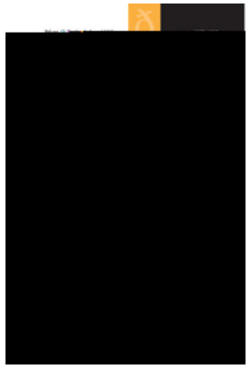


Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong

Christina M. Esposito and Sameer ud Dowla Khan

Department of Linguistics,
Macalester College
esposito@macalester.edu

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

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 f ɮ

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 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).

. f f :

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.

2.3.1 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, 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. [bəd̪ʱũ] ~ [bə̤d̪ṳ̃] ~ [bə̤ðṳ̃] 'whole' (Firth 1957: 235; Pandit 1957: 171; Mistry 1997: 667; Cardona & Suthar 2003: 666).

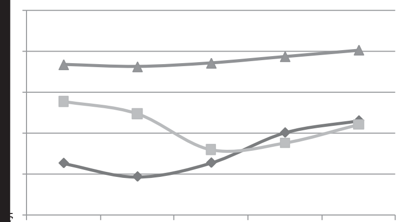
Due to various sociolinguistic pressures, breathy vowels are often not produced as such in particular contexts. Pandit (1957: 170), Dave (1967: 2), Nair (1979: 22), and Cardona & Suthar (2003: 666) report that many speakers have merged the breathy vowels with their corresponding modal vowels in what 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 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 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 successfullApaic

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 [əɦ

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

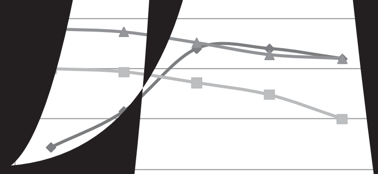
Separate ANOVAs and post-hoc pair-wise comparisons for each measure at each timepoint were used to determine if there was a significant ($p \leq .001$) difference between the vowel



[redacted] at distinguish vowels [redacted] asp [redacted] cons [redacted]
[redacted] e timepoints (T1–T5 [redacted] me [redacted] liste [redacted]
[redacted] between the two ca [redacted]

[redacted] ures that distinguish [redacted] voi [redacted] pira [redacted]
1
CPP
H1*-H2*
CQ
DECPA
H1*-H2*
CQ
DECPA

[redacted]



[redacted]

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 drop9.2716Tm(39.7(w)716Tm(39.2)716lue)op9.271216(v)20.3(o)39.7(w)3335.3(el)0.4(ls) vowel, while the breathiness associated with breathy vowels would be produced across a larger 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 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.

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.

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. Abingdon: 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

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*

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 presented at Acoustical Society of America 159, Baltimore, MD. [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, IL: 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.

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*, vol. 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.