Journal of Phonetics 40 (2012) 152–161

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

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

The timing and sequencing of coarticulated non-modal phonation in English and White Hmong

Marc Garellek*

University of California, Los Angeles, CA 90095-1543, USA

ARTICLE INFO

Article history:
Received 7 September 2010
Received in revised form
26 September 2011
Accepted 2 October 2011
Available online 12 November 2011

ABSTRACT

Despite the growing number of studies on the acoustics of non-modal phonation, little is known about how two distinct non-modal phonations can interact acoustically when coarticulated. This study investigates the acoustics of breathy-to-creaky phonation contours in vowels from a production study of native speakers of English and White Hmong. These languages differ in the nature of the non-modal phonation types. In the English corpus, both the breathiness and creakiness are non-contrastive. In the Hmong corpus, the breathiness can be contrastive or a result of coarticulation with a neighboring segment, but the creakiness is always contrastive.

The contours were analyzed using the three measures of phonation that were found to best differentiate non-modal from modal phonation in these languages: H1*–H2*, H1*–A1*, and harmonics-to-noise ratio. Results from these measures provide support for the presence of breathy–creaky contours in vowels. The timing and sequencing of the breathy and creaky phonation types are largely dependent on whether they are contrastive, with contrastive non-modal phonation being present during more portions of the vowel than non-modal phonation derived from coarticulation. The acoustic results also provide evidence for simultaneous breathy and creaky phonation types in Hmong.

1. Introduction

This study investigates the acoustics of breathy phonation that is coarticulated with creaky phonation in English and (White) Hmong. There is little known about how two distinct non-modal phonations can interact when coarticulated. Studies of non-modal phonation typically focus on experimental (usually acoustic) descriptions of the production of contrastive non-modal phonation, as in Green Mong (Andruski & Ratliff, 2000), White Hmong (Esposito, Ptacek, & Yang, 2009; Huffman, 1987), Khmer (Wayland & Jongman, 2003), Chong (DiCanio, 2009) Zapotec (Avelino, 2010; Esposito, 2010a), and Mazatec (Garellek & Keating, 2011). Additionally, there has been some research on the interaction of phonation and prosody, especially in English (e.g. Epstein, 2002; Huffman, 2005), on the phasing of non-modal phonation (Silverman, 1995), and its duration (Blankenship, 2002). Recently there has also been research on the perception of contrastive phonation types (Abramson, L-Thongkum, & Nye, 2004; Esposito, 2010b; Gerfen & Baker, 2005).

Other research has been directed at non-modal phonation that is non-contrastive, as in English (Ladefoged, 1983; Löfqvist & McGowan, 1992), Swedish (Gobl & Ní Chasaide, 1999; Löfqvist & McGowan, 1992), Trique (DiCanio, 2011) and Tagalog (Blankenship, 1997), to name a few. The studies of non-modal phonation have either implicitly or explicitly dealt with coarticulation of phonation, because non-contrastive non-modal phonation is usually due to coarticulation from adjacent segments, mostly glottalized or aspirated ones. An example is the breathiness of English vowels following aspirated stops (e.g., Epstein, 1999; Löfqvist & McGowan, 1992). However, there is currently little understanding of how one type of non-modal phonation may be coarticulated with another. The goal of this study is therefore to investigate the acoustic consequences of non-modal phonation coarticulation.

One point of interest in studying phonation coarticulation is to investigate the timing and sequencing of the adjacent voice qualities. Blankenship (2002) found that contrastive non-modal phonation lasts longer and is more differentiated from modal voice than non-modal phonation derived from coarticulation. This finding is interesting, because studies of other types of coarticulation have shown that coarticulation can span whole segments and even across segments. For example, Cohn (1990) found that nasal flow in English begins early for a vowel preceding a nasal. Thus, the non-contrastive vowel nasalization of English is not substantially shorter than the vowel nasalization of contrastive nasal vowels in French, though the overall amount of nasalization in English may be less than in French. Additionally, Öhman's study

* Tel.: +1 310 825 0634; fax: +1 310 206 5743.
E-mail address: marcgarellek@ucla.edu

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.10.003

on coarticulation (1966) revealed that in VCV sequences, the first vowel already shows effects of the following vowel in English and Swedish. West (1999) also found coarticulatory effects of English /l/ and /r/ on vowels immediately and more distantly preceding the liquid. The coarticulation of some Mandarin tones, which like phonation involve laryngeal articulation, can also influence large portions of vowels (Xu, 1997).

Blankenship's findings imply that an interval of non-contrastive non-modal phonation should be shorter than an interval of contrastive non-modal phonation even when the two are adjacent. However, if two types of non-contrastive non-modal phonation are adjacent in a given segment, then they should have roughly the same duration. In her study, Blankenship did not study non-contrastive phonation in languages with a phonation contrast. Presumably, the non-modal phonation derived from coarticulation in these languages should be shorter in duration than the contrastive phonation, but not necessarily shorter than the corresponding non-modal phonation in a language without phonation contrasts. On the other hand, studies by Manuel and Krakow (1984) and Manuel (1987, 1990) have shown that the number of contrasts can limit the extent of coarticulation in a given language. Looking at V-to-V coarticulation, they found that coarticulation was greater in languages with fewer vowel contrasts. They interpret this finding in terms of the presence of output constraints on segments. These constraints are determined in part by the number of contrastive phonemes of the language (in their case, the number of contrastive vowels). Coarticulation is faithful to these constraints, and so is found to be less extensive in languages with many vowels occupying a common space. From the point of view of phonation coarticulation, these studies suggest that having a phonation contrast in a given language should result in less coarticulation of non-modal phonation. Thus, non-modal phonation derived from coarticulation should be less extensive in languages with a phonation contrast, but contrastive phonation should always be extensive. This hypothesis differs from Blankenship's, in that it assumes that the extent of coarticulated non-modal phonation will depend on the presence of a phonation contrast in the language, rather than on intrinsic timing differences between contrastive and non-contrastive phonation derived from coarticulation.

Thus, the goals of this study are to describe the acoustics of coarticulated non-modal phonation in English and Hmong, and to determine whether Blankenship's or Manuel and Krakow's and Manuel's predictions hold true under environments of non-modal coarticulation. As I will show, both languages investigated have the potential of showing breathy–creaky contours in vowels, where by 'breathy' and 'creaky' I make no specific claims about the nature of the articulatory gestures involved, which may also differ depending on whether the non-modal phonation is contrastive. For the purposes of the present study, by 'breathy' and 'creaky' I mean that the acoustic measures that are known to differentiate breathy and creaky voice cross-linguistically show statistically significant differences from modal voice.

In this study, the breathy–creaky contours differ in whether the breathiness and creakiness are contrastive or from coarticulation with an aspirated or glottalized segment. Specifically, audio recordings were made of words with expected breathy–creaky contours in English (for which both the breathiness and creakiness are non-contrastive), and in Hmong (for which the creakiness is contrastive but the breathiness may or may not be contrastive). Keeping with Blankenship (2002), I would predict that the phonemic breathiness of Hmong should be more extensive and more differentiated from modal than the non-contrastive breathiness in Hmong or English. Conversely, the phonemic creakiness of Hmong should be more extensive and more differentiated from the modal than the non-contrastive creakiness of English. If indeed there are output constraints on phonation coarticulation similar to those proposed for vowels by Manuel and Krakow (1984) and Manuel (1987, 1990), then the non-contrastive breathiness of Hmong should be more restricted in its timing than that of English, given that Hmong contrasts three types of phonation while English has no phonation contrast.

In Section 2, I will review previous studies of non-modal phonation in English and Hmong, and the acoustic correlates used to characterize phonation cross-linguistically. In Section 3, I describe the experiment and show time courses of the measures used in this study. Section 4 and the conclusion are devoted to a discussion of the results presented in the previous section, interpreting the results in light of the studies mentioned above.

2. Phonation types in English and Hmong and their acoustic correlates

This section reviews non-modal phonation found in English and Hmong, and outlines which acoustic measures have been used to characterize non-modal phonation in these languages and others.

2.1. Phonation in English and Hmong

2.1.1. English

English does not have a phonation contrast, but vowels can show non-modal phonation in certain prosodic conditions (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Epstein, 2002; Huffman, 2005; Kreiman, 1982; Pierrehumbert & Talkin, 1992; Redi & Shattuck-Hufnagel, 2001, among others) and adjacent to certain sounds. Vowels following /h/ and aspirated stops are slightly breathy (Epstein, 1999; Gobl & Ní Chasaide, 1999; Ladefoged, 1983; Löfqvist & McGowan, 1992). By 'breathiness' in English I am not referring to the (voiceless) aspiration noise that can follow voiceless stops, which in fact is a primary cue to the phonemic contrast between voiced and voiceless stops in English (Lisker & Abramson, 1964). Rather, I am referring to the breathy voicing in the vowel following such aspiration noise. Although aspiration noise is known to be a salient cue in the perception of stop contrasts in English, there is little evidence to show that listeners rely on the following breathy voicing as a cue to the preceding stop category. Because the breathy voicing is due to coarticulation following the aspiration noise, it is viewed as non-contrastive in the present study, just as the breathiness in Tagalog vowels following /h/ were treated by Blankenship (2002).

In addition, often appear as allophones of voiceless stops in English, especially after sonorants (nasals, liquids, glides or vowels), or before other obstruents and sonorants in coda position (Cohn, 1993; Epstein, 2002; Huffman, 2005; Selkirk, 1972; Westbury & Niimi, 1979). Vowels before these glottalized stops often have some creaky phonation, derived from coarticulation with the coda (Garellek, 2011a).

Breathiness and creakiness can theoretically co-occur in an English vowel if the environments for both are combined. For example, the vowels in words like *cat* or *hat* will show breathiness at the vowel's onset, because they follow an aspirated stop or /h/. Moreover, the same vowels should show creakiness at the vowel offset, because they are followed by a voiceless stop in coda position. Therefore, my hypothesis is that English words like *cat* or *hat* should show a phonation contour in the vowels, starting from more breathy-like phonation and ending in more creaky-like phonation. Again, by 'breathy' and 'creaky,' I mean that the phonation types are acoustically closer to breathy and creaky voice than to modal. It should matter little, for English as well as for Hmong, whether in reality these contours are breathy-tense, or lax-creaky, for example, although a more extreme phonation (i.e., a phonation that is more differentiated from modal voice) might in fact be inherently longer in duration. The possibility that cross-linguistic timing differences might be a result of differences in the type of non-modal phonation will be assessed in the discussion section below.

2.1.2. Hmong

Whereas English only shows non-modal phonation due to coarticulation, Hmong contrasts both breathy and creaky vowels with modal ones. In Hmong, non-modal phonation is associated with certain tones. Vowels may have a phonemic low creaky tone (e.g. /pà̰/ 'blanket') or a phonemic falling breathy tone, for example /pâ̤/ 'pile'. Both the creaky and breathy tones contrast with tones with similar pitch but different phonation types. Thus, /pà̰/ 'blanket' and /pâ̤/ 'pile' contrast mostly in phonation with modal /pà/ 'stick' and /pâ/ 'flower', respectively, although differences in the pitch contours exist as well, especially for the creaky tone (Esposito et al., 2009; Ratliff, 1992).

Hmong distinguishes unaspirated from aspirated stops at all its places of articulation. The aspirated stops cannot occur before breathy vowels, suggesting that the aspiration noise is perceptually confusable with vowel breathiness. Indeed, both phonemically modal vowels following aspirated stops and phonemically breathy vowels show increased breathiness in comparison to modal vowels (Esposito & Khan, 2010; Fulop & Golston, 2008).

However, creaky vowels may follow aspirated stops in Hmong, for example in /pʰà̰/ 'chubby, fat.' Such words should show a breathy–creaky phonation contour. Unlike in English, the creakiness in such contours is phonemic.

2.2. Phonation measures

Although phonation strictly speaking involves only differences in the mode of vibration of the vocal folds, non-modal phonation is known to sometimes involve other laryngeal and even supralaryngeal postures (Edmondson & Esling, 2006). Moreover, a range of non-modal phonation is possible, even at the level of the vocal folds. For example, breathy phonation can be produced by a wide resting aperture of the vocal folds, by the less abrupt closing of the folds, or by maintaining a constant posterior (inter-arytenoid) gap between the folds (Gordon & Ladefoged, 2001). In their laryngeal articulator model, Edmondson and Esling attribute laryngealized voice and glottalization not only to the vocal folds but also the ventricular folds, arytenoids, and aryepiglottic folds. Additionally, phrase-final creak has been found to involve low sub-glottal pressure (Slifka, 2006), and certain types of creaky voice may include additional effects like vocal fry and period doubling (Gerratt & Kreiman, 2001). Owing to this multi-dimensional nature of non-modal phonation, various acoustic measures have been used to distinguish modal phonation from its non-modal counterparts. By far the most commonly used measure is H1–H2, or the difference in the amplitudes of the first and second harmonics. A higher value of H1–H2 is thought and often found to be correlated with greater glottal open quotient (DiCanio, 2009; Holmberg, Hillman, Perkell, Guiod, & Goldman, 1995; Stevens & Hanson, 1995; Sundberg, Andersson, & Hultqvist, 1999; but cf. Kreiman et al., 2008). Open quotient (OQ) is the proportion of a glottal period during which there is no contact between the vocal folds. H1–H2 as a correlate of OQ should be a good measure for differentiating non-modal phonations from modal, since breathy phonation often has a greater OQ than modal, whereas creaky phonation can involve a more closed glottis. Indeed, for languages with contrastive breathy phonation, H1–H2 (or its formant-corrected counterpart, denoted by asterisks: H1*–H2*) has been shown to be greater in breathy phonation than in modal for a variety of languages (Bickley (1982), for Gujarati; Huffman (1987), for Hmong; Blankenship (1997), for Mazatec; Wayland and Jongman (2003), for Khmer; Miller (2007), for Ju|'hoansi; for other languages, see Esposito (2010b)). For languages that contrast creaky phonation with modal, a lower H1–H2 has also been found for Mazatec (Blankenship, 1997; Garellek & Keating, 2011), Green Mong (Andruski & Ratliff, 2000), Ju|'hoansi (Miller, 2007), Chong (DiCanio, 2009), and Santa Ana del Valle Zapotec (Esposito, 2010a).

In addition to H1–H2, wideband spectral tilt measures comparing H1 to the amplitude of the first formant (A1) or the second or third formants (A2 or A3) have been used. These measures are thought to correlate with the abruptness of vocal fold closure (Stevens, 1977). H1–A1 is also thought to be correlated with the bandwidth of the first formant, which might reflect posterior glottal opening at the arytenoids (Hanson, Stevens, Kuo, Chen, & Slifka, 2001). Taking these studies into account, a higher value of H1–A1 should be an indication of whispery voice, which is produced by means of air flowing through the arytenoids (Laver, 1980), whereas the higher spectral tilt measures like H1 − A2 and H1 − A3 should correlate with speed or abruptness of closure. Blankenship (1997) found that some of these measures could distinguish modal from laryngealized phonation to some degree. As with H1–H2, these measures are more often used for comparing breathy and modal than creaky vs. modal phonation types. For more on the effectiveness of these measures at distinguishing breathy vs. modal phonation in a number of languages, see Esposito (2006). Esposito (2010b) showed that H1–A1 was able to distinguish breathy from modal phonation in several languages. If H1–A1 is a correlate of whispery voice, this suggests that some degree of whispery phonation is present in the breathy phonation of some languages, as is claimed in some studies, e.g. by Fulop and Golston (2008). This is not surprising, given that breathy voice often involves incomplete closure of the vocal folds, which could facilitate inter-arytenoid opening as well.

Noise measures have also been used to distinguish breathy or creaky phonation from modal. De Krom (1993) found that the harmonics-to-noise ratio (HNR) decreased almost linearly as the noise in the signal increased. Noise can be due to aspiration, which is a characteristic of breathy voice, or to aperiodicity, which is a characteristic of creaky voice (Gordon & Ladefoged, 2001). Thus, both breathiness and aperiodicity result in lower HNR values. HNR has been used to distinguish breathy from modal phonation in Javanese (Wayland, Gargash, & Jongman, 1994), Ju|'hoansi (Miller, 2007), and White Hmong (Fulop & Golston, 2008). Miller (2007) also found an effect of HNR for glottalized vowels. Blankenship (2002) showed that a similar measure of noise, cepstral peak prominence (Hillenbrand, Cleveland, & Erickson, 1994), distinguished breathy from modal phonation in Mazatec and Chong, as did Esposito (2010b) for a number of languages.

In sum, breathy–creaky contours are likely to be found in English and Hmong The contours may well be manifested by different acoustic measures, given the success of various studies at characterizing phonation differences using a variety of measures.

3. Experiment

The experiment was designed to compare vowels with breathy–creaky contours to modal vowels in English and Hmong. In addition, breathy–modal and modal–creaky contours were included for comparison in both languages, and contrastive breathy vowels were included for Hmong.

3.1. Method

3.1.1. Stimuli

3.1.1.1. English. The stimuli are divided into four groups based on expected phonation pattern. The target group consists of monosyllabic English words with an expected breathy–creaky contour. These words begin with an aspirated stop /p, t, k/ or /h/, have a low vowel /æ/ or /ɑ/, and end in coda /p, t, k/, for example *pat*.

The next group consists of words with an expected breathy–modal contour. These differ from the breathy–creaky words by having a coda /s/ (or sometimes /st/, /sk/, or /z/) instead of coda stops, for example *pass*. Although /s/ also involves glottal spreading and therefore in principle could induce some breathiness on the vowel, little breathiness was found. Fricative codas were chosen because a pilot study revealed that voiced stops in coda position still resulted in creak, and sonorants were avoided because they were likely to influence the formant tracking. The voicing difference between /s/ and /z/ in codas was deemed trivial, in that it was unlikely to alter the voice quality of the preceding vowel appreciably. Moreover, the following word from the carrier *Say the words _____ for me* begins with voiceless /f/, thus resulting in the partial devoicing of /z/ codas by assimilation. The carrier contained *words* in plural because the target was always preceded by a function word (either *a*, *to*, or *I*), for reasons to be motivated below for target words beginning with /h/. For complex codas, speakers usually elided the second consonant (either /t/ or /k/), given that the following word began with an obstruent.

The third group of words is comprised of those with an expected modal–creaky contour. These differ from the breathy–creaky words in that they begin with a voiced stop /b, d, g/ instead of an aspirated one, for example *bat*.

The last group consists of words with expected modal vowels with little phonation contour. These differ from the modal–creaky words in that they end with /s/ or /z/, for example *boss*.

To increase the likelihood that the word-initial /h/ would be partially voiced, all target words were preceded by a function word ending in a vowel so that /h/ would be intervocalic and voiced. Voiced [ɦ] might be more likely to induce breathiness on the following vowel, given that it involves breathy phonation, whereas voiceless [h] involves mostly aspiration noise. Indeed, phonemic breathy vowels can be reflexes of intervocalic /h/, as in Gujarati (Fischer-Jørgensen, 1967) and Mazatec (Silverman, Blankenship, Kirk, & Ladefoged, 1995). The reason for including voiced [ɦ] was therefore to see whether the breathiness on the following vowel differed in any way from the breathiness following aspirated stops.

3.1.1.2. Hmong. Hmong has no coda consonants except [ŋ], which was avoided for the effects of nasality on the preceding vowels. Thus, all stimuli are words of shape CV, where the vowel carries a tone, written orthographically as *-m* (low creaky), *-g* (falling breathy), and the modal tones *-s, -v, -b,* and *-j*, or null. The stimuli are divided into four groups based on expected phonation pattern. The first group consists of words with an expected breathy–creaky contour. These words begin with an aspirated stop /pʰ, tʰ, kʰ/ or /h/ and have an open or open-mid creaky vowel /à̰/ or /ɔ̰̀/, for example /pʰà̰/ 'chubby, fat.'

The second group consists of words with an expected breathy–modal contour. These differ from the previous group in that their tones were modal, either high, mid, or low level tones, e.g. /pʰa/ 'to lead the way.' The low tone was preferred because its pitch resembles most closely that of the creaky tone, but if such a word could not be found then other level modal tones were used.

The third group consists of words with unaspirated onsets /p, t, k/ and the breathy falling tone, e.g. /pâ̤/ 'pile.'

The fourth and fifth groups consist of words with unaspirated onsets /p, t, k/, but whose tones were creaky (e.g. /pà̰/ 'blanket') and modal (e.g. /pà/ 'stick').

As for English, the carrier word preceding the target stimuli in Hmong ended in a vowel, promoting the voicing of /h/ to [ɦ] in targets beginning with that sound. The Hmong carrier was [ʈɔ̌ haì _____ dua] 'Repeat _____ again.'[1]

The complete list of stimuli for the two languages can be found in the Tables A1 and A2 of Appendix A in the Supplementary Material.

3.1.2. Participants

Twelve speakers of North American English were recorded: six women and six men. Two of the male speakers were from Canada and Washington D.C.; the rest were from California. Most of the speakers described themselves as fluent in at least one other language. The English speakers were recorded in a sound-attenuated booth using a Shure SM10A head-mounted microphone, whose signal ran through an XAudioBox pre-amplifier and A-D device. The recording was done using PCQuirerX at a sampling rate of 22,000 Hz. Thirteen speakers of Hmong were recorded: seven women and six men. One of the women was not included in the analysis because she was a native speaker of Green Mong. Three of the Hmong speakers spoke both White (*Daw*) and Green (*Leng*) natively. The remaining speakers spoke only the White variety. All the Hmong speakers could speak some English as well as Lao or Thai. All were living at the time in California. The speakers were recorded in a sound-attenuated room using a CAD u37 USB microphone and a laptop computer and using Audacity at a sampling rate of 22,000 Hz.[2]

3.1.3. Test sentences and procedure

Speakers were asked to say the target words in a carrier phrase. They were instructed to repeat each phrase before saying the next one. No instructions were given regarding rate of speech or voice quality, and the speakers were asked to utter the sentences as naturally as possible. The English carrier was *Say the words _____ for me*. This carrier phrase was chosen because it ensured that the coda-[t] would be unreleased (and thus likely preceded by creaky voice). The Hmong carrier was [ʈɔ̌ haì _____ dua] 'Repeat _____ again'. In total, 969 English and 773 Hmong tokens were obtained.

3.1.4. Labeling and measurements

The target vowel was labeled in Praat (Boersma & Weenink, 2009). The onset and offset of the vowel were taken to be the beginning and end, respectively, of clear first and second formants, in order to obtain reliable values for voice quality measures that depend on formants. All vowels were coded for quality (either /æ/ or /ɑ/ for English and /a/ or /ɔ/ for Hmong), as well as for preceding and following consonants. In the case of Hmong, the vowel's tone was also coded.

The acoustic measures for the labeled portions were obtained using VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011), which

[1] The effect of Hmong /d/ on the phonation of the preceding word is unclear, but presumably any effect would be the same for all contours. The tones of all the target words were pronounced at the expected pitch level and phonation type, suggesting that no tone sandhi influenced the pronunciation of the target word.

[2] The fact that the Hmong speakers were recorded in a different environment and with different equipment could theoretically affect the noise measures. However, the Harmonics-to-noise measure was actually highest in amplitude for Hmong, indicating that the harmonic amplitudes were well above the noise floor for these recordings.

calculates F0 and phonation measures optionally using the correction algorithm from Iseli, Shue, and Alwan (2007). A variety of measures were initially obtained to determine which ones best distinguish the phonation types. These included F0 using the STRAIGHT algorithm (Kawahara, Masuda-Katsuse, & de Cheveigné, 1999) and spectral tilt measures corrected for formant values as denoted by use of asterisks: H1*–A1*, H1*−A2*, and H1*−A3*. The formant frequencies and their amplitudes are calculated using the Snack Sound Toolkit (Sjölander, 2004). Harmonics-to-noise ratios for four frequency ranges (<500 Hz, <1500 Hz, <2500 Hz, <3500 Hz) are calculated using the algorithm in de Krom (1993).

3.2. Results and discussion

The measures investigated in this study are H1*–H2*, H1*–A1*, and harmonics-to-noise ratio under 500 Hz (HNR). Of the measures obtained by VoiceSauce (see Section 3.1.4), these three measures were found to be the best at distinguishing breathy phonation from modal phonation, as well creaky phonation from modal phonation, for the English and Hmong corpora (Garellek, 2010). In order to compare the time courses of each measure and for each language, VoiceSauce output the averages over ninths of the labeled interval corresponding to the voiced portion of the target vowel. These time courses were then averaged across all words belonging to the expected phonation contour. The time courses are thus in normalized rather than raw duration, which smoothes the data. Another reason for choosing to normalize time was so that the beginnings and ends of the vowels could be compared across contours. Given that preceding and following consonants are known to affect the vowel's voice quality, it would be less informative to compare the measures for different contours at raw duration points if one contour would be influenced by the preceding or following segment but not another.

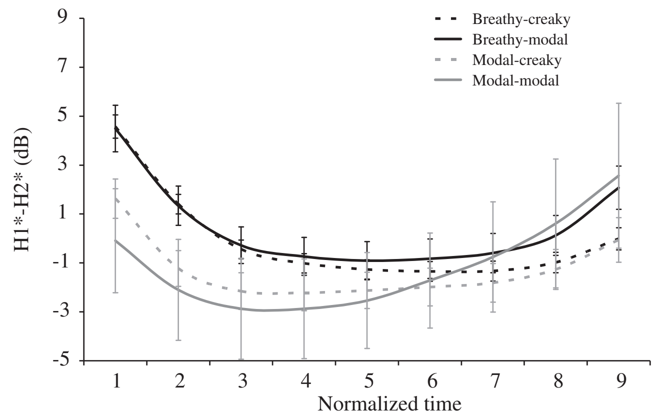
Importantly, time normalization can only allow for accurate cross-contour comparison if the contours are of roughly the same duration. Thus, to avoid any substantial averse effect of the time normalization, the time courses to be described only include a subset of the data for which the vowel durations all fell within one standard deviation from the overall mean duration. In total, 666 English tokens and 672 Hmong tokens were retained. All the vowels analyzed in this study had roughly the same durations, differing by at most 24 ms across contours, as shown in Table 1. Because the durations of the contours differed little within language, it is assumed that the timing of non-modal phonation (where the non-modal phonation is different from modal) is representative of its raw duration. It should also be noted that in general, the contours were much shorter in English than in Hmong, likely due in part to the fact that the Hmong words were open syllables, whereas the English ones had voiceless codas.

The time courses for each measure are plotted in Figs. 1–6, organized by measure. For each language plot there are four lines corresponding to four contours: breathy–creaky, breathy–modal, modal–creaky, and modal–modal. The contours beginning with breathy phonation are dark; those beginning with modal are light. The contours ending in modal phonation have solid lines; those ending in creak are in dotted lines. If a measure differentiates between breathy and modal at the vowel onset and between modal and creaky at the vowel offset, then the figure should show differentiation by line color at the beginning, but differentiation by line texture at the end. For Hmong, an additional breathy–breathy contour, derived from a contrastively breathy vowel, is included in a dashed line.

The modal–modal contour was used as the baseline for comparison with the other three contours. This contour was chosen as a baseline rather than an average of all contours or an average of modal values so that at each ninth, the comparison between modal and non-modal phonation could be made while

Table 1
Average raw duration (with standard deviations in parentheses) in ms by contour, for the subset of data included in the analysis.

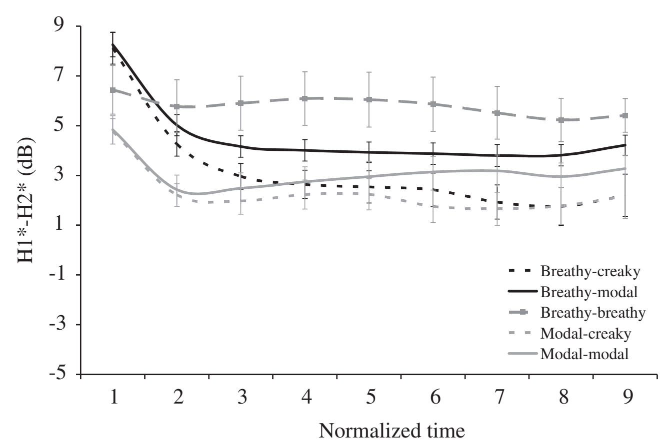
	English	Hmong
Breathy–creaky	113.06 (13.19)	188.84 (27.09)
Breathy–modal	119.83 (14.90)	193.63 (26.39)
Modal–creaky	129.56 (13.74)	201.26 (25.99)
Modal–modal	132.76 (13.98)	212.72 (22.78)
Breathy–breathy	–	206.04 (27.46)



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BC vs. MM	✓	✓	✓	✓	✓				
BM vs. MM	✓	✓	✓	✓	✓				
MC vs. MM									

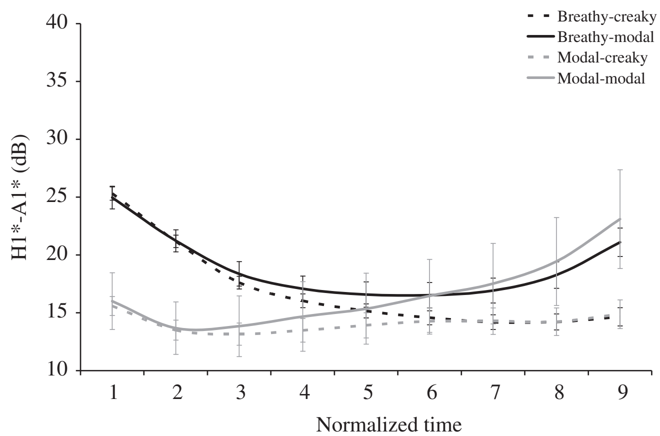
Fig. 1. Time courses of H1*–H2* for English with 95% confidence intervals. Breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. p-Values under 0.01 are checked off.



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BB vs. MM	✓	✓	✓	✓	✓	✓	✓	✓	✓
BC vs. MM	✓	✓					✓	✓	✓
BM vs. MM	✓	✓	✓	✓					
MC vs. MM						✓	✓	✓	✓

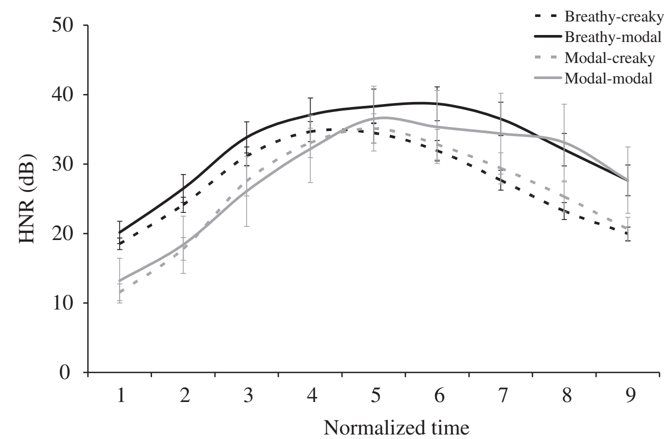
Fig. 2. Time courses of H1*–H2* for Hmong with 95% confidence intervals. Breathy–breathy contour is the dashed line; breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. p-Values under 0.01 are checked off.



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BC vs. MM	✓	✓	✓	✓					✓
BM vs. MM	✓	✓	✓	✓					
MC vs. MM									

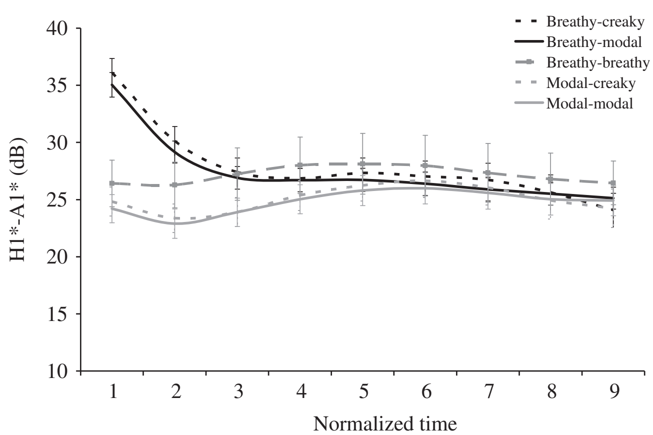
Fig. 3. Time courses of H1*–A1* for English with 95% confidence intervals. Breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. *p*-Values under 0.01 are checked off.



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BC vs. MM	✓							✓	✓
BM vs. MM	✓	✓	✓						
MC vs. MM									✓

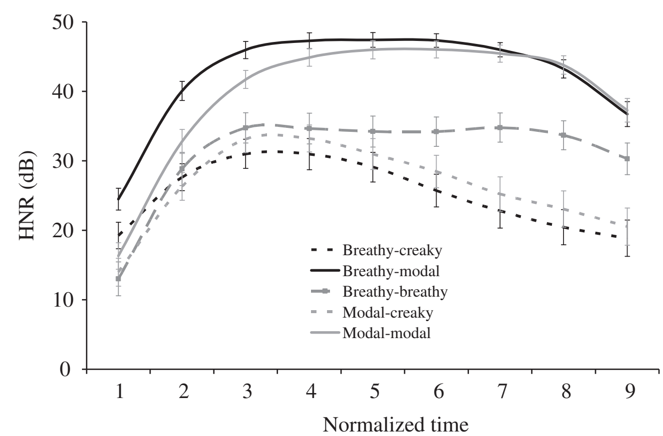
Fig. 5. Time courses of HNR for English with 95% confidence intervals. Breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. *p*-Values under 0.01 are checked off.



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BB vs. MM		✓	✓	✓					
BC vs. MM	✓	✓	✓						
BM vs. MM	✓	✓	✓						
MC vs. MM									

Fig. 4. Time courses of H1*–A1* for Hmong with 95% confidence intervals. Breathy–breathy contour is the dashed line; breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. *p*-Values under 0.01 are checked off.



Statistical differences at each vowel ninth

Vowel ninth	1	2	3	4	5	6	7	8	9
BB vs. MM			✓	✓	✓	✓	✓	✓	✓
BC vs. MM	✓	✓	✓	✓	✓	✓	✓	✓	✓
BM vs. MM	✓	✓	✓	✓					
MC vs. MM		✓	✓	✓	✓	✓	✓	✓	✓

Fig. 6. Time courses of HNR for Hmong with 95% confidence intervals. Breathy–breathy contour is the dashed line; breathy–creaky contour is the dotted dark line; breathy–modal contour is the solid dark line; modal–creaky is the dotted light line; modal–modal is the solid light line. Different colored lines should be compared at the beginning; different textures at the end. *p*-Values under 0.01 are checked off.

still factoring out the influence from the preceding and following segments. At each ninth, a linear mixed-effects model was run in R using the *lmer* function of the *lme4* package (Baayen, Davidson, & Bates, 2008), comparing each contour to the baseline, with the acoustic measure in question as the dependent variable. The largest significant model had subject and item as random effects. The inclusion of additional fixed or random effects such as F0, vowel quality, or onset type did not significantly improve model fit. Thus, although voice quality measures are known to interact with F0 and vowel quality across languages (Garellek & Keating, 2011; Kuang, 2011; DiCanio, 2011), no such interactions were found. No substantial difference was found between aspirated stops and /h/, and so vowels following both were considered to begin breathy. The expected phonation contour was used as the fixed effect of the model. The *p*-values for the coefficient estimates were obtained using the *pvalues.fnc* function in the *languageR* package in R, which approximates *p*-values by conducting Markov chain Monte Carlo (MCMC) sampling with 10,000 simulations (Baayen, 2007). Given the large number of tests performed, the *p*-values were adjusted so that those under 0.01 were considered statistically significant. The exact *p*-values can be found on the author's website, listed in the acknowledgments.

3.2.1. H1*–H2*

The time courses of H1*–H2* are plotted in Fig. 1 for English and Fig. 2 for Hmong. Four contours are plotted in each figure: breathy–creaky, breathy–modal, modal–creaky, and modal–modal. For Hmong, an additional breathy–breathy contour, derived from a contrastively breathy vowel, is included. The higher the value of H1*–H2*, the breathier the phonation. Conversely, the lower the value of H1*–H2*, the creakier the phonation. The statistical differences from the modal–modal (MM) contour for the three or four non-modal ones (abbreviated as BB, BC, BM, and MC) are shown below each figure. Statistically significant differences below 0.01 are checked off in the figures.

The results for H1*–H2* for English show that the measure has higher values for breathy phonation than for modal during the first five ninths of the vowel (depicted by both the solid and dotted black lines). H1*–H2* does not differentiate modal from creaky phonation at the vowel offset. Note that for modal-ending contours (the solid lines), the value of H1*–H2* tends to rise towards the end of the vowel. This is likely due to the coda /s/ in such words, which is [+spread glottis] (Halle & Stevens, 1971).

The results for Hmong in Fig. 2 show that H1*–H2* is higher for breathy than for modal phonation during the first two ninths of the vowel for breathy–creaky words, but for the first four ninths for breathy–modal ones. The contrastively breathy words (represented by a dashed line) have breathier vowels than modal words (represented by a light solid line) for every interval of the vowel. Unlike for English, H1*–H2* for Hmong *does* differentiate modal from creaky phonation at the vowel offset (in the final third), for both modal–creaky and breathy–creaky contours.

3.2.2. *H1*–A1**

The results for H1*–A1* for English in Fig. 3 show that the measure differentiates breathy and modal phonation during the first half of the vowel. For breathy onsets, H1*–A1* is higher, as expected. As with H1*–H2*, H1*–A1* does not consistently differentiate modal from creaky phonation at the vowel offset, although the trend is in the expected direction, with lower values for creaky offsets.

For Hmong, H1*–A1* (in Fig. 4) differentiates between breathy and modal phonation during the first third of the vowel for breathy–creaky and breathy–modal. This finding is similar to that of H1*–H2*, lending further support that words that start with an aspirated stop and have a phonemically modal vowel are breathier than those with unaspirated onsets at the beginning of the vowel. For the breathy–breathy contour, H1*–A1* is higher than the modal–modal contour in the first half of the vowel, but these values are smaller than those for the breathy–creaky and breathy–modal contours in the first third. This is surprising, given that in the latter two contours the breathiness is non-contrastive.

3.2.3. *Harmonics-to-noise ratio*

HNR is expected to be lower for breathy and creaky portions, because the measure is sensitive to both aspiration and aperiodicity. The results for HNR for English in Fig. 5 show that the measure differentiates between breathy and modal phonation during the first third for breathy–modal, though in the opposite direction than expected, with breathy onsets having a higher HNR. However, the higher values of HNR for breathy phonation might be a result of a prominent first harmonic, which is typically found for breathy phonation. Indeed, the higher values of H1*–H2* for breathy phonation in English and Hmong are likely due to the prominence of H1. This effect is not found to be significant for breathy–creaky contours except in the first ninth. HNR is lower in creaky phonation than in modal phonation in the latter two ninths of the vowel for breathy–creaky, and in the final ninth for modal–creaky. The differentiation between modal and creaky phonation is in the predicted direction, with creaky values having lower HNR values, presumably due to decreased periodicity.

The results for HNR for Hmong in Fig. 6 indicate that this measure differentiates between non-contrastive breathy and modal phonation during the first four ninths for breathy–modal. For the modal–creaky contour, HNR is lower than the modal–modal contour after the first ninth, suggesting that the creakiness of the contour is responsible for the early drop in the measure. Indeed, for both contours ending in creaky voice, HNR has lower values for creaky phonation for nearly the whole duration. This suggests that HNR in Hmong is reflecting mostly the noise due to creaky phonation, rather than that of breathy phonation derived from coarticulation with an aspirated segment. Interestingly, it is the non-contrastive breathy onsets that have the higher HNR values at the vowel onset for English and Hmong, suggesting that the breathy voicing after aspiration is more periodic. For the contrastive breathy–breathy contour, HNR is lower than for the modal–modal contour for the final two-thirds of the vowel, which suggests that in contrast to the non-contrastive breathiness, the phonemic breathiness in Hmong is noisy.

3.3. *Spatial differentiation*

As mentioned earlier, Blankenship (2002) also found that contrastive non-modal phonation is more differentiated from modal than the non-modal phonation derived from coarticulation. Since I found non-modal phonation to be most differentiated from modal at the onsets and offsets of the vowel, this claim is tested for breathy–modal contours at the first ninth, and for modal–creaky ones at the final ninth, by comparison with the modal–modal contours. Differences in the level of breathiness are assessed using the best measures found for differentiating breathy from modal phonation in the two languages, H1*–H2* and H1*–A1*. Differences in the level of creakiness are assessed using HNR, which was found to be significant in differentiating modal from creaky phonation in the two languages (Tables 2 and 3).

The HNR difference for Hmong modal vs. creaky is higher than for English. These results confirm that non-modal phonation, when contrastive, can be more differentiated from modal. Although the modal vs. non-contrastive breathy differences in H1*–H2* and H1*–A1* do not differ much between the two languages, it is surprising that the contrastive breathiness of Hmong differs from modal less than the non-contrastive breathiness. This may be due to the fact that phonemic breathy voice in Hmong is differentiated from modal vowels over the entire duration of the vowel, so large differences in degree of breathiness are not needed.

Table 2
Absolute differences (in dB) in measures of breathiness (H1*–H2* and H1*–A1*) between breathy–modal and modal–modal contours at the first ninth. The difference between contrastive breathy–breathy and modal–modal for the Hmong is included in the table for comparison.

	H1*–H2*	H1*–A1*
English—*non-contrastive breathiness*	4.59	8.95
Hmong—*non-contrastive breathiness*	3.42	10.84
(cf. Hmong—*contrastive breathiness*)	1.61	2.21

Table 3
Absolute differences (in dB) in HNR between modal–creaky and modal–modal contours at the final ninth.

	HNR
English—*non-contrastive creakiness*	7.09
Hmong—*contrastive creakiness*	16.75

3.4. Summary of results

The results show that breathy is differentiated from modal phonation mostly by H1*–H2* and H1*–A1*. For English, initially-breathy contours are breathy for the first half of the vowel (about 50–60 ms), in line with previously found results, e.g. those obtained by Löfqvist and McGowan (1992). In Hmong, the breathiness derived from coarticulation is differentiated from modal phonation in more intervals for the breathy–modal than for the breathy–creaky contour, because creakiness starts early in Hmong. For H1*–H2* the effects of phonemic breathiness are seen throughout the vowel's duration. Thus, non-contrastive breathiness can be found in more timing intervals in English than in Hmong, but this depends on the vowel's phonation contour. Further, because the Hmong vowels were about 60% longer than English ones (see Table 1), it is not the case that the breathiness in English is longer in duration than in Hmong, even though it can be found over more portions of the vowel.

Creakiness is manifested most strongly by the HNR values. For English, the difference between modal–creaky and modal–modal vowels is made at the vowel offset, in the latter third of the vowel. For Hmong however, creakiness starts after the first ninth and persists throughout the remainder of the vowel. Interestingly, the creakiness occurs in the same timing intervals for which H1*–H2* and H1*–A1* show breathy phonation. Hmong creaky phonation is also more differentiated from modal phonation than English creaky phonation, but the two language's non-contrastive breathiness have the same degree of differentiation from modal phonation. Surprisingly, Hmong phonemic breathiness is less differentiated from modal than its non-contrastive breathiness.

4. General discussion

4.1. The acoustic measures

Many acoustic studies of voice quality use several measures to characterize non-modal phonation. The use of multiple cues rests on the assumption that phonation is multi-dimensional in its articulation (Edmondson & Esling, 2006; Gordon & Ladefoged, 2001) and in its perception (Esposito, 2006, 2010a; Gerfen & Baker, 2005), as well as from cross-linguistic work showing that listeners of different languages rely on different cues to perceive phonation differences (Esposito, 2006, 2010a).

Since acoustic measures of phonation are thought to reflect various articulatory postures from which non-modal voice arises, we can speculate on the articulatory origins of the acoustic results found in this study. First, the major contribution of H1*–A1* to the breathy category of the contours in the three languages suggests that aspiration noise in English and Hmong contains strong whispery components in addition to incomplete closure of the folds, which is arguably correlated with H1*–H2*. The contrastive breathiness in Hmong is manifested less by H1*–A1* and more by H1*–H2* and HNR, suggesting that it is produced differently than the breathiness resulting from coarticulation. This has been found before for the breathiness in Hmong vowels following a voiced aspirated stop (Esposito & Khan, 2010; Fulop & Golston, 2008), but further study is needed to see how such breathiness differs from the kind shown in this study. Creakiness in English and Hmong pattern similarly, and are best differentiated from modal voice by the harmonics-to-noise ratio measure, but also by H1*–H2*.

It is also interesting that the acoustic measures have generally higher values for Hmong speakers than for English speakers. This could accidental, i.e. it could be that the Hmong speakers sampled in this study happen to have breathier voices than the given English participants. However, it is also possible that in Hmong, the baseline for modal phonation is slightly breathier than in English. Other cross-linguistic work on phonation has shown than languages differ greatly in their realization of modal phonation (Keating, Esposito, Garellek, Khan, & Kuang, 2011). Why languages differ in this regard is still unclear and worthy of further investigation.

4.2. The sequencing of coarticulated non-modal phonations

The results indicate that breathy–creaky contours are observable in English and Hmong. Thus, it is not the case that the presence of one non-modal phonation type on a vowel precludes the presence of some other, even opposing phonation, at least for the languages studied here. In fact, different phonation types are known to co-occur on a single vowel in Chong (DiCanio, 2009) and in some Zapotecan languages (Munro, Lopez, Méndez, Garcia, & Galant, 1999). Although both breathy and creaky portions surface in contours here, the languages differ in their timing of each portion. The results concerning the timing of phonation are partially in agreement with those of Blankenship (2002), in that the non-contrastive breathiness of Hmong is limited in duration compared to the creakiness or the phonemic breathiness. In contrast, English breathiness and creakiness were strongest in the initial and final third of the vowel's duration, respectively. Timing of non-modal phonation here is assumed to be correlated with its duration, because the contours analyzed were of approximately the same duration. Therefore, this study indicates that, when two non-modal phonation types are juxtaposed, the contrastive type will last longer.

Can these timing differences be due to inherent timing disparities across differing types of non-modal phonation? This is unlikely, given that the non-contrastive and contrastive breathiness in Hmong were manifested using the same acoustic cues. The fact that the same measures show differences in timing for contrastive vs. non-contrastive phonation suggests that timing differences are not solely due to different non-modal phonation types.

The presence of modal voice in the breathy–creaky contours is also of interest. In English, the vowels start off breathy, then become modal, and finally end creaky. On the other hand, breathy–creaky contours in Hmong are dominated by creak, which lasts for most of the vowel and co-occurs with the breathiness derived from coarticulation. In Hmong, no modal portions appear in these contours. The simultaneity of breathy and creaky voices runs counter to the basic model of phonation involving only the glottal opening (e.g. in Gordon & Ladefoged, 2001; Ladefoged, 1971), given that the glottis cannot be both open and closed at once. The findings do lend support, however, to models of phonation involving either the vocal folds as a whole (e.g. Hanson et al., 2001; Laver, 1980), or the entire laryngeal system, such as the Laryngeal Articulator Model (Edmondson & Esling, 2006; Esling, 2005). Nevertheless, it should be noted that these acoustic findings are merely presumed to be correlated with aerodynamic or articulatory postures, so future articulatory research would be needed to confirm both the simultaneity of breathy and creaky phonation as well as the nature of these voice qualities for the languages at hand.

4.3. Contrast and phonation coarticulation

This study shows that the creakiness in English is concentrated at the end of the vowel and is not strongly differentiated from modal phonation. This finding is similar to that by Blankenship (2002), who showed for other languages that contrastive non-modal phonation is more pronounced and lasts longer than non-modal phonation resulting from coarticulation. In Hmong, however, the contrastive creakiness is more pronounced and is differentiated from modal phonation for most of the vowel. Since contrastive features require that they be perceptually salient in order to distinguish words of a language, it is not surprising that contrastive non-modal phonation should be present in more intervals and be well-differentiated from modal phonation.

However, Blankenship's predictions do not hold as well for non-contrastive vs. contrastive breathiness in Hmong. The results of this study show that breathiness following aspirated sounds is more differentiated from modal than contrastive breathiness. Nevertheless, the former is restricted to the first third of the vowel, whereas the latter is widespread throughout the vowel. Therefore, although Blankenship's timing predictions are borne out in this study, her predictions about differences in the degree of differentiation between non-contrastive and contrastive phonation types are only found for creakiness. Therefore, it seems as though the timing and spatial properties of non-modal phonation can trade off with one another.

It is also interesting that non-modal phonation which is non-contrastive should consistently be restricted to small portions of the vowel. Coarticulatory effects are not always found to be short-lived. According to Manuel and Krakow (1984) and Manuel (1987, 1990), coarticulation can be strongly influenced by the number of contrasts in a language. Assuming that languages have output constraints on coarticulation that are influenced by the number of contrasts to be maintained, a language should not have to limit its coarticulation of a certain feature, if that feature is not contrastive. It is therefore interesting that English should have such limited non-modal phonation. In Hmong, the breathiness derived from coarticulated aspiration is timed with only the early parts of the vowel perhaps because the language must contrast modal, creaky, and breathy vowels. For English, such an explanation would not hold, because the language does not contrast phonation. Even in the breathy–modal contours, which for English ended with slight breathiness due to the coda /s/, breathiness derived from the aspirated stop drops after a third of the vowel's duration, only to rise again for the /s/. This seems to imply that English vowels have a more modal target. The non-contrastive breathy phonation in English behaves similarly to the breathiness occurring after aspirates in Tagalog and Navaho, which also showed rapid onsets and offsets (Blankenship, 1997). To account for this, Blankenship hypothesized that in these languages, vowels are featurally specified for modal voice. Indeed, Cohn (1990, 1993) found that gestural coarticulation for nasality typically shows a sharp decline when the language has a feature specification corresponding to the gesture, but not when there is underspecification. Another possible explanation for modal target effect seen for English could be one of perceptual ease, given that modal phonation is thought to allow for better cue retrieval by listeners (Silverman, 1995; but cf. Garellek, 2011b). Whether English vowels tend towards modal phonation in order to increase perceptual recoverability or due to feature specification is left for future study.

5. Conclusions

The goals of this study were twofold. The first was to show that breathy–creaky contours in vowels can be found in the world's languages, even in English which lacks contrastive phonation. These contours can be described using common measures of phonation. Measures of spectral tilt like H1*–H2* and H1*–A1*, as well as noise measures like HNR, are useful for distinguishing either the breathy or creaky portions of such contours, or in some cases both types of non-modal phonation.

The second goal was to account for cross-linguistic differences in the timing and sequencing of the contours using previous work on phonation timing and theories of coarticulation. The findings of this study support the findings of Blankenship (1997, 2002) that non-modal phonation is longer when it is contrastive. The creakiness in Hmong is found throughout the vowel and is more pronounced than the non-contrastive form in English. However, this study also shows, contra Blankenship, that contrastive non-modal phonation may be less differentiated from its non-contrastive counterpart, as is the case with contrastive vs. non-contrastive breathiness in Hmong. The more restricted timing of non-modal phonation derived from coarticulation might be attributed to the perceptual benefits of modal phonation. This study adds to our understanding of the production of phonation by showing that codas can alter a vowel's phonation significantly, that rapid changes in phonation within vowels are possible and likely more common than assumed, and that breathiness can be found to co-occur simultaneously with creakiness in vowels.

Acknowledgments

I would like to thank Patricia Keating, Sun-Ah Jun, Kie Zuraw, Megha Sundara, Jody Kreiman, and the anonymous reviewers as well as D.H. Whalen for their insightful comments. I am also grateful to Maxine Van Doren and Kristine Yu for assisting me in recruiting and testing participants, and to Yen-Liang Shue for help with VoiceSauce. For help with piloting the experiment, I would like to thank Mel Bervoets, Chad Vicenik, and Jamie White. This work was partially supported by an FQRSC graduate fellowship 128451 to the author and NSF Grant BCS-0720304. Ancillary materials can be found at the author's website: http://www.linguistics.ucla.edu/people/grads/mgarellek/ancillary.htm.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.wocn.2011.10.003.

References

Abramson, A. S., L-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.
Avelino, H. (2010). Acoustic and electroglottographic analyses of nonpathological, nonmodal phonation. *Journal of Voice*, *24*, 270–280.
Baayen, R. H. (2007). *Analyzing Linguistic Data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
Bickley, C. (1982). Acoustic analysis and perception of breathy vowels. In *Speech communication group working papers* (pp. 71–82). Cambridge: Massachusetts Institute of Technology.
Blankenship, B. (1997). *The time course of breathiness and laryngealization in vowels*. Ph.D. Dissertation. Los Angeles: University of California.
Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*, 163–191.
Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer* (version 5.1.14) [Computer program]. Retrieved August 30, 2009, from 〈http://www.praat.org/〉.
Cohn, A. (1990). *Phonetic and phonological rules of nasalization*. Ph.D. Dissertation. Los Angeles: University of California [Published in UCLA Working Papers in Phonetics, 76].
Cohn, A. (1993). Nasalisation in English: Phonology or phonetics. *Phonology*, *10*, 43–81.
de Krom, G. (1993). A cepstrum-based technique for determining a harmonics-to-noise ration in speech signals. *Journal of Speech and Hearing Research*, *36*, 254–266.

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

DiCanio, C.T. (2011). Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, doi:10.1016/j.wocn.2011.10.006.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.

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

Epstein, M. (1999). *A comparison of linguistic and pathological breathiness using the LF model*. M.A. Thesis. Los Angeles: University of California.

Epstein, M. (2002). *Voice quality and prosody in English*. Ph.D. Dissertation. Los Angeles: University of California.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics*, *50*, 13–44.

Esposito, C. M. (2006). *The effects of linguistic experience on the perception of phonation*. Ph.D. Dissertation. Los Angeles: University of California.

Esposito, C. M. (2010a). Variation in constrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association*, *40*, 181–198.

Esposito, C. M. (2010b). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, *38*, 306–316.

Esposito, C. M., Ptacek, J., & Yang, S. (2009). An acoustic and electroglottographic study of White Hmong phonation. *Journal of the Acoustical Society of America*, *126*, 2223.

Esposito, C. M., & Khan, S. (2010). Contrastive breathiness on consonants and vowels. *Poster presented at the 2nd Pan-American/Iberian meeting on acoustics*. Cancun, Mexico.

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

Fulop, S., & Golston, C. (2008). Breathy and whispery voice in White Hmong. In *Proceedings of meetings on acoustics* (Vol. 4). 060006-1-10.

Garellek, M. (2010). The acoustics of coarticulated non-modal phonation. *UCLA Working Papers in Phonetics*, *108*, 66–112.

Garellek, M. (2011a). Lexical effects on English vowel laryngealization. In *Proceedings of the 17th international congress of phonetics sciences*.

Garellek, M. (2011b). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics*, *109*, 31–39.

Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Mazatec. *Journal of the International Phonetic Association*, *41*, 185–205.

Gerfen, C., & Baker, K. (2005). The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, *33*, 311–334.

Gerratt, B. R., & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics*, *29*, 365–381.

Gobl, C., & Ní Chasaide, A. (1999). Voice source variation in the vowel as a function of consonantal context. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 122–143). Cambridge: Cambridge University Press.

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

Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *MIT Quarterly Progress Report*, *101*, 198–212.

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

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

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

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

Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics*, *33*, 335–362.

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

Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication*, *27*, 187–207.

Keating, P., Esposito, C., Garellek, M., Khan, S., & Kuang, J. (2011). Phonation contrasts across languages. In *Proceedings of the international conference of phonetic sciences (ICPhS 17)*. Hong Kong.

Kreiman, J. (1982). Perception of sentences and paragraph boundaries in natural conversation. *Journal of Phonetics*, *10*, 163–175.

Kreiman, J., Iseli, M., Neubauer, J., Shue, Y.-L., Gerratt, B. R., & Alwan, A. (2008). The relationship between open quotient and H1*–H2*. *Journal of the Acoustical Society of America*, *124*, 2495.

Kuang, J. (2011). *Production and perception of the phonation contrast in Yi*. M.A. Thesis. UCLA.

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

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).

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

Lisker, L., & Abramson, A. S. (1964). Cross-language study of voicing in initial stops: Acoustic measurements. *Word*, *20*, 384–422.

Löfqvist, A., & McGowan, R. S. (1992). Influence of consonantal environment on voice source aerodynamics. *Journal of Phonetics*, *20*, 93–110.

Manuel, S. (1987). *Acoustic and perceptual consequences of vowel-to-vowel coarticulation in three Bantu languages*. Ph.D. Dissertation. Yale University.

Manuel, S. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America*, *88*, 1286–1298.

Manuel, S., & Krakow, R. (1984). *Universal and language particular aspects of vowel-to-vowel coarticulation* (pp. 69–78). Haskins Laboratories Status Report on Speech Research, SR-77/78.

Miller, A. L. (2007). Guttural vowels and guttural coarticulation in Ju|'hoansi. *Journal of Phonetics*, *35*, 56–84.

Munro, P., Lopez, F. H., Méndez, O. V., Garcia, R., & Galant, M. R. (1999). *Di'csyonaary X:tèe'n Dìi'zh Sah Sann Lu'uc (San Lucas Quiaviní Zapotec Dictionary/Diccionario Zapoteco de San Lucas Quiaviní)*. Los Angeles: Chicano Studies Research Center Publications, UCLA.

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, *39*, 151–168.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II* (pp. 90–117). Cambridge: Cambridge University Press.

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

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*, 407–429.

Selkirk, L. (1972). *The phrase phonology of English and French*. Ph.D. Dissertation. Massachusetts Institute of Technology.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In *Proceedings of the international conference of phonetic sciences (ICPhS 17)*. Hong Kong.

Silverman, D. (1995). *Phasing and recoverability*. Ph.D. Dissertation. Los Angeles: University of California [Published by Garland 1997].

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

Sjölander, K. (2004). *The Snack Sound Toolkit [Computer program]*. Retrieved May 25, 2010, from 〈http://www.speech.kth.se/snack/〉.

Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice*, *20*, 171–186.

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

Stevens, K. N., & Hanson, H. M. (1995). Classification of glottal vibration from acoustic measurements. In O. Fujimura, & M. Hirano (Eds.), *Voice quality control* (pp. 335–342). San Diego: Singular Publishing Group, Inc.

Sundberg, J., Andersson, M., & Hultqvist, C. (1999). Effects of subglottal pressure variation on professional baritone singers' voice sources. *Journal of the Acoustical Society of America*, *105*, 1965–1971.

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

West, P. (1999). The extent of coarticulation of English liquids: An acoustic and articulatory study. In *Proceedings of the international conference of phonetic sciences* (pp. 1901–1904). San Francisco.

Westbury, J. R., & Niimi, S. (1979). An effect of phonetic environment on voicing control mechanisms during stop consonants. In J. Wolf, & D. Klatt (Eds.), *Speech communication papers presented at the 97th meeting of the Acoustical Society of America*. New York: Acoustical Society of America.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*, 61–83.