Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong

Christina M. Esposito
Department of Linguistics,
Macalester College
esposito@macalester.edu

Sameer ud Dowla Khan
Department of Cognitive, Linguistic, and Psychological Sciences,
Brown University
sameeruddowlakhan@gmail.com

Gujarati and White Hmong are among a small handful of languages known to maintain a phonemic contrast between breathy and modal voice across both obstruents and vowels. Given that breathiness on stop consonants is realized as a breathy-voiced aspirated release into the following vowel, how is consonant breathiness distinguished from vocalic breathiness, if at all? We examine acoustic and electroglottographic data of potentially ambiguous CV sequences collected from speakers of Gujarati and White Hmong, to determine what properties reliably distinguish breathiness associated with stop consonants from breathiness associated with vowels comparing both within and across these two unrelated languages. Results from the two languages are strikingly similar: only the early timing and increased magnitude of the various acoustic reflexes of breathiness phonetically distinguish phonemic consonantal breathiness from phonemic vocalic breathiness.

1 Introduction

Numerous languages exhibit contrastive breathy-voiced phonation either on obstruent consonants as in Hindi (e.g. Ohala 1983, Dixit 1989), Bengali (Khan 2010a), and Maithili (Yadav 1984) or on vowels as in many Zapotec languages (e.g. Jones & Knudson 1977, Munro & Lopez 1999, Esposito 2010b). However, very few languages preserve this contrast across both obstruent consonants AND vowels. While languages such as Suai (Abramson & Luangthongkum 2001), Jalapa Mazatec (Kirk, Ladefoged & Ladefoged 1993), and Wa (Watkins 1999, 2002) contain both breathy vowels and VOICELESS aspirated consonants, languages that include both breathy vowels and BREATHY-VOICED aspirated consonants (also known as breathy-aspirates) are exceptionally rare. This latter type appears to be limited to some Khoisan languages (e.g. !Xóõ, see Traill 1985; Ju|'hoansi, see Miller 2007), White

Journal of the International Phonetic Association (2012) 42/2
doi:10.1017/S0025100312000047

Hmong, and Gujarati. Languages such as these are particularly interesting because the breathiness on breathy-voiced aspirated stop consonants is typically realized not during the stop closure itself, but as a breathy-voiced release into the following vowel. Thus, both breathy vowels and prevocalic breathy-voiced aspirated stops involve breathy voicing during the vowel.

Previous research suggests that the voice quality of breathy vowels and breathy-voiced aspirated consonant releases should be similar from an articulatory standpoint. Ladefoged & Maddieson (1996) define both breathy-voiced aspirated stops (which they refer to as 'murmur' following the terminology used in Ladefoged 1971) and breathy-voiced vowels in the same way; both involve vocal folds that vibrate without much contact and high rates of airflow. In comparing the Hindi minimal pair [bal] 'hair' and [bʱal] 'forehead', they observed breathy voicing for the first 100 ms following the stop release for the breathy-voiced aspirated stop [bʱ]. In their fiberoptic study of one speaker of Hindi, Kagaya & Hirose (1975) observed that breathy-voiced aspirated consonants were largely identical to plain voiced stops up until the consonant release; after that point, glottal width increased, although not nearly to the extent seen in voiceless aspiration. This intermediate glottal width is key in maintaining voicing while allowing enough space for breathy airflow, further facilitated by the lack of supraglottal constriction. Kagaya & Hirose's acoustic analysis also indicates that breathy-voiced aspirated consonants in Hindi have a significantly lowered f0 at the consonant release relative to all other stop consonant types, consistent with findings in other non-tonal languages indicating a correlation between breathy vowels and lower f0 (see Pandit 1957, Dave 1967, and Fischer-Jørgensen 1967 for Gujarati; Wayland, Gargash & Jongman 1994 for Javanese). Furthermore, in a photo-electroglottographic study on plosives in Hindi, Dixit (1989) found that the breathy-voiced aspirated consonants were produced by slack vocal folds, a moderately open glottis, a high rate of oral airflow, and a random distribution of noise. These four characteristics of breathy-voiced aspirated consonants are also properties of breathy vowels (see Gordon & Ladefoged 2001). Fischer-Jørgensen (1967) also described breathy-voiced aspirated stops in Gujarati as similar to a breathy vowel, with the main difference being the degree of noise.

However, other descriptions suggest that breathy-voiced aspirated consonants and breathy vowels are distinct. Laver (1981) defines the phonation of breathy vowels as involving low muscular effort, thus producing a wide glottis, while he defines the phonation of whispered voice (in which he includes breathy-voiced aspirated consonants) as involving a manipulation of the arytenoids such that the vocal folds vibrate modally along their length but with a posterior gap through which air flows continuously. Esling & Harris (2005), on the other hand, posit that the difference between whispery and breathy voice is not due to degree of glottal constriction, but rather due to an engagement of the aryepiglottic sphincter during whispery voice.

The current study examines acoustic and electroglottographic data collected from Gujarati and White Hmong to determine what properties reliably distinguish vowels following breathy-voiced aspirated consonant releases (e.g. [Cʱ**V**]) from phonemically breathy vowels (e.g. [C**V̤**]), and to explore the phonetic and phonological properties shared between these structures in the two genetically-unrelated languages. Given that breathiness on consonants is typically realized as a breathy-voiced release into the following vowel, how are the two types of breathiness distinguished in CV sequences, if at all?

We hypothesize that the difference between these segments is likely one of timing and/or degree of breathiness. In terms of timing, we predict that the breathiness associated with breathy-voiced aspirated consonants is localized to the consonant release and thus produced at the onset of a following vowel, while the breathiness associated with breathy vowels is produced across a larger portion of the vowel, with language-specific distinctions in its exact localization. We also predict that post-aspirated vowels exhibit a different degree (i.e. more or less breathiness) of breathiness than breathy vowels.

2 Background

Phonation contrasts can be made using a variety of articulatory mechanisms, which produce an array of acoustic effects available to the listener for the perception of linguistic voice quality. To investigate these dimensions of phonation, we begin by reviewing the acoustic properties of phonation contrasts and continue with electroglottographic properties, the two types of measurements used in the current study. To minimize undue repetition, we restrict the following review to languages other than Gujarati and White Hmong; these two languages of interest are discussed in much greater detail in Section 2.3.

2.1 Acoustic properties of phonation

Often the most robust acoustic differences between phonation types can be seen in the spectrum; breathy phonation has a more sharply falling spectrum than modal phonation, while creaky phonation is often characterized as having a nearly flat spectrum. This steepness in the spectrum can be measured as spectral balance or spectral slope. Spectral balance is defined as the difference between the amplitude of the first harmonic (H1) from that of the second harmonic (H2), i.e. H1-H2, and has been used to measure phonation in languages as diverse as Jalapa Mazatec (Blankenship 2002, Garellek & Keating 2010), !Xóõ (Bickley 1982), Chanthaburi Khmer (Wayland & Jongman 2002), Green Mong (Andruski & Ratliff 2000), Takhian Thong Chong (DiCanio 2009), and Santa Ana del Valle Zapotec (Esposito 2010b). Spectral slope is measured as the difference between the amplitude of the first harmonic (H1) and that of harmonics exciting higher formants, i.e. H1-A1, H1-A2, and H1-A3. H1-A1 has been shown to reliably distinguish phonation types in !Xóõ (Ladefoged 1983) while H1-A2 distinguished phonation types in Krathing Chong (Blankenship 2002). More commonly seen in the literature is H1-A3, which distinguishes phonation types in English (Stevens & Hanson 1995), Krathing Chong (Blankenship 2002), Takhian Thong Chong (DiCanio 2009), and Santa Ana del Valle Zapotec (Esposito 2010b). Esposito's (2010a) cross-linguistic study also looked at small sets of data in Krathing Chong, Fuzhou, Green Mong, White Hmong, Mon, San Lucas Quiaviní Zapotec, Santa Ana del Valle Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ, finding that spectral balance (i.e. H1-H2) and one or more of these three common measures of spectral tilt (i.e. H1-A1, H1-A2, H1-A3) differentiated phonation types in each language. In calculating spectral tilt or spectral balance, the amplitudes of harmonics can be corrected for the effects of the frequencies and bandwidths of adjacent formants (Hanson 1995); in this case, an asterisk (*) can be used to signify a corrected amplitude, e.g. H1*-A3*, a convention we adopt here.

Other spectral measures discussed in the voice quality literature include the difference in amplitude between the second and fourth harmonics (H2-H4), for measuring pathological voice quality (Kreiman, Gerratt & Antoñanzas-Barroso 2006), the average of H1-H2 compared to A1, for measuring non-contrastive voice quality in English (Stevens 1988), and formant amplitude differences such as A2-A3 in English (Klatt & Klatt 1990). These are, however, not widely used in studies of linguistically contrastive voice quality.

Acoustic measures of the spectrum have been associated with various physiological characteristics. Holmberg et al. (1995) showed that H1-H2 correlated with the open quotient (OQ) of the glottal cycle, i.e. the portion of time the vocal folds are open per cycle. The larger the open quotient (i.e. the longer the vocal folds are apart), the greater the amplitude of the first harmonic over that of the second harmonic. Thus, the value (in dB) of H1-H2 is higher for breathy phonation than for modal or creaky phonation. Furthermore, Stevens (1977) suggested that spectral tilt measures could be correlated with the abruptness of vocal fold closure. More abrupt vocal fold closure excites the higher harmonics; for breathy phonation, which typically involves less abrupt vocal fold closure, the higher harmonics are weakened, and thus spectral tilt measures are higher for breathy phonation than for modal or creaky phonation.

Depending on the language, dialect, vowel quality, tone, speaker sex or gender, and other factors, not all spectral measures will distinguish phonation types. In Mpi, for example, H1-H2 distinguishes phonation types on high tone vowels more reliably than on mid or low tone vowels (Blankenship 2002). In Santa Ana del Valle Zapotec, H1-H2 successfully distinguishes breathy, modal, and creaky phonation in female speech but not in male speech (Esposito 2010b). Kreiman, Gerratt & Antoñanzas-Barroso (2007) showed that f0 was positively correlated with H1*-H2* in non-disordered and pathological productions of the vowel [a], while Iseli, Shue & Alwan (2007) found that H1*-H2* was positively correlated with f0 only for speakers whose pitch was lower than 175 Hz.[1] Because females generally speak in a higher pitch than males, some of the sex-specific effects of H1*-H2* may be due to its complex relation to f0.

In addition to spectral measures, measures of noise and/or aperiodicity in the signal can also measure differences in voice quality. One such measure, cepstral peak prominence (CPP), has been used in English (Hillenbrand, Cleveland & Erickson 1994); Krathing Chong, Jalapa Mazatec, Mpi, and Tagalog (Blankenship 2002); and for a small set of data from Krathing Chong, Fuzhou, Green Mong, White Hmong, Mon, San Lucas Quiaviní Zapotec, Santa Ana del Valle Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ (Esposito 2010a).

2.2 Electroglottographic properties of phonation

When invasive methods of articulatory research are either unavailable or inappropriate, an electroglottograph (EGG) can be used as an indicator of the degree of contact between the vocal folds over time, which can in turn help describe and categorize phonation types. The EGG has been used to measure linguistic voice quality in Maa (Guion, Post & Payne 2004), Vietnamese (Michaud 2004), Santa Ana del Valle Zapotec (Esposito 2005), Tamang (Michaud & Mazaudon 2006), Takhian Thong Chong (DiCanio 2009), and Yi (Kuang 2010, 2011), and non-linguistic voice quality in speakers with and without voice disorders. For example, Childers & Lee (1991) used EGG measures to determine the characteristics of the voice source during modal voice, vocal fry, falsetto, and breathy voice in both normal and pathologically-disordered voices. They found that breathy phonation (as well as falsetto) was produced with a longer glottal pulse width, lower pulse skewing (the ratio of the opening phase to the closing phase), and less abrupt glottal closure than modal phonation. Using acoustic data, they also found that breathy phonation was produced with high turbulent noise, not seen in the other voice qualities.

The most common measure derived from the EGG is CQ, variously referred to as contact quotient, closed quotient, and closing quotient. CQ is a ratio of the portion of time the vocal folds are in a greater degree of contact over the total duration for a complete glottal cycle. In the current study, calculating the edges of this portion of this 'greater degree of contact' involves a hybrid method with a 25% threshold (see Rothenberg & Mahshie 1988, Orlikoff 1991, Howard 1995, and Herbst & Ternström 2006). This means that the beginning of the contact/closure phase (the portion with the 'greater degree of contact') is defined as the point at which the first derivative of the EGG (dEGG) is at its peak, and the end of the contact/closure phase is defined as the point 25% from the point of greatest opening (where 25% is calculated from the time from closure peak to opening peak). CQ is the inverse of the open quotient measure (OQ). Acoustic and electroglottographic studies of contrastive voice quality/register in Takhian Thong Chong (DiCanio 2009) and White Hmong (Esposito, in press) compared OQ with H1-H2 and H1-A3, finding that OQ was more closely correlated with H1-H2 than with H1-A3, confirming Holmberg et al.'s (1995) study. Assuming a unidimensional

[1] The Iseli et al. (2007) study also found that H1*-H2* was dependent on vowel height for speakers whose pitch was higher than 175Hz.

model of phonation based on glottal opening (Ladefoged 1971, Gordon & Ladefoged 2001), phonations with a wider opening (e.g. breathy voice) are expected to have lower CQ values than do phonations with greater vocal fold contact (e.g. modal voice, creaky voice).

The first derivative of the EGG, dEGG, is also useful in measuring voice quality. The peak positive value in the dEGG for each glottal pulse represents the amplitude of the increase in contact between the vocal folds; this value is variably referred to as Peak Increase in Contact (PIC; see Keating et al. 2010) or as dEGG Closure Peak Amplitude (DECPA; see Michaud 2004 for Mandarin, Naxi and Vietnamese; see Vũ-Ngọc, d'Alessandro & Michaud 2005 for Vietnamese). In this way, DECPA can represent the speed of the vocal folds during the closing phase; phonations produced with faster glottal closure have greater DECPA values than phonations produced with slower glottal closure.[2] Of course, the vocal folds need not actually fully close to derive a DECPA value, as what is being measured is the increase in contact between the folds. It is not uncommon in breathy phonation and similar voice qualities for the folds to come into contact while still leaving a partially open glottis, allowing air to pass through.

2.3 About the languages

2.3.1 Gujarati

Gujarati is an Indo-European language (Indo-Iranian branch, Central Indic group) spoken primarily in Gujarat state in India, with significant minority populations in other central-western Indian states including Maharashtra (with a large community in Mumbai), Rajasthan, Karnataka, and Madhya Pradesh, and in long-established immigrant communities throughout the UK, North America, East Africa, and elsewhere (Lewis 2009).

Like other Indic languages, Gujarati has a four-way contrast in voicing and aspiration in stops and affricates, including voiceless unaspirated, voiceless aspirated, modally-voiced unaspirated, and breathy-voiced aspirated consonants across five places of articulation: bilabial, dental, retroflex, alveolopalatal (affricate), and velar (Nair 1979, Masica 1993, Cardona & Suthar 2003).[3] In the vocalic inventory, the most conservative dialects show an eight-vowel system [i e ɛ a ə ɔ o u] in modal phonation, while other dialects (e.g. Saurashtra) show a six-vowel system [i e a ə o u] (Firth 1957: 231–232; Pandit 1961: 62–63). Gujarati also has a set of breathy vowels, most of which are modern reflexes of what were once sequences of vowels and breathy consonants (Pandit 1957: 169–170; Dave 1967: 1–2; Fischer-Jørgensen 1967: 73; Nair 1979: 9; Masica 1993: 120; Mistry 1997: 666–669; Cardona & Suthar 2003: 665–666).[4] Breathy vowels that derive from such structures come in four types, based on their historical source sequence. One very common source is [əɦV]; breathy vowels [ə̤ a̤ ɛ̤ ɔ̤ ə̤j ə̤w] are the modern reflex of what is historically and orthographically [əɦə əɦa əɦe əɦo əɦi əɦu], respectively (e.g. [ba̤ɾ] 'outside', orthographically બહાર ⟨bəɦaɾə⟩).

Less frequent sources of breathy vowels include [Vɦə], [#ɦ], and [VCʱ]. Historical and orthographic [Vɦə] is optionally rendered as a single breathy vowel in modern Gujarati, e.g. [ʋa̤n] ~ [ʋaɦən] 'vehicle'. In very casual speech, a third type of breathy vowel comes as the result of the optional lenition of word-initial [ɦ], as in [ṳlːəɽ] ~ [ɦulːəɽ] 'riot'. Lastly, post-vocalic breathy-voiced aspirated consonants [bʱ d̪ʱ ɖʱ dʑʱ gʱ] optionally lose their aspiration in

[2] For more information on EGG measures see Childers & Krishnamurthy (1985), Baken & Orlikoff (2000), and Henrich et al. (2004).

[3] Some dialects do not preserve all stop/affricate consonant contrasts; many speakers produce fricatives in place of (typically aspirated) consonants, including [f] in place of [pʰ], [z] in place of [ɟʑ] and/or [ɟʑʱ], and [ɕ] in place of [cɕʰ] (Firth 1957: 235; Cardona & Suthar 2003: 663–665).

[4] A small set of words such as [kə̤ɾo] 'wall' and [na̤nũ] 'small' contain breathy vowels believed to not be derived from sequences of modal vowels and [ɦ] (Masica 1993: 147; Mistry 1997: 668; Cardona & Suthar 2003: 666).

very casual speech, with their associated breathiness transferred to surrounding vowels; they can also be lenited to fricatives or approximants in these situations, e.g. [bəd̤ʱũ] ~ [bə̤d̤ṳ̃] ~ [bə̤ðṳ̃] ‘whole’ (Firth 1957: 235; Pandit 1957: 171; Mistry 1997: 667; Cardona & Suthar 2003: 666).

Due to various sociolinguistic pressures, breathy vowels are often not produced as such in particular contexts. Pandit (1957: 170), Dave (1967: 2), Nair (1979: 22), and Cardona & Suthar (2003: 666) report that many speakers have merged the breathy vowels with their corresponding modal vowels in what is often described as an ‘educated’ speech register, producing [bɛ̤n] ‘sister’ as [bɛn]. Turner (1921: 529), Dave (1967: 4), Masica (1993: 120), and Cardona & Suthar (2003: 665–666) also report that speakers are more likely to produce breathy vowels as disyllabic sequences reflecting their orthographic representation, especially in formal settings or when reading, e.g. producing [bɛ̤n] ‘sister’ as [bəɦen] or [bəɦɛn], orthographically બહેન ⟨bəɦenə⟩. Breathy vowels with a [əɦV] source are the least likely to be pronounced as a disyllabic sequence, but even words of this source have been reported to be produced in a spelling pronunciation (i.e. disyllabically) in studies such as Dave (1967: 4), where subjects were told to read words directly from a script.[5]

Due to well-known constraints on aspiration in Indic languages (i.e. Grassmann’s Law, see Whitney 1889, Wackernagel 1896), Gujarati does not have monomorphemic sequences of breathy-voiced aspirated consonants and breathy vowels (i.e. *[CʱV̤]); furthermore, the low frequency of breathy segments in borrowed words means that new additions to the lexicon are unlikely to change this characteristic of the language.

Acoustic studies of breathy phonation in Gujarati have been primarily focused on breathy vowels, and less so on breathy aspirated consonants. Fischer-Jørgensen (1967) examined various acoustic measures to determine what properties reliably distinguished breathy vowels from their modal counterparts. Spectral balance, as measured by the amplitude difference between the first and second harmonics (i.e. H1-H2), and three measures of spectral tilt, as measured by the amplitude difference between the first harmonic and the first, second, and fourth formants (i.e. H1-A1, H1-A2, and H1-A4, respectively), were all found to be more sharply falling in breathy vowels. Furthermore, a slightly lowered f0, lower overall intensity (as measured by RMS energy) were found to be characteristics of breathy vowels; aspiration noise was also found in some breathy productions, although this was assessed only visually. An earlier study by Pandit (1957) also found both low f0 and an increase in aspiration noise at higher frequencies to be associated with breathy vowels, while a later study by Bickley (1982) also confirmed that a higher H1-H2 value was a reliable indicator of breathiness. Dave (1967) focused on the formant structure of breathy vowels, finding that they are largely indistinguishable from modal vowels in vowel quality. In the acoustic component of Khan (2010b, 2012), a study of ten Gujarati speakers’ voice quality, it was further confirmed that breathy vowels have a significantly steeper spectral balance (as measured by H1*-H2*) and spectral tilt (as measured by H1*-A3*) than their corresponding modal vowels, concurring with previous studies. Unlike previous studies of Gujarati, however, the data examined in Khan (2010b, 2012) were collected in a more naturalistic setting (as was done for the current study), and the spectral measures were corrected for the effects of formant frequencies and bandwidths (Hanson 1995) using Iseli et al.’s (2007) algorithm, as indicated with the asterisk (*). Furthermore, Khan (2010b) found that the midpoints of breathy vowels had lower CPP values than modal vowels, as well as significantly steeper rises in intensity.

Perception studies of Gujarati breathy vowels largely concur with the main predictions of acoustic studies: while f0 and aspiration noise can have some influence on voice quality categorization, a high H1-H2 value is consistently found to be the strongest cues for breathy voice. In the listening component of her study, Fischer-Jørgensen (1967) determined that the perception of synthesized breathy vowels in Gujarati was largely dependent on the

[5] For an analysis distinguishing dialects based on phonation type, see Modi (1987).

fundamental, which had a low frequency at the onset of the vowel but a high amplitude (H1) throughout (measured relative to the rest of the spectrum); other acoustic cues were determined to be less important for perception. In the perception component of Bickley's (1982) study, it was found that Gujarati speakers rely solely on spectral balance (H1-H2) when categorizing the voice quality of synthesized vowels; aspiration noise did not appear to influence voice quality categorization. Furthermore, in a cross-linguistic study of the perception of linguistic voice quality by speakers of English, Spanish, and Gujarati, Esposito (2010a) also found that Gujarati-speaking listeners rely primarily on H1-H2 differences when categorizing vowels excised from various non-Indic languages (i.e. Krathing Chong, Fuzhou, Green Mong, White Hmong, Mon, Santa Ana Del Valle Zapotec, San Lucas Quiaviní Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ), even in cases where the phonation contrasts in those other languages were not made using differences in H1-H2. Considering this strong bias amongst Gujarati speakers to attend to H1-H2 differences when categorizing vowels, it follows that Gujarati speakers are in fact more sensitive to very small changes in H1-H2 than are speakers of other languages, a hypothesis supported in Kreiman, Gerratt & Khan's (2010) perception study of speakers of English, Thai, and Gujarati.

While the majority of studies of breathy phonation in Gujarati have focused on its acoustic properties and their perception by native speakers, a handful of articulatory studies can also be found in the literature. Fischer-Jørgensen's (1967) study incorporated two articulatory components on a subset of her subjects, including an EGG analysis of two speakers and an aerodynamic analysis of three speakers. She found that breathy vowels are produced with greater airflow and shorter closed phase and possibly a wider glottis. Modi (1987) used x-ray data of the word [kɛ̤ʋət̪] 'proverb' to determine that breathy phonation in 'murmur dialects' such as Standard Gujarati involves a lowered and widened glottis. Most recently, the first large-scale EGG study of Gujarati vowels (Khan 2010b, 2012) found that breathy vowels have a significantly lower contact quotient (CQ) than corresponding modal vowels, signifying that breathy phonation involves a more open glottis than modal vowels. In a further cross-linguistic extension of the Khan (2010b, 2012) study, Keating et al. (2010) showed that this difference in CQ closely resembled the EGG properties of other languages with a phonemic distinction between modal and breathy vowels. To date, there has not been an EGG study of breathy-voiced aspirated consonants in Gujarati.

2.3.2 White Hmong

White Hmong is a Hmong-Mien language spoken in Laos, Thailand, and by a large immigrant community in the US. It contrasts seven tones: rising (45), mid (33), low (22), mid-rising (24), high-falling (52), low-falling (21), and falling (42). Two of the tones are associated with non-modal phonation: the low-falling tone (21) is creaky and the falling tone (42) is breathy.

In addition, White Hmong has a large consonant inventory which includes voiced, voiceless, and prenasalized plosives. A unique feature of White Hmong, that is not found in other varieties such as Green Mong, is a four-way stop contrast within the non-prenasalized alveolar place of articulation [t tʰ d dʱ]; the last consonant of that set, [dʱ], is characterized as a 'whispery voiced alveolar stop, with optional aspiration' in Jarkey (1987: 66). The voiced unaspirated and breathy-voiced aspirated alveolar stops [d dʱ] of White Hmong are modern reflexes of laterally-released velar stops in Proto-Western Hmong [kˡ k^{ɬ}] (Mortensen 2000: 14–15); these correspond to laterally-released alveolar/velar stops [tˡ ~ kˡ t^{ɬ} ~ k^{ɬ}] in other Western Hmong dialects such as Green Mong (Golston & Yang 2001; Mortensen 2004: 3). There is a restriction on the co-occurrence of breathy-voiced aspirated [dʱ] and following vowels bearing the falling breathy tone (*[Cʱ V̤ 42]) or the high-falling tone (*[Cʱ V 52]).

Previous research on the acoustic and electroglottographic properties of phonation in White Hmong showed that the amplitude of the first harmonic (H1*) and derivative-EGG closure peak amplitude (DECPA) are the most successful measures of phonation in that they distinguish all three phonation types (i.e. breathy, modal, creaky), though not all at the same point in the vowel. Other measures distinguish at least two of the three phonation categories. Of particular interest to the current study are the measures that distinguish breathy from modal phonation. The amplitude of the first harmonic (H1*), the amplitude of the first harmonic minus the amplitude of the second harmonic (H1*-H2*) and closed quotient (CQ) distinguish breathy from modal phonation at the beginning of the vowel, while DECPA, H1*, H1*-H2*, and CQ distinguish these phonations at the middle of the vowel and CQ, at the end of the vowel (Esposito 2010c). An additional study, Keating et al. (2010), found that CQ, DECPA (i.e. 'PIC'), and H1*-H2* successfully distinguished the phonation types of White Hmong averaging across the entire vowel duration. To date, there have not been any studies on the perception of phonation by White Hmong listeners.

One study, Fulop & Golston (2008), examined vowels with breathy voice, modal voice, and after breathy-voiced aspirated stops (which they called 'whispery voiced plosives') as produced by two speakers of White Hmong. They measured the amplitude of the first harmonic minus the amplitude of the second (H1-H2) and third harmonics (H1-H3) as well as harmonicity during (i) the consonant release and (ii) the consonant closure phase. During the consonant release, all three measures distinguished all three vowels types. However, during the closure phase, H1-H3 and harmonicity failed to distinguish any of the phonations, while H1-H2 only distinguished the modal from the breathy vowels. Results support the idea that breathy vowels are distinct from vowels after breathy-voiced aspirated consonants. In addition, the higher harmonicity values for vowel after the breathy-voiced aspirated consonants supports Laver's definition for breathy aspiration/whispery phonation, which is posited to involve continuous airflow.

2.4 Previous work on consonant aspiration and vowel breathiness

Apart from the Fulop & Golston (2008) mentioned above, previous research investigating vowel breathiness and consonant aspiration has compared breathy-voiced vowels to modal vowels following VOICELESS aspirated consonants (as opposed to vowels following BREATHY-VOICED aspirated consonants). For example, Watkins (1999) studied phonation in Wa and compared CQ values for breathy vowels to those produced after voiceless aspirated consonants for five timepoints within a vowel. Results showed that there were timing differences between the breathy vowels and the vowels after the aspirated consonants. Breathy vowels began with a higher CQ (i.e. more contact) than vowels after aspirated consonants. However, for the remainder of the vowel, breathy vowels have a lower CQ (i.e. less contact) than the vowels after the aspirated consonants. This trend continued until the last measured timepoint, when the two vowels types had roughly the same CQ. In addition, Garellek & Keating (2010) found that Jalapa Mazatec breathy vowels and modal vowels after voiceless aspirated consonants shared similar values in H1*-H2*, H1*-A1*, H1*-A2*, and CPP.

3 Current study

3.1 Methods

3.1.1 Speakers

3.1.1.1 *Gujarati*

Ten native speakers of Gujarati (three male, seven female) were recorded at the Phonetics Laboratory at UCLA's Linguistics Department.[6] All but two subjects were in their 20s or 30s

[6] As our research questions do not bear on the use of voice quality measures in gender identification or in gender-specific phonation properties, we did not balance the number of speakers across gender lines. See Klatt & Klatt (1990) and Iseli et al. (2007) among others for studies of voice quality differences

Table 1 Speakers, gender, approximate age, birthplace, and number of years in the US.

Language	Speaker	Gender	Age (years, approx.)	Birthplace	Years in the US
Gujarati	1	M	24	India	<1
	2	F	22	India	<1
	3	M	21	India	<1
	4	F	20	India	<1
	5	F	50	India	26
	6	F	29	India	3
	7	F	23	India	<1
	8	F	24	India	<1
	9	F	30	India	<1
	10	M	25	India	<1
White Hmong	1	M	44	Laos	20
	2	M	58	Laos	35
	3	M	35	Laos	20
	4	M	38	Laos	30
	5	M	24	US	24
	6	M	58	Laos	30
	7	F	28	Thailand	16
	8	F	50	Laos	10
	9	F	34	Laos	30
	10	F	24	Laos	17
	11	F	27	US	27
	12	F	28	Laos	24

and had spent the majority of their lives in India, having only recently (<1 year) moved to the US at the time of the recording. Of the remaining two speakers, one was in her 50s and had lived in the US for 26 years, and another was in her 20s and had lived in the US for three years. All subjects were also fluent speakers of English as well as various Indic languages, most commonly Hindi and Marathi, although all reported their first language to be Gujarati. Native fluency in Gujarati was assessed by asking the potential subject questions regarding his or her place of origin and length of stay in the US. All subjects reported that they continued to speak Gujarati on a daily basis and all subjects were fully literate in Gujarati.

3.1.1.2 *White Hmong*

Twelve native speakers of White Hmong (six male, six female) were recorded at the Hmong American Partnership (St. Paul, Minnesota). Speakers ranged from 24 to 58 years of age and were born in Laos, Thailand, or the US, and resided in Minneapolis/St. Paul, Minnesota, at the time of the experiment. Eleven of the speakers spoke English in addition to White Hmong; the reported age of English onset ranged from 5 to 18 years of age. One speaker (Speaker 8) was a monolingual White Hmong speaker. Native fluency in White Hmong was assessed by asking the potential subjects questions regarding his or her place of origin and length of stay in the US. All speakers reported that they used White Hmong daily and all were fully literate in White Hmong.

Table 1 summarizes the background information on the speakers; gender, approximate age, country of birth, and number of years in the United States are given for both Gujarati and White Hmong subjects.

---

across genders. Pilot work for the Gujarati component of the current study (Khan 2010b) suggests that males and females do not use different cues for voice quality distinctions in Gujarati, unlike the case in languages such as Santa Ana del Valle Zapotec (Esposito 2010b).

Table 2 Gujarati and White Hmong wordlist. Gujarati words are written in the Gujarati alphasyllabary and White Hmong words are written in the Hmong Romanized Popular Alphabet. Words from both languages are transcribed in IPA under the orthographic representation, and glossed into English below the IPA transcription.

Gujarati			White Hmong		
Breathy V CV̤	Post-aspirated V CʱV	Modal CV	Breathy V CV̤	Post-aspirated V CʱV	Modal CV
બહાર ba̤ɾ 'outside'	ભાર bʱaɾ 'burden'	બાર baɾ 'twelve'	dag da̤ 42 'lie; fool'	dhas dʱa 22 'separate'	daj da 52 'yellow'
બહાનું ba̤nũ 'excuse'	ભાન bʱan 'consciousness'	બાણ baɳ 'arrow'	dig di̤ 42 'probe; dig with a stick'	dhis dhus dʱi 22 dʱu 22 'the bubbling sound of boiling food'	
ડહોળું ɖɔ̤ɭũ 'polluted'	ઢોળવું ɖʱoɭʋũ 'to spill'	ડોળો ɖoɭo 'eyeball'	dog dig dɔ̤ 42 di̤ 42 'average'	dhos dʱɔ 22 'fits together'	dos dɔ 22 'onion'

3.1.2 Speech materials

Both the Gujarati and White Hmong data sets consisted of three types of words, categorized by their target consonant–vowel (CV) sequence: (i) a voiced unaspirated consonant followed by a breathy vowel (i.e. [CV̤], 'Breathy V'); (ii) a breathy-voiced aspirated consonant followed by a modal vowel (i.e. [CʱV], 'Breathy-aspirated C'); or (iii) a voiced unaspirated consonant followed by a modal vowel (i.e. [CV], 'Modal'). For the sake of convenience, we use the term 'post-aspirated vowel' as equivalent to 'modal vowel following a breathy-voiced aspirated consonant' henceforth. The wordlist for both languages is presented in Table 2.

3.1.2.1 *Gujarati*

Gujarati words were elicited in the following method. First, the investigator revealed a flashcard displaying the target word written in Gujarati orthography (with an English translation below) for no more than two seconds. The speaker then had to create a sentence immediately beginning with the word. The recording was then started, and the speaker produced the sentence as many times as possible within a fixed ten-second window. To familiarize this method to the speakers, a flashcard displaying કુત્રો 'dog' ([kut̪ɾo]) was provided, after which, the investigator (acting as a subject) would create the sentence [kut̪ɾo bʱagi gəjo] 'The dog ran away', as an example, and produce it as many times as possible in ten seconds as an illustration of the task. Later, measurements were taken (as explained below) of all repetitions of these target words, and these measurements were then averaged across repetitions of each word before proceeding with the statistical analysis.[7]

By asking the subjects to produce the words in a sentence of their own creation and by keeping the orthographic representation hidden for the duration of the recording, it was possible to minimize the effects of spelling pronunciation (i.e. disyllabic pronunciation) associated with reading tasks in Gujarati, while the increased speech rate and use of

[7] The spectral measurements of a total of 14 of 438 Gujarati tokens were removed from the averaging, due to machine mistracking of F1 (which was necessary to correct the spectral measures for the effects of formant frequencies and bandwidths), ultimately leading to the removal of [bʱaɾ] 'burden' from Speaker 7's spectral data; the EGG and CPP data for those tokens were unaffected by the mistracking and thus included in later analyses. The EGG data for Speaker 7 was corrupted during the recording of [ɖʱoɭʋũ] 'to spill', and thus CQ and PIC were not measured for that word. Lastly, the word [ɖɔ̤ɭũ] 'polluted' was not recorded at all from Speaker 8 due to unfamiliarity with the term.

familiar vocabulary items helped discourage the use of formal register (i.e. breathy–modal neutralization). Furthermore, of the four sources of breathiness in Gujarati (i.e. [əɦV], [Vɦə], [#ɦ], [VCʱ]), the last three are largely restricted to very casual, lenited speech, inappropriate for a laboratory setting; thus, all target words come from the more stable [əɦV] source.

3.1.2.2 *White Hmong*

All White Hmong words were uttered in the frame *Rov hais* ____ *dua* [ʈɔ24 hai22 ____duə33] ‘Say ____ again’. The onset consonants of the target words were limited to alveolars as that is the only place of articulation with non-prenasalized breathy-voiced aspirated consonants in White Hmong. Six of the words were monosyllabic; the other two words were disyllabic, in which only the first syllable was examined.

Because breathy vowels only occur on the falling tone (42), and breathy-voiced aspirated consonants cannot cooccur with the breathy-falling tone (42) or the high-falling tone (52), the tones examined in the current study include the high-falling tone (52) for modal consonant–vowel sequences, the low tone (22) for breathy-voiced aspirated consonants followed by modal vowels, and the falling tone (42) for the unaspirated consonants followed by breathy vowels.

3.1.3 Measurement

For both languages, simultaneous audio and electroglottographic recordings were made using a Glottal Enterprises two-channel electroglottograph and a head-mounted microphone. Acoustic and electroglottographic measurements were taken automatically using VoiceSauce (Iseli et al. 2007, Shue, Keating & Vicenik 2009) and EGGWorks (Tehrani 2009), respectively. The acoustic and electroglottographic parameters along which the data were measured were chosen based on their reported success in distinguishing modal and breathy vowels in the most recent studies of Gujarati (Keating et al. 2010; Khan 2010b, 2012) and White Hmong (Esposito, in press, Keating et al. 2010).[8] The three acoustic parameters included H1*-H2*, H1*-A3*, and cepstral peak prominence (CPP) as defined in Hillenbrand et al. (1994). Both spectral measures (i.e. H1*-H2* and H1*-A3*) were corrected for surrounding formant frequencies and bandwidths (Hanson 1995) using the Iseli et al. (2007) method. The two electroglottographic parameters included CQ – measured using the hybrid method with a 25% threshold as explained above – and DECPA, defined as the peak positive value for each glottal pulse in the first derivative of the electroglottographic signal.

Measurements were made by dividing each vowel into nine parts with equal duration and then averaging the value for a given measure within each part. Only the first five parts (essentially, the beginning and the middle of the vowel) were examined as we reasoned that the effects of breathy-voiced aspirated consonants would be localized to the beginning and, to a lesser extent, the middle of the vowel.[9] Because they are defined as the first five-ninths of the vowel’s duration, these five timepoints are not of equal duration across tokens. They are in effect normalized for overall vowel duration.

---

[8] Additional acoustic and electroglottographic measures were automatically taken by VoiceSauce and analyzed for statistical significance. Measures such as first formant quality (Q1), calculated as the first formant frequency divided by its bandwidth (see Pennington 2005), were calculated and found to not statistically distinguish vowel types in either language; thus, their results are not reported here.

[9] Given that the consonants we examined are phonetically oral stops and thus do not have meaningful spectral properties, and given that this is an acoustic and electroglottographic study rather than an imaging study (e.g. photo-electroglottography), we did not investigate the phonetic properties of the consonants themselves, focusing instead on their release and the following vowel. See Esposito et al. (2007) for an investigation of the acoustic, aerodynamic, and electroglottographic properties of breathy oral stops and breathy nasal stops in three related Indic languages: Marathi, Hindi, and Bengali.

Table 3 Measures that distinguish vowels after breathy-voiced aspirated consonants from either breathy vowels or modal vowels at five timepoints (T1–T5) in Gujarati. All measures listed showed a statistically significant difference ($p \leq .001$) between the two categories in question.

	Measures that distinguish vowels after breathy-voiced aspirated consonants from timepoints				
	1	2	3	4	5
Breathy vowels		CPP H1*-H2* H1*-A3*	CPP H1*-H2*		
Modal vowels	H1*-H2* H1*-A3* CQ	CPP H1*-H2* H1*-A3* CQ	CPP H1*-H2* H1*-A3* CQ	CPP H1*-H2* H1*-A3* CQ	CPP H1*-H2* H1*-A3* CQ

3.2 Results

Separate ANOVAs and post-hoc pair-wise comparisons for each measure at each timepoint were used to determine if there was a significant ($p \leq .001$) difference between the vowel types (i.e. breathy, post-aspirated, and modal). The current study is particularly concerned with comparing post-aspirated vowels to breathy and modal vowels. For a detailed analysis comparing breathy vowels to modal vowels, see Khan (2010b, 2012) for Gujarati and Esposito (in press) for White Hmong.

3.2.1 Gujarati

The results of the acoustic and EGG measures across five timepoints for Gujarati are presented in Table 3. Only measures that show a significant difference between the phonation categories (i.e. post-aspirated vowels compared to either breathy or modal vowels) are given. Graphs of the average values for statistically successful measures ($p \leq .001$) across five timepoints are presented in Figure 1. DECPA is not included in the graphs for Gujarati, as this measure was not statistically successful in distinguishing any two of the three categories.

3.2.1.1 *Results of acoustic measures*

Spectral results indicate that breathy, modal, and post-aspirated vowels are not significantly different at the consonant release, but the three soon separate into distinct categories in the next few timepoints; by the midpoint of the vowel, however, breathy and post-aspirated vowels become indistinguishable in their spectral properties, while remaining distinct from modal vowels. Specifically, H1*-H2* is significantly higher in post-aspirated vowels than in breathy vowels at timepoints 2 and 3; the same measure also distinguishes both categories from modal vowels, which have the lowest H1*-H2* values in the set. By timepoint 4, however, breathy and post-aspirated vowels are no longer statistically distinct from each other, while they both remain significantly higher than modal vowels. In terms of overall changes in mean values, modal vowels exhibit a low H1*-H2* across all five timepoints, while post-aspirated vowels show a dynamic shape, reaching their highest H1*-H2* values at timepoint 2; breathy vowels are similar to post-aspirated vowels in terms of this dynamic shape, although with a delayed peak, exhibiting their highest H1*-H2* values at timepoints 4 and 5.

Like H1*-H2*, spectral tilt as measured by H1*-A3* is significantly higher (i.e. more steeply falling) in post-aspirated vowels than in breathy vowels at timepoints 2 and 3, while both categories exhibit a significantly steeper tilt than modal vowels across timepoints 3, 4, and 5. Like the overall changes in means of H1*-H2* values, modal vowels maintain a low, flat H1*-A3* value, while post-aspirated vowels reach their highest value at timepoints 2 and 3 and breathy vowels show a similar but delayed peak at timepoint 4.

As a measure of the strength of the signal over noise across the spectrum, CPP is expected to be lower in breathier phonations. Indeed, CPP measures are generally lower for

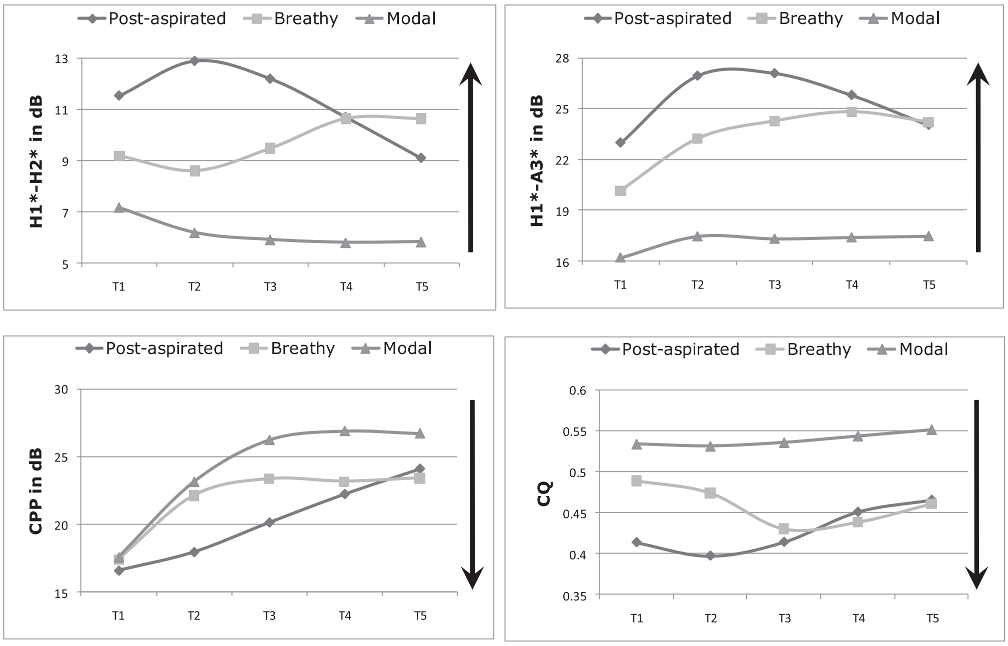

Figure 1 Graphs of the average H1*-H2* (dB), H1*-A3* (dB), CPP (dB), and CQ (unspecified units) values for vowels after breathy-voiced aspirated consonants (labeled 'post-aspirated' in the graphs above), breathy vowels, and modal vowels at each of the first five of nine timepoints (T1–T5) in Gujarati. For each graph, the arrow points in the direction of increased breathiness.

post-aspirated vowels and higher for modal vowels, with an intermediate value for breathy vowels. However, the specifics are slightly different from the patterns seen in the measures of spectral tilt and spectral balance. All three categories show a rise in CPP across all five timepoints, both starting and ending with statistically non-distinct values at timepoints 1 and 5; however, the sharpness of their rises across the intermediate timepoints is different. At timepoint 2, both breathy and modal vowels sharply rise in CPP and are not distinguished from one another, while post-aspirated vowels show a shallower rise and are thus significantly lower in CPP than the other two categories. By timepoint 3, however, all three categories separate and are statistically distinct from one another, with post-aspirated vowels having the lowest CPP value (i.e. noisiest and/or least periodic) and modal vowels having the highest. The pattern is then reversed at timepoint 4, with breathy and post-aspirated vowels not distinguished but modal vowels continuing to exhibit a significantly higher value.

3.2.1.2 *Results of electroglottographic measures*

EGG data also reveal distinctions in the production of modal, breathy, and post-aspirated vowels in Gujarati. Post-aspirated vowels have a lower CQ than modal vowels at all timepoints, but are not statistically distinguished from breathy vowels at any timepoint. Post-aspirated vowels exhibit their lowest (i.e. breathiest) CQ value at timepoint 2. At timepoint 3, breathy vowels reach their lowest CQ value, essentially merging with the post-aspirated category. After timepoint 3, both categories gradually become less breathy but still significantly distinct from modal vowels, with the average CQ values of the breathy and post-aspirated vowels appearing indistinguishable from one another.

DECPA does not distinguish any two of the three categories in Gujarati following our use of $p \leq .001$ as a benchmark for statistical significance; however, post-aspirated and

Table 4 Measures that distinguish vowels after breathy-voiced aspirated consonants from either breathy vowels or modal vowels at five timepoints (T1–T5) in White Hmong. All measures listed showed a statistically significant difference ($p \leq .001$) between the two categories in question.

	Measures that distinguish vowels after breathy-voiced aspirated consonants from timepoints				
	1	2	3	4	5
Breathy vowels	CPP	CPP	CPP	CPP	CPP
	H1*-H2*				H1*-H2*
	CQ	CQ	CQ	CQ	CQ
	DECPA	DECPA	DECPA	DECPA	DECPA
Modal vowels	H1*-H2*	H1*-H2*			
	CQ	CQ			
	DECPA	DECPA	DECPA		

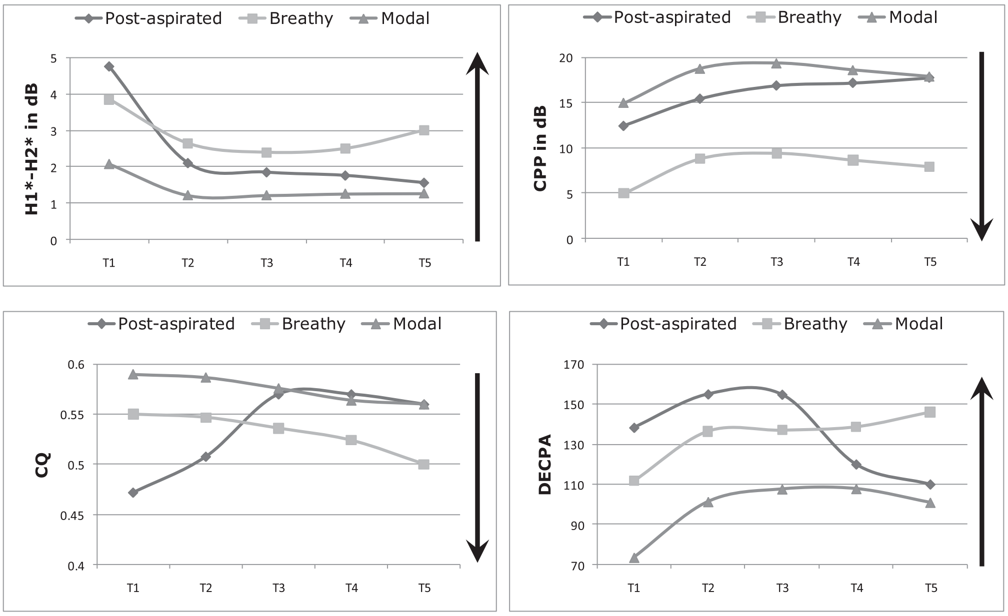

Figure 2 Graphs of the average H1*-H2* (dB), CPP (dB), CQ (unspecified units), and DECPA (unspecified units) values for vowels after aspirated consonants (labeled 'post-aspirated' in the graphs above), breathy vowels, and modal vowels in each of the first five of nine timepoints (T1–T5) in White Hmong. For each graph, the arrow points in the direction of increased breathiness.

modal vowels are distinguished in DECPA in the first timepoint, with post-aspirated vowels exhibiting the higher value, following a lower standard for significance ($p = .005$).

3.2.2 White Hmong

The results of the acoustic and EGG measures across five timepoints for White Hmong are presented in Table 4. Only measures that show a significant difference between the phonation categories are given. Individual graphs of the average values for CPP, H1*-H2*, CQ, and DECPA across five timepoints are presented in Figure 2. The results of H1*-A3* are not shown in the graphs because this measure does not significantly distinguish any two of the three categories in White Hmong.

3.2.2.1 *Results of acoustic measures*

On the measure H1*-H2*, vowels after breathy-voiced aspirated consonants were breathier than phonemically breathy vowels at the first timepoint. (That is, they had a significantly higher H1*-H2* value.) By point 2, however, the H1*-H2* value for the post-aspirated vowels is no longer significantly higher than that of breathy vowels. In fact, on points 2, 3, and 4, there is no significant difference between the post-aspirated and breathy vowels. However, by point 5, the average H1*-H2* value for the breathy vowels increases, while the value for the post-aspirated vowels decreases. These two vowel types are significantly different at this timepoint, with the post-aspirated vowels having a modal-like H1*-H2* value. The post-aspirated vowels and modal vowels are significantly different on the first two timepoints, when the post-aspirated vowels have a higher, breathy-like, H1*-H2* value. By point 3, when the H1*-H2* value drops for the post-aspirated vowels, this vowel type is no longer significantly different from modal vowels.

Like modal vowels, post-aspirated vowels have a significantly higher CPP (i.e. they are less noisy and/or more periodic) than breathy vowels throughout the five timepoints. In fact, there was no significant difference between the CPP values for modal vowels and post-aspirated vowels at any of the timepoints. Thus, along the CPP dimension, there are essentially two categories in White Hmong: breathy vowels and modal vowels, the latter category also including vowels preceded by breathy-voiced aspirated [dʱ].

3.2.2.2 *Results of electroglottographic measures*

Post-aspirated vowels have a significantly different CQ value from breathy vowels on all five timepoints and are significantly different from modal vowels on the first two timepoints. On points 1 and 2, the post-aspirated vowels have even less vocal fold contact than the breathy vowels. But, by timepoint 3, the CQ value of the post-aspirated vowels increases such that it is significantly higher than breathy phonation and no longer significantly different from modal phonation.

For DECPA, the post-aspirated vowels are significantly breathier than either breathy or modal phonation until point 4.[10] During points 4 and 5, the DECPA value for post-aspirated vowels drops and becomes more modal-like, becoming significantly lower than that of breathy vowels, but not significantly different from modal phonation.

4 Discussion

In the current study, we hypothesized that the difference between post-aspirated vowels – phonemically modal vowels following breathy-voiced aspirated stops ([ʱV]) – and phonemically breathy vowels ([V̤]) would be manifested in the timing and/or degree of breathiness. Both timing and degree distinctions were found in both languages. Language-specific differences in the overall resemblance of post-aspirated vowels to other categories were also found, as well as differences in the reliability of specific acoustic and electroglottographic measures to distinguish categories. This section expands on these cross-linguistic comparisons, and explores possible explanations for the cross-linguistic differences.

We hypothesized that the breathiness associated with breathy-voiced aspirated consonants would be localized to the consonant release and thus produced at the onset of the following

---

[10] Contrary to expectation, the results for DECPA are higher for breathier phonation than they are for modal phonation. A similar trend is reported for breathy phonation in White Hmong (Esposito 2010c) and lax phonation in Yi (Kuang 2010, 2011). It is hypothesized that the vocal folds must move more quickly during breathy phonation due to their greater glottal aperture. This is in contrast with creaky phonation, which is produced by vocal folds that are close together, and therefore do not need to move as quickly to reach a state of glottal closure.

vowel, while the breathiness associated with breathy vowels would be produced across a larger portion of the vowel, with language-specific distinctions in the exact localization. Results confirmed this hypothesis. In both languages, there was a brief, early realization of breathy phonation (as indicated along multiple acoustic and EGG dimensions) after the breathy-voiced aspirated consonant; post-aspirated vowels generally began very breathy, but became more modal at the vowel midpoint, reflecting the fact that the breathiness is phonologically associated with the consonant and not with the vowel. The localization of breathy phonation in the breathy vowel, however, was more language-specific. In White Hmong, breathy vowels are uniformly breathy across the first half of their duration, while in Gujarati, breathy vowels start out with a more modal-like phonation, but become breathier by the midpoint. This dynamic realization of breathy vowels in Gujarati may be due to the historical source of vocalic breathiness in that language. Most breathy vowels in Gujarati derive from disyllabic sequences of vowels with intervocalic [ɦ]; thus, although truly disyllabic productions (e.g. [VɦV]) were not found in the current study (presumably due to the precautionary measures taken in the experimental setup), monophthongal breathy vowels in Gujarati may still exhibit the strongest breathiness near the midpoint, while still remaining breathier than modal vowels throughout the duration.

In addition to the difference in timing, we also hypothesized that post-aspirated vowels would show a different degree of breathiness than breathy vowels; however, the direction of magnitude was not inherently obvious from previous research. The results of our acoustic and electroglottographic analyses confirm that the two categories are distinguished by degree in both languages, clearly demonstrating that post-aspirated vowels begin with even greater breathiness than breathy vowels across various measures; in effect, the beginning of the post-aspirated vowel is breathier than that of a breathy vowel. This greater magnitude of breathiness in post-aspirated vowels is likely related to its short, early realization; given its association to the preceding consonant rather than to the vowel itself, the breathiness from breathy-voiced aspirated stops must be produced in a limited duration, and may thus require compensatory amplification to be reliably perceived by the listener.

The similarity of post-aspirated vowels to other categories was found to be language-specific. Post-aspirated vowels in both languages begin breathier and become more modal towards the vowel midpoint, but their overall resemblance to breathy or modal vowels differs. Along the various measures across the five timepoints, post-aspirated vowels were statistically more similar to breathy vowels in Gujarati, but statistically more similar to modal vowels in White Hmong. The strong similarity between modal vowels and post-aspirated vowels in White Hmong may be one of the reasons why Jarkey (1987) believed the aspiration associated with [dʱ] to be optional; the aspiration produced following the release of [dʱ] may simply be less salient in White Hmong due to its limited duration, typically not lasting beyond the first timepoint. Post-aspirated vowels in Gujarati, on the other hand, maintain their breathiness through the second or third timepoint (depending on the measure), and thus presumably have a more salient period of breathiness, more closely resembling breathy vowels.

In the absence of photographic data that would have been obtained through laryngoscopic or other invasive means in the current study, our findings on the articulation of breathy phonation in Gujarati and White Hmong breathy vowels and breathy-voiced aspirated stops are based on electroglottographic data as well as established correlations between acoustic outputs and source characteristics. For example, claims that breathy phonation involves a more open glottis are supported by the lower CQ values and higher H1*-H2* values in both breathy vowels and post-aspirated vowels across the two languages. Claims that breathy phonation involves a less abrupt glottal closure, however, require further investigation. The two measures related to the nature of the glottal closure are DECPA and H1*-A3*, which are expanded on below.

One may assume that DECPA, the peak positive value in the first derivative of the EGG signal, should be higher when the vocal folds approximate one another (i.e. in the closure phase) more quickly, but surprisingly, it was found to be higher in both types of breathy sequences in White Hmong. (The measure was not reliable in distinguishing phonation categories in Gujarati.) A higher DECPA was also found for lax phonation (a phonation similar to breathy) in Yi (Keating et al. 2010; Kuang 2010, 2012) and in an additional study on White Hmong (Esposito, in press). A visual inspection of the EGG signal for the White Hmong data collected for the current study confirms that breathy phonation is characterized by a longer open quotient and shorter closing quotient (although it is presumed the glottis may not be completely closed in breathy phonation), with a very sharp transition between the two. Keating et al. (2010: 93) suggests that the higher DECPA values for breathy/lax phonation is due to a principle of 'the further, the faster'; the greater degree of glottal opening in breathy phonation might require the vocal folds to move more quickly to return to a (semi-)closed state. It may be that this shortening of the transition to the closure phrase makes it possible for White Hmong speakers to elongate their open phase without elongating the entire glottal pulse, which would in turn significantly lower the pitch in this lexical tone language. In Gujarati, on the other hand, the longer open phase presumably comes about through a longer overall glottal pulse, generating the lower f0 long established as a property of Gujarati breathiness (Pandit 1957, Dave 1967, Fischer-Jørgensen 1967) while not significantly affecting the DECPA.

Because of the strengthening of the fundamental and the weakening of higher harmonics due to the less-abruptly closed glottis in breathy voice, H1*-A3* is often cited as an acoustic correlate of the abruptness of glottal closure (Stevens 1977). Indeed, this measure successfully distinguishes both types of breathy phonation from modal phonation in Gujarati in the expected direction (although this was not the case in Fischer-Jørgensen's 1967 study using uncorrected spectral measures), but it is not a useful measure in White Hmong. It is unclear why H1*-A3* is unsuccessful in White Hmong, although one possibility would be that the higher frequency aperiodic noise generated in White Hmong breathy phonation is loud enough to boost the value of A3* to a level not significantly distinct from that of modal phonation. Indeed, the authors' intuitions would characterize the breathy vowels of White Hmong as far noisier than those of Gujarati.

5 Conclusions and directions for further study

Despite their geographical and genetic distance, Gujarati and White Hmong have both independently generated a cross-linguistically unusual contrast between breathy and modal phonation in both voiced stop consonants ([C] vs. [Cʱ]) and vowels ([V] vs. [V̤]), a distinction so rare that it is not even shared by closely related languages such as Hindi or Green Mong. Both Gujarati and White Hmong have derived CV sequences in which phonetic breathiness can be associated phonologically to the consonant ([CʱV]) or to the vowel ([CV̤]) – but not to both (*[CʱV̤]), due to similar phonotactic restrictions in both languages. Both acoustically and articulatorily, these two types of sequences (i.e. [CʱV] and [CV̤]) are distinguishable within the first half of the vowel: in both languages, breathiness associated with stops is characterized by a short period of extreme breathiness concentrated at the onset of the consonant into the vowel, while breathiness associated with vowels is characterized by a less extreme production of breathiness, spread more evenly across the first half of the vowel (with some dynamic behavior in Gujarati). Naturally, the phonetic details are more language-specific, with some acoustic and electroglottographic measures being better indicators of breathiness in one language over the other (e.g. H1*-A3* in Gujarati, DECPA in White Hmong), and with vowels following breathy-voiced aspirated consonants ([ʱV]) more closely resembling phonemically breathy vowels ([V̤]) in Gujarati while they more closely resemble phonemically modal vowels ([V]) in White Hmong.

The results of the present study show that timing and degree of breathiness are reliable in distinguishing breathy vowels from post-aspirated vowels. The question arises: If presented with breathiness in a CV sequence, can listeners rely on either timing or degree of breathiness alone to determine the segment to which the breathiness is phonologically associated (i.e. [CʱV]) or [CV̤]), if they perceive the breathiness at all? There are secondary cues to distinguishing these segments in both languages. In White Hmong, all breathy vowels bear the falling tone 42, so f0 could play a vital role in the perception of breathiness, while in Gujarati, duration could play an important role in distinguishing these segments in that most breathy vowels in that language derive from disyllables and can be produced as such in certain registers, while post-aspirated vowels are not derived from such sequences. A follow-up to the current study would be a perception experiment where speakers are asked to identify and/or discriminate between [CʱV] and [CV̤] sequences (and possibly modal [CV] as another option). This future extension would allow us to determine how both cross-linguistic and language-specific cues assist native speakers of these two typologically rare languages in perceiving breathy voice and determining its segmental association in ambiguous contexts.

Acknowledgments

This study is supported by NSF grant BCS-0720304 for a project entitled ‘Production and Perception of Linguistic Voice Quality’ (Principal Investigators: Patricia A. Keating, Abeer Alwan, Jody Kreiman, and Christina M. Esposito; 2007–2012), exploring aspects of phonation in various languages. The authors would like to thank (i) our consultants for their participation, (ii) Victoria Thatte, Ramesh Goud, and Ranjan Goud for their input, (iii) the staff and clients at the Hmong American Partnership, the members of the Immanuel Hmong Lutheran Church, Eric Bestrom, Claire Moore-Cantwell, Alice Ware, and Oleh Zaychenko for their assistance, and (iv) John H. Esling, Adrian P. Simpson, Ewa Jaworska, and three anonymous reviewers for their valuable comments and suggestions.

References

Abramson, Arthur S. & Theraphan Luangthongkum. 2001. Phonation types in Suai. *Journal of the Acoustical Society of America* 110(5), 2763.

Andruski, Jean & Martha Ratliff. 2000. Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association* 30(1/2), 37–61.

Baken, Ronald J. & Robert F. Orlikoff. 2000. *Clinical measurement of speech and voice*. San Diego, CA: Singular Thomson Learning.

Bickley, Corine. 1982. Acoustic analysis and perception of breathy vowels. *Speech Communication Group Working Papers*, *Research Laboratory of Electronics*, 73–93. Boston, MA: MIT.

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

Cardona, George & Babu Suthar. 2003. Gujarati. In George Cardona & Dhanesh Jain (eds.), *The Indo-Aryan languages*, 659–697. Routledge.

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

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

Dave, Radhekant. 1967. A formant analysis of the clear, nasalized, and murmured vowels in Gujarati. *Indian Linguistics* 28, 1–30.

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

Dixit, R. Prakash. 1989. Glottal gestures in Hindi plosives. *Journal of Phonetics* 17, 213–237.

Esling, John H. & Jimmy Harris. 2005. States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In William J. Hardcastle & Janet Beck (eds.), *A figure of speech: A Festschrift for John Laver*, 347–383. Mahwah, NJ: Lawrence Erlbaum.

Esposito, Christina M. 2005. An acoustic and electroglottographic study of phonation in Santa Ana del Valle Zapotec. Poster presented at the 79th meeting of the Linguistic Society of America.

Esposito, Christina M. 2010a. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38(2), 303–316.
Esposito, Christina M. 2010b. Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40(2), 181–198.
Esposito, Christina M. (In press). An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics*, doi.org/10.1016/j.wocn.2012.02.007. Published by Elsevier, 29 March 2012.
Esposito, Christina M., Sameer ud Dowla Khan & Alex Hurst. 2007. Breathy Nasals and /Nh/ Clusters in Bengali, Hindi, and Marathi. *Indian Linguistics* 68(3–4), 275–299. Pune: Linguistic Society of India.
Firth, John Rupert. 1957. Phonetic observations on Gujarati. *Bulletin of the School of Oriental and African Studies* 20(1), 231–241.
Fischer-Jørgensen, Eli. 1967. Phonetic analysis of breathy (murmured) vowels. *Indian Linguistics* 28, 71–139.
Fulop, Sean & Chris Golston. 2008. Breathy and whispery voicing in White Hmong. *Proceedings of Meetings on Acoustics* 4, 1–10.
Garellek, Marc & Patricia A. Keating. 2010. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *UCLA Working Papers in Phonetics* 108, 143–163.
Golston, Chris & Phong Yang. 2001. Hmong loanword phonology. In Caroline Féry (ed.), *The Holland Institute for Generative Linguistics Phonology Conference V*, 40–57. Potsdam: University of Potsdam.
Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.
Guion, Susan G., Mark W. Post & Doris L. Payne. 2004. Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics* 32, 517–542.
Hanson, Helen M. 1995. *Glottal characteristics of female speakers*. Ph.D. dissertation, Harvard University.
Henrich, Nathalie, Christophe D'Alessandro, Boris Doval & Michèle Castellengo. 2004. On the use of the derivative of electroglottographic signals for characterization of non-pathological phonation. *Journal of the Acoustical Society of America* 115(3), 1321–1332.
Herbst, Christian & Sten Ternström. 2006. A comparison of different methods to measure the EGG contact quotient, *Logopedics Phoniatrics Vocology* 31, 126–138.
Hillenbrand, James M., Ronald A. Cleveland & Robert L. Erickson. 1994. Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research* 37, 769–778.
Holmberg, Eva, Roger Hillman, Joseph Perkell, Peter Guiod & Susan Goldman. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research* 38, 1212–1223.
Howard, David M. 1995. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice* 9, 163–172.
Iseli, Markus, Yen-Liang Shue & Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America* 121, 2283–2295.
Jarkey, Nerida. 1987. An investigation of two alveolar stop consonants in White Hmong. *Linguistics of the Tibeto-Burman Area* 10, 57–70.
Jones, Ted E. & Lyle M. Knudson. 1977. Guelavia Zapotec phonemes. In William Merrifield (ed.), *Studies in OtoManguean phonology*, 163–80. Dallas, TX: SIL/Univeristy of Texas, Arlington.
Kagaya, Ryohei & Hajime Hiroshi. 1975. Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants. *Annual Bulletin Research Institute of Logopedics and Phoniatrics* 9, 27–45.
Keating, Patricia A., Christina M. Esposito, Marc Garellek, Sameer ud Dowla Khan & Jianjing Kuang. 2010. *Phonation contrasts across languages* (UCLA Working Papers in Phonetics 108). Los Angeles, CA: University of California, Lost Angeles.
Khan, Sameer ud Dowla. 2010a. Bengali (Bangladeshi Standard). *Journal of the International Phonetic Association* 40(2), 221–225.
Khan, Sameer ud Dowla. 2010b. Breathy phonation in Gujarati: An acoustic and electroglottographic study. Poster presentation, Acoustical Society of America 159, Baltimore. [Abstract: *Journal of the Acoustical Society of America* 127(3), 2021.]

Khan, Sameer ud Dowla. 2012. The phonetics of contrastive phonation in Gujarati. Ms., Brown University.

Kirk, Paul L., Jenny Ladefoged & Peter Ladefoged. 1993. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In Anthony Mattina & Timothy Montler (eds.), *American Indian linguistics and ethnography in honor of Laurence C. Thompson*, 435–450. Missoula, MT: Universtiy of Montana Press.

Klatt, Dennis H. & Laura C. Klatt. 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.

Kreiman, Jody, Bruce R. Gerratt & Norma Antoñanzas-Barroso. 2006. Analysis and synthesis of pathological voice quality. http://www.surgery.medsch.ucla.edu/glottalaffairs/files/GASoftwareManual2006.pdf (accessed 1 February 2011).

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

Kreiman, Jody, Bruce R. Gerratt & Sameer ud Dowla Khan. 2010. Effects of native language on perception of voice quality. *Journal of Phonetics* 38(4), 588–593.

Kuang, Jianjing. 2010. An acoustic and electroglottographic study of phonation contrast in Yi. Poster presentation, Acoustical Society of America 159, Baltimore. [Abstract: *Journal of the Acoustical Society of America* 127(3), 2022.]

Kuang, Jianjing. 2011. Production and Perception of the Phonation Contrast in Yi. Ms., University of California, Los Angeles.

Ladefoged, Peter. 1971. *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter. 1983. The linguistic use of different phonation types. In Diane Bless & James Abbs (eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 351–360. San Diego, CA: College-Hill Press.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Malden, MA: Blackwell.

Laver, John. 1981. *Phonetic description of voice quality*. Cambridge: Cambridge University Press.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world*, 16th edn. Dallas, TX: SIL International. http://www.ethnologue.com/ (accessed 1 March 2011).

Masica, Colin P. 1993. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Michaud, Alexis. 2004. A measurement from electroglottography: DECPA, and its application in prosody. *Speech Prosody 2004*, Nara, Japan, 633–636.

Michaud, Alexis & Martine Mazaudon. 2006. Pitch and voice quality characteristics of the lexical word-tones of Tamang, as compared with level tones (Naxi data) and pitch-plus-voice-quality tones (Vietnamese data). *Speech Prosody 2006*, Dresden, Germany, 823–826.

Miller, Amanda L. 2007. Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics* 35, 56–84.

Mistry, Purushottam J. 1997. Gujarati phonology. In Alan S. Kaye (ed.), *Phonologies of Asia and Africa*, vol. 2, 653–674. Winona Lake, IN: Eisenbrauns.

Modi, Bharti. 1987. Rethinking of 'murmur in Gujarati'. *Indian Linguistics* 47, 39–55.

Mortensen, David. 2000. Sinitic loanwords in two Hmong dialects of Southeast Asia. Honors thesis, Utah State University. http://www.pitt.edu/∼drm31/sinitic_loans_in_hmong.pdf (accessed 1 February 2011).

Mortensen, David. 2004. Preliminaries to Mong Leng (Hmong Njua) Phonology. Ms., University of California, Berkeley. http://www.pitt.edu/∼drm31/mong_leng_phonology.pdf (accessed 1 February 2011).

Munro, Pamela & Felipe Lopez, with Olivia Mendez, Rodrigo Garcia & Michael Galant. 1999. *Di'syonaary X:tee'n Diizh Sah Sann Lu'uc (San Lucas Quiaviní Dictionary/Diccionario Zapoteco de San Lucas Quiaviní)*. Los Angeles, CA: Chicano Studies Research Center Publications.

Nair, Usha. 1979. *Gujarati phonetic reader*. Mysore: Central Institute of Indian Languages.

Ohala, Manjari. 1983. *Aspects of Hindi phonology*. Delhi: Motilal Banarsidass.

Orlikoff, Robert F. 1991. Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects. *Journal of Speech and Hearing Research* 34, 1066–1072.

Pandit, Prabodh B. 1957. Nasalization, aspiration and murmur in Gujarati. *Indian Linguistics* 17, 165–172.

Pandit, Prabodh B. 1961. Historical phonology of Gujarati vowels. *Language* 37(1), 54–66.

Pennington, Mark. 2005. *The phonetics and phonology of glottal manner features*. Ph.D. dissertation, Indiana University.

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

Shue, Yen-Liang, Patricia A. Keating & Chad Vicenik. 2009. VoiceSauce: A program for voice analysis. *Journal of the Acoustical Society America* 124(4), 2221. [Abstract]

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

Stevens, Kenneth. 1988. Modes of vocal fold vibration based on a two-section model. In Osamu Fujimura (ed.), *Vocal physiology: Voice production mechanisms and functions*, 357–371. New York: Raven Press.

Stevens, Kenneth & Helen Hanson 1995. Classification of glottal vibration from acoustic measurements. In Osamu Fujimura & Minoru Hirano (eds.), *Vocal fold physiology: Voice quality control*, 147–170. San Diego, CA: Singular Publishing Group.

Tehrani, Henry. 2009. EGGWorks: A program for automated analysis of EGG signals. http://www.linguistics.ucla.edu/faciliti/facilities/physiology/EggWorksSetup.exe (accessed 1 September 2010).

Traill, Anthony. 1985. *Phonetic and phonological studies of !Xóõ Bushman*. Hamburg: Helmut Buske.

Turner, Ralph Lilley. 1921. Gujarati phonology. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 3, 329–365.

Vũ-Ngọc, Tuấn, Christophe d'Alessandro & Alexis Michaud. 2005. Using open quotient for the characterization of Vietnamese glottalized tones. *Eurospeech-Interspeech 2005: 9th European Conference on Speech Communication and Technology*, Lisbon, 2885–2889.

Wackernagel, Jakob. 1896. *Altindische Grammatik, Teil I: Lautlehre*. Göttingen: Vandenhoeck und Ruprecht.

Watkins, Justin. 1999. CQ of laryngeal gestures and settings in Wa. *The 14th International Congress of Phonetics Sciences* (ICPhS 14), San Francisco, August 1999, 1017–1020.

Watkins, Justin. 2002. *The phonetics of Wa: Experimental phonetics, phonology, orthography and sociolinguistics*. Canberra: Pacific Linguistics.

Wayland, Ratree, Scott Gargash & Allard Jongman. 1994. Acoustic and perceptual investigation of breathy voice. *Journal of the Acoustical Society of America* 97, 3364.

Wayland, Ratree & Allard Jongman. 2002. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31(2), 181–201.

Whitney, William Dwight. 1889. *Sanskrit grammar*. Republished 1964, Cambridge: Harvard University Press.

Yadav, Ramawatar. 1984. Voicing and aspiration in Maithili: A fiberoptic and acoustic study. *Indian Linguistics* 45, 1–30.