et al., 1994) and Green Mong (Andruski & Ratliff, 2000). This may also explain why breathy phonation appears to be consistently associated with lowered tone in many languages reviewed by Hombert, Ohala, & Ewan (1979).

### 3.6. *Acoustic intensity*

Cross-linguistic investigations of phonation types generally show that breathy phonation is associated with a decrease in overall acoustic intensity in many languages including Gujarati (Fischer-Jorgensen, 1967), Kui, and Chong (Thongkum, 1988). This finding is, however, not universal. Wayland et al. (1994), for example found that breathy vowels in Javanese are associated with an increase in overall acoustic intensity.

### 3.7. *Duration*

In some languages, nonmodal phonation types are associated with increased vowel duration (Gordon & Ladefoged, 2001). For example, breathy vowels are longer than clear vowels in Kedang (Samely, 1991), and Jalapa Mazatec (Kirk et al., 1993; Silverman, Blankenship, Kirk, & Ladefoged, 1995), and Javanese (Wayland et al., 1994). This trend is, however, not found in Hmong (Huffman, 1987) and Lucas Quiaviní Zapotec (Gordon & Ladefoged, 2001).

### 3.8. *Summary of previous studies*

In summary, previous studies on phonation suggest the possibility of a conglomerate of cues to the production and perception of breathy phonation. Variations in acoustic cues associated with a breathy phonation may stem from ways in which this phonation is articulatorily manifested in different languages or even different speakers in the same language. As pointed out by Hanson (1995), some speakers may adjust their glottal configuration in such a way that a larger open quotient results while rate of decrease of airflow at glottal closure remains nearly the same. Thus, the difference between $H_1$ and $H_2$ increases, but the tilt stays nearly the same or changes only a small amount due to a change in the skewness of the glottal pulse. This may explain why some languages or some speakers' production of a breathy phonation may show either an increase in the differences between $H_1$ and $H_2$, or spectral tilt, but not both.

Variations in perceptual cues associated with a breathy phonation may arise from several factors including listeners' linguistic background, or relative experience, and the use of natural versus synthetic stimuli. As pointed out by Kreiman & Gerratt (1998) "voice quality is an interaction between an acoustic voice stimulus and a listener; the acoustic signal itself does not possess vocal quality, it evokes it in the listener" (p. 1598).

In summary, the literature reviewed above suggests that a conglomerate of cues may convey the breathy and clear phonation distinction. Besides spectral noise, dynamic spectral cues, namely the difference in amplitude between the first two harmonics ($H_1 - H_2$), the first harmonic and the most prominent harmonic in the $F_1$ region ($H_1 - A_1$), and the difference in amplitude between the first harmonic and the most prominent harmonic in the $F_3$ region ($H_1 - A_3$) emerge as the most likely cues.

## 4. Acoustic measurements

In this section, an acoustic analysis of Chanthaburi Khmer vowels is presented. All acoustic and perceptual correlates reviewed above, namely amplitude of the first harmonic, additive noise, spectral tilt, and tracheal coupling were examined. Moreover, vowel duration, vowel RMS amplitude, and vowel fundamental frequency ($F_0$) were also measured.

### 4.1. Speakers

Five (three female and two male) native speakers of Chanthaburi participated in this study. They were between 26 and 64 years of age. The female and male speakers were designated W1, W2, W3, and M1 and M2, respectively. Besides Khmer, all participants also speak Thai and were exposed to Isaan (a Laotian dialect spoken in Northeastern Thailand) and Chong (a language in the Pearic branch of the Austroasiatic language family). To assure that all participants were native Khmer speakers, only participants whose ancestors of three or more generations were Khmer are included. Their spouses were also Khmer and Khmer was the main language spoken in the home.

### 4.2. Stimuli

Stimuli consisted of 23 (near) minimal pairs[1] of breathy and clear vowels (see the appendix). The word list was constructed based on the knowledge of the oldest female speaker (W1, 64 years old). A few words on the original wordlist that were not recognized by some younger speakers were replaced by words of similar syllable structures. As a result, the wordlist varied slightly for each subject.

---

[1] It was not possible to get only minimal pairs. With the exception of /ɛ/, /o/, and /a/, the distinction between the so-called 'first or clear' and 'second or breathy' register vowels in this dialect of Khmer is accompanied by a difference in vowel height and/or diphthongization. 'Clear' vowels are usually lower (and/or diphthongized) than 'breathy' vowels.

Since speakers of this Khmer dialect were Thai-Khmer bilingual and illiterate in Khmer, they were asked to say the Khmer words based on the definition given in Thai by the first author. While the majority of the target words were produced in isolation, some words were produced as a phrase or a compound. For example, the speakers insisted that they added the word for 'hair' to the target word /cɛ..ːk/ 'to part (hair)' to make it sound more natural. The benefit of obtaining a natural production of target words outweighed the risk of a (yet-to-be studied) coarticulatory effect of the phonation of neighboring vowels on the phonation of the target vowel. Every word was repeated twice and recorded using a high-quality cassette recorder (Marantz, Model PMD 222) and microphone (AKG D310). All recordings took place in the informants' homes.

### 4.3. Procedure

Recordings of the wordlist were digitized on a Sun Sparc station LX at 11 kHz at the Phonetics Laboratory, Cornell University. Each word was stored as a separate file to be processed by the commercial software package ESPS/WAVES +. The beginning and end of the target vowel of each word was marked by examining both waveforms and wide-band spectrograms. Vowel onset was taken to be the onset of periodicity in the waveform. Vowel offset was indicated by the loss of the second formant ($F_2$) on the spectrogram. Cursors were automatically placed at three other locations in the vowel, namely at 30%, 50% and 70% in the vowel.

An autocorrelation method was used to compute $F_0$ values at 5 ms intervals. Vowel $F_0$ was then obtained by averaging these $F_0$ values from the beginning to the end of the vowel. Overall RMS amplitude of the vowel was computed from the digitized waveforms and was averaged for the entire vowel. Formant frequencies ($F_1$, $F_2$ and $F_3$) to be used in the algorithms to correct the effect of $F_1$ amplitude on $H_1$ and $H_2$ amplitudes, and of $F_2$ and $F_3$ amplitudes on $A_3$ amplitude were measured from LPC spectra using a Hamming window of 25.6 ms with eight poles and pre-emphasis of 0.98.[2] $H_1$ and $H_2$ amplitudes were taken from a DFT spectrum using the same parameters.[3] Amplitudes of the most prominent harmonic in the $F_1$ and $F_3$ regions were measured using DFT spectra, supplemented by LPC spectra to ensure accurate locations of $F_1$ and $F_3$ (see Fig. 1).

### 4.4. Acoustic parameters measured

Breathy voice can be quantified through a number of phonetic measurements and cross-linguistic investigation of the realization of different phonations has revealed both similarities and

---

[2] We thank James Hillenbrand for pointing out that LPC may not be appropriate for signals generated with tracheal coupling. There may, however, be no better alternative. A precise estimate of formant frequency values using spectrograms is difficult and inaccurate since a signal generated with tracheal coupling is likely to exhibit a less prominent $F_1$ peak, with a wide bandwidth. Moreover, a relatively high degree of aspiration noise in a breathy signal makes it difficult for higher formants ($F_2$, $F_3$) to be visually discerned and measured with any accuracy.

[3] We are grateful to Peter Ladefoged for pointing out that pre-emphasis should not be used when computing spectra in which a comparison of the amplitudes of the harmonics is to be made since it provides the spectrum with a nonlogarithmic slope. However, since the same degree of pre-emphasis was used to derive spectra of both clear and breathy vowels, a comparison of measurements obtained for both types of vowels, which is the goal of the study, remains viable.

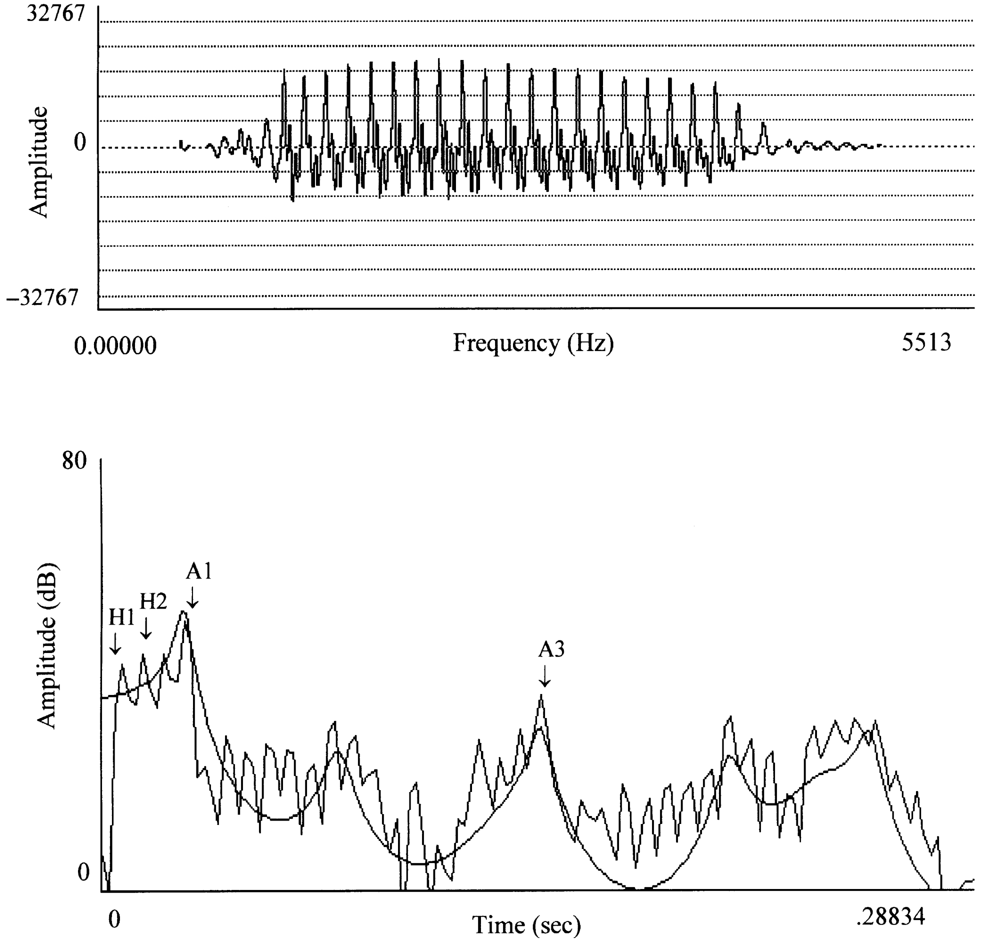

Fig. 1. The top panel shows the waveform of the word [pə̤ːp] "to run into someone" spoken by speaker W1. The bottom panel shows LPC (smooth curve) and DFT (jagged curve) spectra derived at approximately the middle of the vowel and illustrates how the amplitude of $H_1$ and $H_2$, as well as the most prominent harmonic in the $F_1$ and $F_3$ regions ($A_1$ and $A_3$), respectively, were obtained.

differences across languages (Gordon & Ladefoged, 2001). Therefore, in order to accurately assess whether or not the vowel system of this dialect of Khmer has preserved the earlier breathy and clear phonation contrast, it is necessary that their acoustic characteristics be thoroughly examined. Consequently, acoustic measurements were taken at different locations in the vowel as well as at different frequency ranges.

The following acoustic parameters of the vowels were measured:

1. $H_1 - H_2$ (an indicative measure of open quotient) at 30%, 50% and 70% in the vowel.
2. $H_1 - A_1$ (an indicative measure of $F_1$ bandwidth) at 30%, 50% and 70% in the vowel.
3. $H_1 - A_3$ (a measure of spectral tilt) at 30%, 50% and 70% in the vowel.
4. Harmonics-to-noise ratio using de Krom's (1993) algorithm.
5. RMS amplitude.

Table 1
Mean difference between first- and second-harmonic amplitude ($^*H_1 - {}^*H_2$ in dB) at 30%, 50% and 70% in the vowel for all five speakers

| | Breathy vowels | | | Clear vowels | | |
|---|---|---|---|---|---|---|
| Speaker | 30% | 50% | 70% | 30% | 50% | 70% |
| W1 | 2.70 | 2.17 | 1.74 | −0.57 | 0.13 | −0.01 |
| W2 | 2.65 | 3.25 | 1.66 | 0.13 | 0.35 | 0.74 |
| W3 | 6.13 | 6.54 | 7.02 | 2.3 | 3.22 | 4.1 |
| M1 | −0.83 | −1.08 | −0.65 | −3.16 | −3.07 | −3.01 |
| M2 | −7.36 | −8.23 | −8.79 | −9.75 | −10.51 | −10.14 |

6. Fundamental frequency ($F_0$).
7. Duration.

### 4.5. Results

#### 4.5.1. $H_1$–$H_2$

To assess the relative amplitude of $H_1$ of Khmer Chanthaburi vowels, the amplitude of the second harmonic ($H_2$) was used as a reference. $H_1 - H_2$ also seems to be an appropriate measurement of open quotient (Hanson, 1995). The difference between the amplitude of $H_1$ and $H_2$ of a breathy vowel is expected to be greater than that of a clear vowel. However, because of the potential 'boosting' effect of $F_1$ on the amplitude of $H_1$ and $H_2$ due to the proximity of $F_1$ and $H_1$ and $H_2$, especially among high vowels, the raw values obtained for the amplitude of $H_1$ and $H_2$ were corrected using the algorithm given in Hanson (1995), yielding a normalized measure named $^*H_1 - {}^*H_2$.[4]

Mean $^*H_1 - {}^*H_2$ values at three different locations in the vowel for all five speakers are reported in Table 1. As expected, among female speakers, $^*H_1 - {}^*H_2$ amplitudes are greater for breathy vowels than for clear vowels at all three locations. However, the negative values in the male speakers' data indicate that the amplitude of the $H_2$ is consistently greater than that of $H_1$, and that the degree of $H_2$ prominence is greater among clear vowels than breathy vowels. According to Ní Chasaide & Gobl (1997), a relatively strong $H_2$ can indicate tense or creaky voice. The male speakers' data thus suggest that clear vowels may be more tense than breathy vowels. This is also true for $^*H_1 - {}^*H_2$ measured at 30% and 70% in the vowel for W1.

Two-way ANOVAs with Phonation Type and Location in the vowel as independent variables performed on the data obtained for individual speakers yielded a significant main effect of Phonation Type, but not of Location for all speakers with $F(1,82)$ ranging from 13.08 to 25.35, and $p$ values ranging from 0.0007 to 0.0001. No significant interaction was obtained. Tests of

---

[4] The algorithm's requirement that twice the value of $F_0$ not exceed the value of $F_1$ renders it impossible to obtain the corrected $^*H_1$ and $^*H_2$ values for high vowels as well as diphthongs with high vowels as the first element. Thus, only 15 out of the original 23 minimal pairs were included in the analysis. This limitation also applies to the $^*H_1 - A_1$ and $^*H_1 - A_3$ measurements.

simple effects of Phonation Type suggested that all speakers distinguished breathy and clear phonation type on the basis of $^*H_1 - {}^*H_2$ measured at 30% in the vowel [$F(1,84)$ ranging from 6.2 to 11.03, $p$ values ranging from 0.01 to 0.0007], while $^*H_1 - {}^*H_2$ measured at the center and at the end of the vowel did so only for some speakers. For example, W2, W3, M1 and M2 distinguished breathy and clear phonations on the basis of $^*H_1 - {}^*H_2$ measured at the center of the vowel ($F(1,84)$ ranging from 5.60 to 13.41, $p$ values ranging from 0.02 to 0.0004), while $^*H_1 -{}^* H_2$ measured at the end of the vowel separated the two phonation types for W3 and M1 only, [$F(1,84) = 6.40$ and 9.48, $p<0.01$ and 0.002, respectively].

### 4.5.2. $F_1$ bandwith ($^*H_1 - A_1$)

Following Hanson (1995), a difference between the amplitude of $H_1$ and the amplitude of the most prominent harmonic in the $F_1$ region ($^*H_1 - A_1$) was used as an indirect measure of $F_1$ bandwidth. $^*H_1 - A_1$, reflecting $F_1$ bandwidth, was expected to be greater for a breathy vowel than for a clear vowel.

Mean differences of $^*H_1$ and $A_1$ amplitudes at 30%, 50% and 70% in the vowel are reported in Table 2. Negative values of the data indicate that the amplitude of $A_1$ is greater than that of $^*H_1$. According to Hanson (1995), a relatively stronger or more prominent $F_1$ peak (or greater $A_1$) indicates a narrower $F_1$ bandwidth. From these data, it can be seen that clear vowels have a narrower $F_1$ bandwidth than breathy vowels. The $F_1$ bandwidth also appears to increase toward the end (70%) of the vowels. This is true for both breathy and clear vowels. However, two-way ANOVAs performed on individual speakers' data resulted in a significant main effect of Phonation Type, but not of Location for every speaker [$F(1,84)$ ranging from 5.8 to 54.57, and $p$ values ranging from 0.02 to 0.0001]. No significant interaction was obtained. Tests of simple effects of Phonation Type revealed that the $^*H_1 - A_1$ value that distinguished breathy and clear vowels occurred at different locations in the vowel for different speakers. Speakers W2 and W3 distinguished the two phonation types on the basis of this parameter at all three locations in the vowel [$F(1,84)$ ranging from 5.35 to 20.53, $p$ values ranging from 0.02 to 0.0001]. $^*H_1 - A_1$ measured at the 30% and 50% in the vowel served to separate between breathy and clear vowels for W1 [$F(1,84) = 13.5$ and 10.26, $p<0.0004$ and 0.002, respectively]. However, only $^*H_1 - A_1$ measured at 30% distinguished breathy vowels from clear vowels for M2, $F(1,84) = 8.8$, $p<0.004$, and only $^*H_1 - A_1$ measured at the end of the vowel did so for M1, $F(1,84) = 4.30$, $p<0.04$.

Table 2
Mean difference between first harmonic and the most prominent harmonic in the $F_1$ region ($^*H_1 - A_1$ in dB) at 30%, 50% and 70% in the vowel for all five speakers

| | Breathy vowels | | | Clear vowels | | |
|---|---|---|---|---|---|---|
| Speaker | 30% | 50% | 70% | 30% | 50% | 70% |
| W1 | −4.1 | −3.9 | −4.6 | −8.5 | −7.7 | −6.5 |
| W2 | −6.6 | −6.2 | −5.6 | −11.7 | −12.0 | −11.0 |
| W3 | −5.5 | −7.2 | −7.6 | −11.1 | −11.8 | −10.8 |
| M1 | −12.0 | −12.9 | −11.4 | −13.7 | −13.9 | −14.0 |
| M2 | −13.4 | −14.3 | −14.7 | −17.3 | −16.8 | −16.2 |

Table 3
Mean difference between first harmonic and the most prominent harmonic in the $F_3$ region ($^*H_1 - {}^*A_3$ in dB) at 30%, 50% and 70% in the vowel for all five speakers

| | Breathy vowels | | | Clear vowels | | |
|---|---|---|---|---|---|---|
| Speaker | 30% | 50% | 70% | 30% | 50% | 70% |
| W1 | −4.2 | 10.6 | 10.2 | 7.2 | 8.0 | 7.5 |
| W2 | 11.2 | 14.2 | 12.8 | 5.9 | 8.4 | 9.8 |
| W3 | 11.2 | 11.0 | 12.0 | 3.6 | 4.5 | 7.2 |
| M1 | 0.27 | 1.1 | 3.2 | 3.3 | 2.1 | 4.4 |
| M2 | −9.0 | −0.5 | −8.5 | −10.3 | −8.8 | −8.5 |

### 4.5.3. Spectral tilt

Following Stevens & Hanson (1994) and Hanson (1995), the difference in dB between the amplitude of $H_1$ and the most prominent harmonic in the $F_3$ region was taken as a measure of spectral tilt. The amplitude of the most prominent harmonic in the $F_3$ region ($A_3$) was corrected for potential boosting effects of the first and second formant, using an algorithm given in Hanson (1995). The average difference between the first harmonic and the most prominent harmonic in the $F_3$ region ($^*H_1 -^* A_3$) was expected to be greater for breathy vowels than for clear vowels.

Mean $^*H_1 -^* A_3$ values for all three locations in the vowel are presented in Table 3. Negative values indicate that the amplitude of $H_1$ is lower than that of $A_3$. According to Hanson (1995), when the amplitude of $H_1$ is low, the difference between $H_1$ and $A_3$ amplitudes is no longer a good measure of spectral tilt. For this reason, average $^*H_1 -^* A_3$ values at 30% in the vowel for speaker W1 and at all three locations in the vowel for speaker M2 were excluded from further analysis. The missing data at 30% in the vowel for speaker W1 necessitated the elimination of the data from this location in the vowel for all remaining speakers.

Results of two-way ANOVAs performed on individual speakers' data revealed a significant main effect of Phonation Type for speakers W2 and W3 only, [$F(1,56) = 4.49$, and 7.63, $p<0.04$ and 0.001 respectively]. No significant main effect of Location was found for any speaker and there was no significant interaction. Tests of simple effects of Phonation Type suggested that the difference between breathy and clear vowels along this dimension was significant for W3 at 50% in the vowel [$F(1,56) = 5.10, p<0.03$], but only marginally so for W2 [$F(1,56) = 3.9, p > 0.053$]. Neither speaker conveyed the distinction between the two phonations along this dimension at the end of the vowel.

### 4.5.4. Additive noise

To measure the ratio between energy caused by harmonic excitation and by noise in the waveform of breathy and clear vowels of Chanthaburi Khmer, the harmonics-to-noise ratio algorithm developed by de Krom (1993) was used. The harmonics-to-noise ratio (HNR) of breathy vowels should be relatively low compared to that of clear vowels. The relatively lower HNR of a breathy vowel is expected to occur around the $F_3$ or higher formant frequencies. However, since different vowel qualities with varying $F_3$ frequencies were used in this analysis, and since additive noise may be present at both relatively low and relatively high frequencies in Khmer, HNRs were measured at seven intervals from 60 to 5000 Hz.

Table 4
Mean HNRs at seven different frequency ranges of breathy and clear vowels for all five speakers

| Speaker | 60–5000 Hz | 60–1000 Hz | 500–1000 Hz | 1–2 kHz | 2–3 kHz | 3–4 kHz | 4–5 kHz |
|---|---|---|---|---|---|---|---|
| Breathy vowels | | | | | | | |
| W1 | 17.0 | 31.5 | 40.6 | 26.4 | 18.6 | 12.9 | 11.1 |
| W2 | 20.9 | 33.0 | 43.4 | 30.3 | 24.1 | 19.0 | 14.3 |
| W3 | 21.6 | 33.8 | 42.6 | 32.3 | 26.4 | 20.3 | 15.4 |
| M1 | 19.1 | 36.4 | 44.2 | 31.1 | 20.6 | 14.0 | 10.5 |
| M2 | 20.0 | 33.7 | 41.4 | 29.3 | 22.8 | 17.1 | 15.0 |
| Clear vowels | | | | | | | |
| W1 | 15.6 | 28.0 | 37.4 | 24.4 | 16.5 | 11.8 | 10.0 |
| W2 | 19.9 | 31.3 | 40.7 | 29.0 | 23.0 | 17.9 | 13.7 |
| W3 | 22.2 | 32.1 | 39.9 | 32.4 | 27.4 | 21.4 | 15.7 |
| M1 | 19.1 | 34.0 | 40.8 | 29.0 | 21.1 | 14.8 | 11.1 |
| M2 | 19.4 | 33.2 | 42.1 | 29.0 | 22.2 | 16.3 | 13.6 |

Table 4 shows mean HNR values at all seven frequency ranges for breathy and clear vowels for all five speakers. Unexpectedly, HNRs are higher for breathy vowels than clear vowels for all five speakers. Two-way ANOVAs performed on each speaker's data yielded a significant main effect of Phonation Type for speaker W1 only, [$F(1,308) = 11.33, p < 0.0009$]. A significant main effect of Frequency Range, on the other hand, was obtained for every speaker [$F(6,308)$ ranging from 78.58 to 209.73, $p < 0.0001$]. There was no significant interaction. Tests of simple effects of Phonation Type for W1 suggested that the HNR value measured at the 60–1000 Hz range was significantly higher for breathy vowels than for clear vowels [$F(1,308) = 4.72, p < 0.03$], while HNR measured at 500–1000 Hz showed a marginal effect [$F(1,308) = 3.78, p > 0.052$]. Tests of simple effects of Frequency Range were significant for both clear vowels and breathy vowels for all speakers [$F(6,308)$ ranging from 36.62 to 119.40, $p$ values ranging from 0.001 to 0.0001].

### 4.5.5. Vowel RMS amplitude

Mean RMS amplitude of breathy and clear vowels for all speakers are reported in Table 5. Averaged across speakers, breathy vowels (77 dB) had greater intensity than clear vowels (75 dB). Results of one-way ANOVAs performed on each speaker's data revealed that all speakers, except speaker W2 produced breathy vowels with a significantly greater RMS amplitude than clear vowels [$F(1,44)$ ranging from 5.09 to 21.37, $p$ values ranging from 0.03 to 0.0001]. A strong trend in the same direction was found for W2 [$F(1,44) = 3.81, p > 0.057$].

### 4.5.6. Vowel average $F_0$

Mean $F_0$ for both breathy and clear vowels for all five speakers is presented in Table 6. With the exception of speaker W2, all speakers produced breathy vowels (157 Hz) with higher $F_0$ than clear vowels (153 Hz). Given the universal trend mentioned earlier, this result is unexpected. This difference reached significance for W1 and W2 [$F(1,44) = 7.63$ and 7.24, $p < 0.008$ and 0.01 respectively].

Table 5
Mean RMS amplitude and standard deviations (in parentheses), in dB, of breathy and clear vowels for all five speakers

| Speaker | Breathy vowels | Clear vowels |
|---|---|---|
| W1 | 76 (2) | 73 (4) |
| W2 | 78 (4) | 76 (3) |
| W3 | 73 (2) | 75 (3) |
| M1 | 77 (3) | 74 (5) |
| M2 | 79 (1) | 78 (1) |

Table 6
Mean $F_0$ and standard deviations (in parentheses), in Hz, of breathy and clear vowels for all five speakers

| Speaker | Breathy vowels | Clear vowels |
|---|---|---|
| W1 | 135 (6) | 130 (6) |
| W2 | 185 (9) | 176 (15) |
| W3 | 202 (7) | 201 (8) |
| M1 | 156 (12) | 152 (9) |
| M2 | 108 (4) | 107 (5) |

Table 7
Mean duration and standard deviations (in parentheses), in ms, of breathy and clear vowels for all five speakers

| Speaker | Breathy vowels | Clear vowels |
|---|---|---|
| W1 | 207 (89) | 186 (81) |
| W2 | 158 (75) | 151 (62) |
| W3 | 173 (64) | 161 (60) |
| M1 | 164 (72) | 136 (57) |
| M2 | 173 (74) | 161 (55) |

### 4.5.7. *Vowel duration*

Mean duration of breathy and clear vowels for all five speakers is presented in Table 7. All five speakers produced breathy vowels (175 ms) with longer duration than clear vowels (159 ms). This finding is in agreement with previous findings (e.g., Wayland et al., 1994; Andruski & Ratliff, 2000). However, the difference did not reach significance for any of the speakers.

In summary, results of the acoustic analyses just presented suggest that certain acoustic parameters measured (e.g., $^*H_1 -^* H_2$, $^*H_1 - A_1$, vowel RMS amplitude) are more successful than others (e.g., HNRs, vowel duration, vowel average $F_0$) in differentiating between breathy and clear vowels of Chanthaburi Khmer. Moreover, it is obvious that the use of these acoustic cues varied from speaker to speaker.

## 5. Discussion

Our discussion of the acoustic analyses is based on the results for individual speakers. Individual speakers' phonetic implementations of breathy and clear vowels along all acoustic parameters measured are summarized in Table 8. As can be seen, ${}^*H_1 -{}^* H_2$ consistently differentiates between breathy and clear vowels for all speakers. Recall, however, that for female speakers, amplitude of $H_1$ is always greater than that of $H_2$ and the difference between $H_1$ and $H_2$ amplitude is greater (suggesting a higher degree of open quotient) in breathy vowels than in clear vowels. For male speakers, on the other hand, $H_2$ amplitude is greater than $H_1$ amplitude and the difference between $H_1$ and $H_2$ amplitude is greater for clear vowels than for breathy vowels. If $H_2$ prominence indicates a tense or creaky voice, the present findings suggest that the distinction made by the male speakers is that of a tense versus lax voice rather than a breathy versus clear voice.

The table also shows that ${}^*H_1 -{}^* H_2$ measured at the beginning of the vowel is more robust than ${}^*H_1 -{}^* H_2$ measured at the center, and especially at the end of the vowel in differentiating between breathy and clear phonations in that it distinguishes breathy and clear vowels for all speakers. However, individual differences exist. ${}^*H_1 -{}^* H_2$ measured at all three locations in the vowel distinguishes between breathy and clear vowels for W3 and M1, while only ${}^*H_1 -{}^* H_2$ measured at vowel onset and at vowel center separate the two phonations for W2 and M2. Moreover, only ${}^*H_1 -{}^* H_2$ measured at vowel onset distinguishes between the two phonation types for W1.

${}^*H_1 - A_1$ is another cue that successfully distinguishes between breathy and clear vowels. Again, ${}^*H_1 - A_1$ measured at vowel onset appears to be more robust than ${}^*H_1 - A_1$ measured at other locations. It separates breathy from clear vowels for four out of five speakers. Three out of

Table 8
Summary of the acoustic analyses for each speaker

| Acoustic parameters | W1 | W2 | W3 | M1 | M2 |
|---|---|---|---|---|---|
| ${}^*H_1 - {}^*H_2$ at 30% | ** | ** | ** | ** | ** |
| ${}^*H_1 - {}^*H_2$ at 50% | – | ** | ** | ** | * |
| ${}^*H_1 - {}^*H_2$ at 70% | – | – | ** | ** | – |
| ${}^*H_1 - A_1$ at 30% | ** | ** | ** | – | ** |
| ${}^*H_1 - A_1$ at 50% | ** | ** | ** | – | – |
| ${}^*H_1 - A_1$ at 70% | – | ** | * | * | – |
| ${}^*H_1 - {}^*A_3$ at 30% | NA | NA | NA | NA | NA |
| ${}^*H_1 - {}^*A_3$ at 50% | – | – | * | – | NA |
| ${}^*H_1 - {}^*A_3$ at 70% | – | – | – | – | NA |
| HNR 60–5000 Hz | – | – | – | – | – |
| HNR 60–1000 Hz | *√ | – | – | – | – |
| HNR 500–1000 Hz | – | – | – | – | – |
| HNR 1–2 kHz | – | – | – | – | – |
| HNR 2–3 kHz | – | – | – | – | – |
| HNR 3–4 kHz | – | – | – | – | – |
| HNR 4–5 kHz | – | – | – | – | – |
| Vowel RMS amplitude | ** | – | ** | * | ** |
| Vowel average $F_0$ | ** | ** | – | – | – |
| Vowel duration | – | – | – | – | – |

*Significant at $p < 0.05$.

**Significant at $p < 0.01$.

√ Significant, but in unexpected direction.

five speakers also differentiate between breathy and clear vowels based on $^*H_1 - A_1$ measured at vowel center, and three speakers do so based on $^*H_1 - A_1$ measured at the end of the vowel.

$^*H_1 -^* A_3$, a measure of spectral tilt, does not separate breathy vowels from clear vowels for most speakers. The only exception is the $^*H_1 -^* A_3$ measured at 50% in the vowel. It distinguishes between the two phonations for speaker W3 and marginally so for W2 ($p > 0.053$). As for HNRs, with the exception of W1's HNR measured between 60 and 1000 Hz, HNRs do not differentiate between breathy and clear vowels in Chanthaburi Khmer. Moreover, the significant difference was in the unexpected direction.

Vowel RMS amplitude is another important cue that differentiates between breathy and clear vowels in Chanthaburi Khmer. Four out of five speakers produced breathy vowels with significantly greater amplitude than clear vowels. While the difference does not reach significance for W2, there is a strong trend in that direction ($p > 0.057$). On the other hand, only two speakers (W1, W2) produced breathy vowels with higher average $F_0$ than clear vowels. Moreover, all speakers produced breathy vowels with a relatively longer duration than clear vowels. The difference, however, was not statistically significant.

## 6. Conclusions

The main goal of the acoustic analyses reported in this study was to examine whether or not the Chanthaburi dialect of Khmer spoken in Thailand has preserved the putative historical distinction between breathy and clear phonation in its vowel system. The present results showed that $^*H_1 -^* H_2,^* H_1 - A_1,^* H_1 -^* A_3$, and vowel RMS amplitude were significant acoustic correlates of the distinction between breathy and clear vowels. $^*H_1 -^* H_2$ measured at the beginning of the vowel was the most robust cue. The use of these acoustic cues varied from speaker to speaker, suggesting individual differences in the phonetic implementation of breathy and clear phonations. This finding is consistent with the observation made by Ní Chasaide & Gobl (1997) that breathy voice "will occur to differing degree across languages or even for different speakers of one language/dialect" (p. 454). For example, breathy vowels produced by W1 showed a relatively greater open quotient ($^*H_1 -^* H_2$) at the beginning of the vowel and wider $F_1$ bandwidth ($^*H_1 - A_1$) at the beginning and middle of the vowel. Speakers W2 and W3, on the other hand, produced breathy vowels with a relatively greater open quotient, both at the beginning and in the middle of the vowel, and a wider $F_1$ bandwidth throughout the vowel.

The present results also suggest that male speakers make a distinction in terms of tense versus lax voice rather than breathy versus clear voice. This conclusion is based on the $^*H_1 -^* H_2$ measurement. As noted, both male speakers produced clear vowels with relatively enhanced $H_2$ amplitude. As suggested by Ní Chasaide & Gobl (1997), $H_2$ prominence may be a characteristic of a tense or creaky phonation. Based on the first author's observation, it seems that both male speakers produced clear vowel with a tense rather than a creaky phonation.

It was also found that among female speakers there was a difference in the location of the robust acoustic cues ($^*H_1 - ^*H_2$ and $^*H_1 - A_1$) in the vowel. For the oldest female speaker (W1), these two acoustic cues were found at the beginning ($^*H_1 - ^*H_2$) and extended up to the middle of the vowel ($^*H_1 - A_1$) only, but never at the end of the vowel. For the younger female speakers (W2, W3), on the other hand, both cues were found throughout the vowel, including at the end of the vowel. It is possible that this difference merely reflects individual differences in the phonetic implementation of breathy and clear vowels. However, the fact that breathiness did not appear to extend beyond the center of the vowel in the older female speaker might also suggest that the breathy quality of the vowel was the result of the coarticulatory effect of neighboring segments (i.e., the initial consonant). According to this hypothesis, the earlier devoicing process of initial stop consonants may not have been completed in the older female speaker's speech, and it is the difference in the phonetic characteristics of the initial consonant that results in the observed clear and breathy distinction in the vowel (see Wayland & Jongman, 2002) for intermediate stages of the devoicing process in Khmer and its influence on the following vowel). On the other hand, the devoicing process was complete among the younger female speakers and the breathy and clear voice distinction had been transferred to the following vowels. This hypothesis, however, cannot be confirmed without further acoustic analysis of the initial consonants. Nonetheless, the present results suggest that the earlier breathy and clear phonation distinction in Khmer is preserved among female speakers of Chanthaburi Khmer. In addition, this distinction may be disappearing or have become a tense versus lax distinction among male speakers.

## Acknowledgements

This research was conducted as part of a doctoral dissertation at Cornell University by the first author under the direction of the second author. Portions of this study were reported at the 133rd meeting of the Acoustical Society of America (1997, State College, PA). We thank Marie Huffman, Peter Ladefoged and James Hillenbrand for their many helpful comments on an earlier version of this paper. We also thank Scott Gargash and Eric Evans for their technical assistance.

## Appendix

The following 23 minimal or near-minimal breathy-clear vowel pairs were used in this study.

| Breathy | | Clear | |
|---|---|---|---|
| 1. /kmpɛːk/ | 'to be bald' | /pɛːt/ | 'eight' |
| 2. /plo̤w/ | 'buttock' | /plow/ | 'way, path' |
| 3. /pɔ̤ːk/ | 'bumped' | /pɔːk/ | 'by chance' |
| 4. /pɔ̤ːŋ/ | 'blistered' | /pɔːŋ/ | 'balloon' |
| 5. /pə̤ːp/ | 'to meet' | /pəːp/ | 'cry of a barking deer' |
| 6. /pra̤h/ | 'buddha image' | /prah/ | 'to lie down' |
| 7. /preːŋ | 'ancient' | /preːng/ | 'oil' |
| 8. /pri̤ɛp/ | 'dove' | /priɛp/ | 'to compare' |
| 9. /pro̤h/ | 'to sow (seed)' | /proh/ | 'male, man' |
| 10. /cnri̤ɛŋ/ | 'eave' | /cmriɛŋ/ | 'song' |
| 11. /ti̤c/ | 'to sting' | /ntic/ | 'a little' |
| 12. /ti̤ɛn/ | 'candle' | /tiɛn/ | 'to blame' |

| | | | |
|---|---|---|---|
| 13. /to̤ːc/ | ‘gibbon’ | /toːc/ | ‘small’ |
| 14. /tṳət/ | ‘great great grandparents’ | /cuət/ | ‘to wrap around’ |
| 15. /tṳm/ | ‘to perch’ | /tom/ | ‘aunt, uncle’ |
| 16. /ka̤t/ | ‘he’ | /kat/ | ‘to cut’ |
| 17. /ko̤t/ | ‘complete’ | /kot/ | ‘a crown’ |
| 18. /cɛ̤ːk/ | ‘to part (hair) | /cɛːk/ | ‘to divide up’ |
| 19. /cmo̤h/ | ‘mangoose’ | /cmoh/ | ‘name’ |
| 20. /cɔ̤ʔ/ | ‘to smoke’ | /cuʔ/ | ‘to plug’ |
| 21. /ma̤t/ | ‘eye’ | /mat/ | ‘mouthful’ |
| 22. /ŋkɔ̤t/ | ‘a crown’ | /kot/ | ‘to note down’ |
| 23. /nca̤t/ | ‘to scoop up’ | /cat/ | ‘arrange’ |

## References

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.

Bickley, C. (1982). Acoustic analysis and perception of breathy vowels. In: *Speech communication group working papers* (pp. 71–82). Cambridge: Massachusetts Institute of Technology.

Chantharupanth, T., & Phromjagarin, C. (1978). *Surin Khmer–English dictionary*. Bangkok: Institute of Indigeneous Language Research, Chulalongkorn University Press.

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

Diffloth, G. (1974). Austroasiatic language. *Encyclopedia Britannica*. (15th ed.) (pp. 480–484). Encyclopedia Britannica.

Diffloth, G. (1994). Khmer. Presentation given at the Cornell Linguistics Circle, May, 1994, Cornell University.

Fairbanks, G. (1940). *Voice and articulation drill book*. New York: Harper and Brothers.

Fant, G., Ishizaka, K., Lindqvist, J., & Sundberg, J. (1972). Subglottal formants. *Speech Transmission Laboratory-Quarterly Progress Status Report*, Vol. 1, Royal Institute of Technology, Stockholm, pp. 85–107.

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

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

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

Headley Jr., R. K., Chhor, K., Lim, L. K., Lim, H. K., & Chun, C. (1977). *Cambodian-English dictionary*. Washington D.C: The Catholic University of America Press.

Henderson, E. J. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies*, *14*(1), 149–174.

Hillenbrand, J. (1988). Perception of aperiodicities in synthetically generated voices. *Journal of the Acoustical Society of America*, *83*, 2361–2371.

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

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal Quality: dysphonic voices and continuous speech. *Journal of Speech and Hearing Research*, *39*, 311–321.

Hillenbrand, J., Metz, D. E., Colton, R. A., & Whitehead, R. L. (1990). *A high speed film and acoustic study of breathy voice*. Paper presented at the meeting of the American speech-language-hearing association, Seattle, WA.

Hombert, J. M., Ohala, J., & Ewan, W. (1979). Phonetic explanations for the development of tones. *Language*, *55*(1), 37–58.

Horii, Y. (1980). Vocal shimmer in sustained phonation. *Journal of Speech and Hearing Research*, *23*, 202–209.

Huffman, F. (1967). *An outline of Cambodian grammar*. Unpublished dissertation, Cornell University.

Huffman, F. (1976a). The relevance of lexicostatistics to Mon-Khmer languages. In P. N. Jenner, et al. (Ed.), *Austroasiatic studies I* (pp. 539–574). Honolulu: University of Hawaii Press.

Huffman, F. (1976b). The register problem in fifteen Mon-Khmer languages. In P. N. Jenner, et al. (Ed.), *Austroasiatic studies I* (pp. 575–590). Honolulu: University of Hawaii Press.

Huffman, F. (1977). *Cambodian reader and glossary (Yale linguistic series)*. New Haven and London: Yale University Press.

Huffman, F. (1978). *Synchronic evidence for the history of Khmer vowels*. Paper presented at second international conference on austroasiatic linguistics, Mysore, India.

Huffman, M. K. (1987). Measures of phonation types in Hmong. *Journal of the Acoustical Society of America*, *81*(1), 495–504.

Huffman, F., & Proum, I. (1977). *English–Khmer dictionary*. New Haven, CT: Yale University Press.

Jacob, J. M. (1960). The structure of the world in old Khmer. *Bulletin of the School of Oriental and African Studies*, *23*, 351–368.

Jacob, J. M. (1963). Prefixation and infixation in old Mon, Old Khmer, and Modern Khmer. In H. L. Shorto (Ed.), *Linguistic comparison in Southeast Asia and the Pacific* (pp. 62–70). London: Luzac and Co. Ltd.

Jacob, J. M. (1965). Notes on the numerals and numeral coefficients in Old, Middle, and Modern Khmer. In Milner, & Henderson (Eds.), *Indo-Pacific linguistic studies II* (pp. 143–162). Amsterdam: North-Holland Publishing Co.

Jacob, J. M. (1976a). An examination of the vowels and final consonants in correspondences between pre-Angkor and Modern Khmer. In N. D. Liem (Ed.), *South-East Asian linguistic studies 2* (pp. 19–38). Canberra: Australian National University.

Jacob, J. M. (1976b). Affixation in Middle Khmer, with Old and Modern comparisons. In P. Jenner, et al. (Ed.), *Austroasiatic studies I* (pp. 591–624). Honolulu: University of Hawaii Press.

Jacob, J. M. (1976c). Some problems arising from the orthography of consonants on the Khmer inscriptions. In P. Jenner, et al. (Ed.), *Austroasiatic studies I* (pp. 625–658). Honolulu: University of Hawaii Press.

Jacob, J. M. (1977). Sanskrit loan-words in pre-Angkor Khmer. In P. Jenner, et al. (Ed.), *Mon-Khmer studies VI* (pp. 151–168). Honolulu: University of Hawaii Press.

Jenner, P. N. (1974a). The value of 'au' and 'ai' in Middle Khmer. In N. D. Liem (Ed.), *Southeast Asian linguistic studies I* (pp. 157–173). Canberra: Australian National University.

Jenner, P. N. (1974b). The development of the registers in Standard Khmer. In N. D. Liem (Ed.), *Southeast Asian linguistic studies I* (pp. 47–60). Canberra: Australian National University.

Jenner, P. N. (1974c). Observations of the Surin dialect of Khmer. In N. D. Liem (Ed.), *Southeast Asian linguistic studies I* (pp. 61–73). Canberra: Australian National University.

Jenner, P. N. (1975). The final liquids of Middle Khmer. *Zeitschrift fur Phonetik und allgemeine Sprachwissenschaft*, *28*(3–4), 599–609.

Jenner, P. N. (1976a). The relative dating of some Khmer Cpa'pa'. In P. Jenner, et al. (Ed.), *Austroasiatic studies II* (pp. 693–710). Honolulu: University of Hawaii Press.

Jenner, P.N. (1976b). The value of i, i, u, and u in middle Khmer. In K. Gregerson & D. Thomas (Eds.), *Mon-Khmer Studies V* (pp. 101–134). Summer Institute of Linguistics.

Kirk, P. L., 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*. Missoula, MT: University of Montana Press.

Klatt, D. H. (1986). Detailed spectral analysis of female voice. *Journal of the Acoustical Society of America*, *81*(Suppl 1), S80.

Klatt, D. H., & Klatt, C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*(2), 820–857.

Klich, R. J. (1982). Relationships of vowel characteristics to listener ratings of breathiness. *Journal of Speech and Hearing Research*, *25*, 574–580.

Kreiman, J., & Gerratt, B. (1998). Validity of rating scale measures of voice quality. *Journal of the Acoustical Society of America*, *104*(3), 1598–1608.

Ladefoged, P. (1981). The relative nature of voice quality. *Journal of the Acoustical Society of America*, *69*(Suppl. 1), S67.

Ladefoged, P. (1982). *A course in phonetics* (2nd ed.). New York: Hartcourt Brace Jovanovich.

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

Ladefoged, P., & Antañanzas-Barroso, N. (1985). Computer measures of breathy voice quality. *UCLA Working Papers in Phonetics*, *61*, 79–86.

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, *35*(1), 344–353.

Martin, M. A. (1975). Le dialect cambodgien parlé à Tatêy, massif des Cardamomes. *Asie du Sud-est et Monde Insulindien*, *6*(4), 71–79.

Maspéro, G. (1915). *Grammaire de la langue khmère (Cambodgien)*. Paris: Imprimerie Nationale VIII.

Ní Chasaide, A., & Gobl, C. (1997). Voice source variation. In W. Hardcastle, & J. Laver (Eds.), *The handbook of phonetics sciences*. Oxford: Blackwell Publishers Ltd.

Phon-ngam, P. (1987). *A phonological comparison of spoken Central Khmer* (*Phnom Penh*) *and Northern Khmer* (*Surin*). MA Thesis, Mahidol University, Thailand.

Pinnow, H. (1957). Sprachgeschichtliche Erwägungen zum Phonemsystem des Khmer. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*, *10*(4), 318–391.

Premsrirat, S. (1995). Phonetic variation of final trill and final palatals in Khmer dialects of Thailand. *Mon-Khmer Studies*, *24*, 103–130.

Samely, U. (1991). *Kedang (Eastern Indonesia), some aspects of its grammar*. Hamburg: Helmut Buske Verlag.

Silverman, D., Blankenship, B., Kirk, P., & Ladefoged, P. (1995). Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics*, *37*, 70–88.

Smalley, W. (1964). *Ethnolinguistic Survey of Northern Khmer speaking people in Northeast Thailand*, (*with data on Kuy*). Unpublished manuscript.

Smalley, W. (1976). The problem of vowels: Northern Khmer. In W. Smalley (Ed.), *Phoneme and orthography: language planning in ten minority languages of Thailand* (pp. 43–84). Canberra: Australia National University.

Stevens, K. (2000). *Acoustic phonetics*. Massachusetts: MIT Press.

Stevens, K., & Hanson, H. (1994). Classification of glottal vibration from acoustic measurements. *Paper presented at the 8th vocal fold physiology conference*, Kurume, Japan, April 7–9, 1994.

Thomas, D., & Headley, R. (1970). More on Mon-Khmer sub-grouping. *Lingua*, *25*(4), 398–418.

Thomas, D., & Tienmee, V. (1990). An acoustic study of Northern Khmer vowels. *Mon-Khmer Studies*, *17*, 99–108.

Thongkum, T. (1988). Phonation types in Mon-Khmer languages. In O. Fujimura (Ed.), *Vocal fold physiology: voice production, mechanisms and functions* (pp. 319–334). New York: Raven Press.

Thongkum, T.L. (1991). An instrumental study of Chong register. In J.H.C.S. (Ed.), *Austroasiatic languages*: *essays in honor of H.L. Shorto* (pp. 141–160). London: School of Oriental and African Studies, University of London.

Wayland, R. (1998). An acoustic study of Battambang Khmer vowels. *Mon-Khmer Studies*, *28*, 43–62.

Wayland, R., Gargash, S., & Jongman, A. (1994). Acoustic and perceptual investigation of breathy voice. *Journal of the Acoustical Society of America*, *97*, 3364.

Wayland, R., & Jongman, A. (2001). Chanthaburi Khmer Vowels: phonetic and phonemic analyses. *Mon-Khmer Studies*, *31*, 65–82.

Wayland, R., & Jongman, A. (2002). Registrogenesis in Khmer. *Mon-Khmer Studies*, *32*, 101–114.

Zemlin, W. R. (1968). *Speech and hearing science: anatomy and physiology*. New Jersey: Prentice-Hall.