Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

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

The phonetics of contrastive phonation in Gujarati ☆

Sameer ud Dowla Khan *

Department of Linguistics, Reed College, 3203 SE Woodstock Boulevard, Portland, OR 97202-8199, USA

ARTICLE INFO

Article history:
Received 20 December 2011
Received in revised form
3 July 2012
Accepted 9 July 2012

ABSTRACT

The current study examines (near-)minimal pairs of breathy and modal phonation produced by ten native speakers of Gujarati in connected speech, across different vowel qualities and separated by nine equal timepoints of vowel duration. The results identify five spectral measures (i.e. H1–H2, H2–H4, H1–A1, H1–A2, H1–A3), four noise measures (i.e. cepstral peak prominence and three measures of harmonics-to-noise ratio), and one electroglottographic measure (i.e. CQ) as reliable indicators of breathy phonation, revealing a considerably larger inventory of cues to breathy phonation than what had previously been reported for the language. Furthermore, while the spectral measures are consistently distinct for breathy and modal vowels when averaging across timepoints, the efficacy of the four noise measures in distinguishing phonation categories is localized to the midpoint of the vowel's duration. This indicates that the magnitude of breathiness, especially in terms of aperiodicity, changes as a function of time. The current study supports that breathy voice in Gujarati is a dynamic, multidimensional feature, surfacing through multiple acoustic cues that are potentially relevant to the listener.

1. Introduction

Due to the highly gradient nature of voice quality, and the masking effects of prosody and other components of speech, the acoustic properties associated with linguistic phonation contrasts are extremely complex, and vary considerably both within and across languages. One language whose phonation contrast has been of much interest is Gujarati, which distinguishes modal ("clear") and breathy ("murmured") vowels (e.g. [baɾ] 'twelve' vs. [ba̤ɾ] 'outside'). While studies of Gujarati agree that these phonations are phonologically contrastive and that speakers can reliably perceive the distinction, a comprehensive study of its acoustic and articulatory properties is long overdue. Certainly, early acoustic studies of Gujarati breathiness are well-known (e.g. Bickley, 1982; Dave, 1967; Fischer-Jørgensen, 1967; Pandit, 1957, among others; see Section 2.3 for a review), yet many important questions remain. What acoustic and electroglottographic measures reliably distinguish phonation types in Gujarati? Are certain acoustic and electroglottographic differences between phonation types limited to particular parts of the vowel? How does voice quality (i.e. phonation) interact with vowel quality (i.e. height, backness)? And what can acoustic and electroglottographic investigation tell us about the articulatory processes behind this contrast?

The goal of the current study is to answer these questions by examining acoustic and electroglottographic (EGG) data collected from ten Gujarati speakers in a semi-naturalistic laboratory setting. The paper is divided as follows: background and motivation for the current study are provided in Section 2, the methods of data collection and analysis are provided in Section 3, the results of which are provided in Section 4, and lastly a discussion of the results is provided in Section 5.

2. Background and motivation for current study

2.1. General characteristics of breathy voice

Visual observation of the pharynx and glottis by Esling and Harris (2005) and Edmondson and Esling (2006) has identified six "valves" that can be manipulated to adjust voice quality: vocal fold adduction and abduction, ventricular incursion, sphincteric compression of the arytenoids, epiglotto-pharyngeal constriction, laryngeal raising, and pharyngeal narrowing. They describe the configuration for canonical breathy voice as involving vibrating vocal folds and partially adducted arytenoids to allow a posterior opening in the glottis, evidently collapsing the categories of breathy voice (murmur) and whispery voice of earlier studies (Catford, 1977; Fant, Liljencrants, & Lin, 1985; Gordon & Ladefoged, 2001; Laver, 1980; Pennington, 2005). Similarly, in the current study, the labels "breathy voice",

☆ *Data collection and funding support from:* Department of Linguistics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095, USA.
* Tel.: +1 7143627261.
E-mail address: sameeruddowlakhan@gmail.com

0095-4470/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2012.07.001

"whispery voice", and "murmur" are simply collapsed as "breathy voice" to avoid a discussion of nomenclature. Indeed, Laver (1980) describes "breathy voice" (Ladefoged's "murmur") as "the range of qualities produced with a low degree of laryngeal effort, and where only a slight amount of glottal friction is audible", and "whispery voice" as "phonations produced with a greater degree of laryngeal effort, and where a more substantial amount of glottal friction…is audible" (p. 134), suggesting that the two can loosely be considered weaker and stronger variants (respectively) of one voice quality.

This general configuration produces more rounded glottal pulses and continuous airflow, involving a larger open quotient (the more "open" proportion of each glottal cycle) and less abrupt glottal closure, while the uninterrupted airflow past the vibrating vocal folds generates aperiodic noise. Acoustically, this produces a fundamental frequency (f0) with high amplitude (H1) and an increase in aperiodic noise at higher harmonics; thus, most acoustic studies of breathy voice focus on measures of spectral amplitude (Section 2.2.1) and measures of periodicity (Section 2.2.2), although additional acoustic measures (Section 2.2.3) are also found. However, due to the multidimensional nature of the articulatory configurations that can produce breathy voice, it is no surprise that its acoustic output is also multidimensional and highly variable within and across languages (Keating, Esposito, Garellek, Khan, & Kuang, 2010).

2.2. Phonologically-conditioned breathy voice

Much of the established literature on phonation is based on studies of non-phonological voice quality differences. These studies can be divided into two general categories: (1) studies of pathologically-disordered phonation (Childers & Lee, 1991; Kreiman, Gerratt, & Antoñanzas-Barroso, 2006, 2007; Lieberman, 1963) and (2) purely phonetic studies of non-modal phonation in English (Hanson, 1995, 1997; Hanson & Chuang, 1999; Hillenbrand, Cleveland, & Erickson, 1994; Iseli, Shue, & Alwan, 2007; Klatt & Klatt, 1990). However, phonological uses of phonation within languages—both allophonic and contrastive—are arguably distinct (Blankenship, 2002; Hillenbrand et al., 1994, p. 777); like other areas of phonology, contrastive and allophonic voice quality can vary by localization (i.e. timing; Blankenship, 2002), prosodic structure (Choi, Hasegawa-Johnson, & Cole, 2005; Epstein, 2002), sociolinguistic register (Podesva, 2007), phonological environment (Pfitzinger, 2008), and various other factors. These factors are typically artificially controlled in studies of non-phonological voice quality, and thus are less suitable for investigating contrastive and allophonic uses of phonation. Given the research questions of the current study, the following literature review is restricted to studies dealing with phonologically-relevant breathy phonation.

2.2.1. Spectral measures

The increased amplitude of the first harmonic in breathy vowels is typically normalized by comparing it to the amplitude of the second harmonic, i.e. H1–H2. A higher H1–H2 (a measure of spectral balance) was shown in Holmberg et al. (1995) and Henrich (2001) to be a reliable acoustic correlate of a larger open quotient—the relative length of time of each glottal cycle's more open portion—and has been identified as the most salient acoustic property of breathy voice in !Xóõ (Bickley, 1982; Ladefoged & Antoñanzas-Barroso, 1985), White and Green (H)mong[1] (Andruski & Ratliff, 2000; Huffman, 1987), Shanghainese (Ren, 1992), Jalapa Mazatec (Blankenship, 2002; Garellek & Keating, 2010; Kirk, Ladefoged, & Ladefoged, 1993), San Lucas Quiaviní Zapotec (Gordon & Ladefoged, 2001), Xhosa (Jessen & Roux, 2002), Krathing Chong (Blankenship, 2002), Wa (Watkins, 2002), Chanthaburi Khmer (Wayland & Jongman, 2003), Suai/Kuai (Abramson, Luangthongkum, & Nye, 2004), Javanese (Thurgood, 2004), Ju|'hoansi (Miller, 2007), Takhian Thong Chong (DiCanio, 2009), Santa Ana del Valle Zapotec (Esposito, 2010b), and Southern Yi (Kuang, 2011).

Breathy phonation is also associated with a steeper spectral tilt; this is typically measured as the difference between the amplitude of the first harmonic and that of one of the first three formants. H1–A1 is thought to be correlated with posterior glottal opening at the arytenoids (Blankenship, 2002; Hanson, Stevens, Kuo, Chen, & Slifka, 2001), while H1–A2 and H1–A3 are thought to be correlated with the abruptness of vocal fold closure (Stevens, 1977); the more sinusoidal, less abrupt glottal closure involved in breathy phonation would amplify f0 and dampen the higher harmonics, which altogether produces a higher spectral tilt. Accordingly, some or all of these spectral tilt measures characterize breathy phonation in !Xóõ (Ladefoged, 1983; Ladefoged & Antoñanzas-Barroso, 1985), Mon (Thongkum, 1987a), Jalapa Mazatec (Blankenship, 2002; Garellek & Keating, 2010; Kirk et al., 1993), San Lucas Quiaviní Zapotec (Gordon & Ladefoged, 2001), Xhosa (Jessen & Roux, 2002), Krathing Chong (Blankenship, 2002), Chanthaburi Khmer (Wayland & Jongman, 2003), Suai/Kuai (Abramson et al., 2004), Javanese (Thurgood, 2004), Takhian Thong Chong (DiCanio, 2009), Santa Ana del Valle Zapotec (Esposito, 2010b), and Southern Yi (Kuang, 2011).

In addition to these more familiar spectral measures, the amplitude difference between the second and fourth harmonics (H2–H4) has been proposed as an auxiliary measure of breathiness. Kreiman et al. (2007) first introduced the measure as one of four measures that accounted for 76.6% of the variance in English voice quality as produced by speakers (most of whom had vocal pathology). This measure was later applied in Esposito's (2010a) cross-linguistic study of phonologically contrastive phonation, finding that it distinguished breathy and modal vowels in Chong, Fuzhou, Mon, and San Lucas Quiaviní, although not as well as other measures such as H1–H2. H2–H4 has also been associated with high pitch, e.g. falsetto in English and Mandarin (Bishop & Keating, 2010) and high tone in Southern Yi (Kuang, 2011). However, to the author's knowledge, H2–H4 has not yet been demonstrated to be a reliable measure of phonological breathy phonation beyond Esposito's (2010a) study.

2.2.2. Periodicity

While breathy voice is often associated with an increase in aperiodic noise, especially at higher frequencies, the complexities of the various established noise and periodicity measures tend to obscure this generalization. Measures of noise and periodicity proposed in the contrastive breathiness literature include cepstral peak prominence (CPP)—defined as a measure of peak harmonic amplitude normalized for overall amplitude as introduced in Hillenbrand et al. (1994)—and harmonic-to-noise ratio (HNR) as measured in de Krom (1993). Both of these measures are predicted to be lower in breathy phonation due to the added noise of increased airflow; accordingly, CPP was found to be lower in breathy phonation in Krathing Chong (Blankenship, 2002), Jalapa Mazatec (Blankenship, 2002; Esposito, 2010a; Garellek & Keating, 2010), and Southern Yi (Kuang, 2011), while HNR was

[1] Hmong dialects are traditionally named by descriptors that often include colors. Andruski and Ratliff (2000) investigates the dialect traditionally called Blue Mong or Green Mong (Hmong Njua). Esposito (2010c), Esposito and Khan (2010, 2012), and Khan and Esposito (2011) investigate White Hmong (Hmong Daw). Huffman (1987) collapsed speakers of both of these dialects into one group, as explained in Huffman (1985, p. 6).

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001

found to be lower in breathy or lax phonation in Ju|'hoansi (Miller, 2007) and Javanese (Wayland, Gargash, & Jongman, 1994) but not in Chanthaburi Khmer (Wayland & Jongman, 2003). Another measure established in Hillenbrand et al. (1994) is Pearson *r* at autocorrelation peak (RPK), but this measure has not been used in studies of linguistically contrastive breathiness, presumably because it was shown to largely replicate the results of the CPP measure with less reliability due to frequent errors and increased sensitivity to lower-frequency energy; consequently, RPK is not explored in the current study.

2.2.3. Other acoustic measures

Additional acoustic measures, including intensity, duration, pitch, the frequency of the first formant (F1), and the effects of tracheal coupling are less reliable in distinguishing linguistic voice quality categories in a consistent way. Overall acoustic intensity, for example, which is typically measured as root-mean-squared (RMS) energy, is often cited in voice quality studies; however, different languages yield conflicting results. While RMS energy was found to be lower for lax voice in Kui/Suai and Chong (Thongkum, 1987b), it was significantly higher in the breathy vowels of Chanthaburi Khmer (Wayland & Jongman, 2003).

Similarly, breathiness is associated with an increase in vowel duration in Kedang (Samely, 1991), Jalapa Mazatec (Kirk et al., 1993), Chanthaburi Khmer (Wayland & Jongman, 2003), Khmu' Rawk (Abramson, Nye, & Luangthongkum, 2007), and in checked syllables in Mon (Thongkum, 1987a), but not in White and Green (H)mong (Huffman, 1987), San Lucas Quiaviní Zapotec (Gordon & Ladefoged, 2001), Suai/Kuai (Abramson et al., 2004), or non-checked syllables in Mon (Thongkum, 1987a).

As phonation and pitch are both primarily manipulated by many of the mechanisms within the larynx, systematic changes in f0 are often seen in non-modal phonation crosslinguistically. Most typically, f0 is lowered during breathy vowels (Laver, 1980) in Mon (Thongkum, 1987a), Nyah Kur (Thongkum, 1987b), Kui/Suai (Thongkum, 1987b), Western A-Hmao (Johnson, 1999), and Khmu' Rawk (Abramson et al., 2007), and accompanies low tone in lexical tone languages such as Green Mong (Andruski & Ratliff, 2000). However, there are exceptions to this general tendency, with languages like Jalapa Mazatec showing no effect of phonation on f0 (Garellek & Keating, 2010), and even languages like Chanthaburi Khmer exhibiting higher f0 on breathy vowels (Wayland & Jongman, 2003). In fact, Madiesson and Hess (1987) document the inconsistent relation between lax phonation and f0 in five minority languages of China: while Jingpho, Yi, and Lahu exhibit f0 lowering during lax phonation, Wa and Lisu show no such effect.

The frequency of the first formant (F1) has been found to be lower in breathy vowels cross-linguistically; lower F1 indicates an acoustically higher vowel. This has been found in Chong (Thongkum, 1987b), Kedang (Samely, 1991), and several Nilotic languages (Denning, 1989), as well as in vowels following weak, non-contrastive consonant breathiness (often called "slack voice") in Shanghainese (Ren, 1992), Xhosa (Jessen & Roux, 2002), and the non-high vowels of Javanese (Thurgood, 2004). However, F1 shows no consistent association with breathy phonation in some Southeast Asian languages such as Mon[2] (Thongkum, 1987a), Nyah Kur (Thongkum, 1987b), Kui/Suai (Thongkum, 1987b), Green Mong (Andruski & Ratliff, 2000), or Khmu' Rawk (Abramson et al., 2007), or in Jalapa Mazatec (Garellek & Keating, 2010). Of course, there are many difficulties inherent in measuring formant frequencies—amplitudes can be weakened in high rounded vowels, surrounding consonants introduce formant transitions, and nasalization introduces nasal formants and antiformants—and thus it is also quite possible that further investigation will reveal a clearer picture of the effect of phonation on vowel quality.

Due to the acoustic damping effects of the tracheal coupling associated with unimpeded airflow through the glottis (Fant, 1972), a wider first formant bandwidth (increased B1) is also associated with breathiness, and has been shown to be a reliable indicator of the phonation type in Southern Yi (Kuang, 2011). Wider F1 bandwidth can also be inferred by measuring H1–A1 (Hanson, 1997, p. 471); a larger H1–A1 in Chanthaburi Khmer breathy vowels was taken to be indicative of damping (Wayland & Jongman, 2003).[3]

Measures of cycle-to-cycle regularity, including jitter (a measure of period consistency across cycles, Lieberman, 1963) and shimmer (a measure of amplitude consistency across cycles, Horii, 1980), primarily distinguish tense and creaky phonation from other voice qualities, and are not considered reliable unless performed on purely monophthongal vowels produced in isolation and with sustained pitch (Andruski & Ratliff, 2000, p. 55; Horii, 1982). For these two reasons, they are poor indicators of breathiness in naturalistic speech, and are consequently not explored in the current study.

2.2.4. Electroglottographic measures

Electroglottography is a non-invasive technique for investigating the articulatory properties of the glottis. This involves sending an electrical signal between two electrodes placed on either side of the subject's larynx, using a collar around the neck. This configuration produces a signal representing the degree of electrical conductance between the electrodes: higher conductance corresponds to greater vocal fold contact (as human tissue is a relatively good conductor of electricity) while lower conductance corresponds to lesser contact. In this way, degrees of greater and lesser contact due to different phonation types can be studied directly from the EGG signal; breathier phonations are predicted to have less contact than modal phonations, which are predicted to have less contact than creakier phonations.

Like acoustic studies of voice quality, electroglottographic studies can be characterized as investigations of either non-phonological or phonological uses of breathy phonation. Examples of the former type of study, which establish the connection between the electroglottographic signal and vocal fold movement, include Scherer, Druker, and Titze (1988) and Baken and Orlikoff (2000), among others. Examples of the latter type, which investigate how electroglottographic analysis can distinguish linguistic phonation types, include Guion, Post, and Payne's (2004) study of Maa, Mazaudon and Michaud's (2008) study of Tamang (whispery voice), DiCanio's (2009) study of Takhian Thong Chong, Esposito's (2010c) study of White Hmong, Kuang's (2011) study of Southern Yi (lax voice), and Esposito and Khan's (2010, in press) comparison of Gujarati and White Hmong. These studies have all found that breathier phonation types involve a smaller Contact Quotient (also known as closing quotient, closed quotient, or CQ), the relative portion of time that the glottis is more closed during each pulse. (See Henrich, d'Alessandro, Doval, & Castellengo, 2004; Herbst, 2005; Herbst & Ternström, 2006 for discussions on how to measure CQ.) These results suggest a more

[2] Thongkum (1987a) does find some reliable associations between F1 and phonation in Mon, but the effects are different across different parts of the vowel space, and interact with syllabic structure: the pair /i i̤/ can be distinguished by F1 in syllables closed by a voiceless consonant, while the pair /ɛ ɛ̤/ can be distinguished by F1 in open syllables and syllables closed by a voiced consonant. The pair /o o̤/ can be distinguished by F1 in all three syllable types.

[3] Also see Pennington (2005) for a discussion of two related measures: first formant quality (Q1), measured as first formant frequency divided by bandwidth (F1/B1) and first formant peak energy factor (PE1).

open glottis in breathier phonations. In addition, Esposito (2010c), Esposito and Khan (2010, 2012), and Kuang (2011) found that breathier phonation types in White Hmong and Southern Yi involve a significantly higher value in a measure known as Derivative-EGG Closure Peak Amplitude (DECPA), measured as the peak positive value in the first derivative of the EGG signal (Michaud, 2004; Vũ-Ngọc et al., 2005); this suggests a faster speed of vocal fold closing in those voice qualities. DECPA is also known as Peak Increase in Contact (PIC) in Keating et al. (2010).

2.3. Contrastive breathy voice in Gujarati

2.3.1. Background

Like most Indic languages, Gujarati (Indo-European; western India; 46 million: Lewis, 2009) distinguishes two voiced consonant types (i.e. voiced unaspirated vs. voiced aspirated) in addition to two voiceless types (i.e. voiceless unaspirated vs. voiceless aspirated). The former two distinguish modal voice[4] and breathy voice (e.g. [baɾ] 'twelve' vs. [bʱaɾ] 'burden'), respectively. Less common in Indic languages is the Gujarati contrast in modal [i e ɛ a ɔ o u ə][5] and breathy/"murmured" [i̤ e̤ ɛ̤ a̤ ɔ̤ o̤ ṳ ə̤] vowels (e.g. [baɾ] 'twelve' vs. [ba̤ɾ] 'outside').[6] While a handful of examples of breathy vowels are believed to have been derived independently (e.g. [kə̤ro] 'wall', [na̤nũ] 'small', see Cardona & Suthar, 2003, p. 666; Masica, 1993, p. 147; Mistry, 1997, p. 668), the vast majority occur as a result of various processes involving sequences of modal vowels and breathy consonants (Cardona, 1965, pp. 29–30, 50; Cardona & Suthar, 2003, pp. 665–666; Dave, 1967, pp. 1–2; Fischer-Jørgensen, 1967, p. 73; Masica, 1993, p. 120; Mistry, 1997, pp. 666–669; Pandit, 1957, pp. 169–170).[7] The four most stable breathy vowels [ɛ̤ a̤ ɔ̤ ə̤] primarily emerged historically from the fusion of sequences of the type [əɦV]: [əɦe əɦa əɦo əɦə], e.g. [bɛ̤n] 'sister' from historical [bəɦen].[8] (Breathy diphthongs [ə̤j ə̤w] also arise from this source, specifically from sequences of [əɦi əɦu], respectively.) A second source of breathy vowels is the result of another historical sequence of vowels and an intervening breathy consonant, [Vɦə]; the sequences [eɦə aɦə oɦə] are typically produced as [e̤ a̤ o̤] in modern Gujarati dialects, but are subject to more variation, e.g. [ʋaɦən]~[ʋa̤n] 'vehicle'. In very casual speech, two additional sources of breathy vowels are found; breathy vowels optionally arise from word-initial sequences of [#ɦV], e.g. [ɦaɾ]~[a̤ɾ] 'necklace', or from the transfer of breathiness from an underlyingly breathy-voiced consonant to its surrounding vowels [VCʱV] with optional lenition, e.g. [ʋəɖʱu]~[ʋə̤ɖṳ]~[ʋə̤ðṳ] 'more' (Cardona, 1965, pp. 23, 50; Cardona & Suthar, 2003, p. 666; Firth, 1957, p. 235; Mistry, 1997, p. 667; Pandit, 1957, p. 171). Through [#ɦV] and [VCʱV] sources, all vowel qualities can be produced with breathy phonation [i̤ e̤ ɛ̤ a̤ ɔ̤ o̤ ṳ ə̤], albeit with less reliability due to the requirement of a casual register and fast speech rate.

With the exception of a handful of words that show no orthographic representation of breathiness, e.g. ક્યારે ⟨kjare⟩ [kja̤ɾe] 'when' (Cardona, 1965, p. 57), most instances of breathy vowels are still written in Gujarati orthography using sequences of vowels and intervocalic હ ⟨ɦ⟩. Thus, speakers sometimes read words as they are spelled instead of producing a single breathy vowel (Cardona & Suthar, 2003, pp. 665–666; Masica, 1993, p. 120; Turner, 1921, p. 529), e.g. occasionally pronouncing [pɛ̤lũ] 'first (neut.)' as [pəɦelũ]~[pəɦɛlũ] due to the spelling પહેલું ⟨pə.ɦe.lũ⟩. Unsurprisingly, this pronunciation is most common in read speech and more formal registers, and was identified as a problematic factor in the recording of read speech for Dave's (1967, p. 4) formant analysis. Furthermore, because the production of breathy vowels from the sources [Vɦə], [#ɦV], and [VCʱV] is heavily dependent on a casual register and fast rate, these words are the most likely to be produced in a spelling pronunciation when in a formal, laboratory setting. Moreover, in some sociolinguistic registers of particular dialects, breathy vowels can be produced as monophthongs, but with modal phonation, e.g. [bɛɾaʃ] for બહેરાશ [bɛ̤ɾaʃ] 'deafness' (Cardona & Suthar, 2003, p. 666; Dave, 1967, p. 2); this has been noted as a marker of an "educated" speech style in Pandit (1957, p. 170); Nair (1979, p. 22), and Masica (1993, p. 120). For additional references on the historical and phonological perspective of breathy vowels in Gujarati, see Pandit (1969, p. 117), Dave (1977), and Vyas (1977, pp. 74–76).

2.3.2. Acoustic studies

Most of the established knowledge of the phonetic realization of the Gujarati phonation contrast comes from four acoustic studies, which have in many ways shaped our general understanding of both the acoustics and the perception of linguistic voice quality cross-linguistically. Pandit (1957) first characterized breathy vowels in Gujarati as involving "voiced breath, low pitched (p. 169)" based on his native intuition, and "random distribution of energy, more noticeable at higher frequencies (p. 172)", using kymographic evidence. Fischer-Jørgensen (1967) quantified Pandit's findings, determining that lower f0 (particularly at the vowel onset), lower RMS energy, and the occasional presence of aspiration noise (assessed visually) were all moderately reliable cues to breathy vowels; however, she found that comparing the relative amplitude of the first harmonic (H1) with that of the second harmonic (H2) and with that of the three formants A1, A2, and A4 was a far more reliable measure of breathiness. Fischer-Jørgensen concluded that indeed, breathy vowels have a steeper spectral balance (i.e. higher H1–H2) and a steeper spectral tilt (i.e. higher H1–A1, H1–A2, H1–A4), and that only these spectral measures—not the measures of pitch, intensity, or aspiration noise—were found to be reliable indicators of breathiness across tokens and speakers. Like Fischer-Jørgensen, Dave (1967) also found that breathy vowels may be associated with lower f0, but the effect was small, and restricted to the vowels [i̤ ə̤]; he also showed that breathy vowels are slightly more open than their corresponding modal vowels,[9] and involve a reduction in energy in the higher frequencies, especially around F4 (p. 29). Later, Bickley (1982) identified a high H1–H2 as a strong characteristic of breathy vowels produced by four speakers,

[4] When not adjacent to breathy vowels, modally-voiced stops can be produced with faint implosion followed by a brief period of tense voice (Firth, 1957, pp. 234–235), e.g. [baɾ] 'twelve' produced as [ɓa̰͡aɾ]. This presumably amplifies the contrast with breathy-voiced consonants (e.g. [bʱaɾ]~[ba̤͡aɾ] 'burden') and vowels (e.g. [ba̤ɾ] 'outside'), neither of which are characterized by implosion or tense voice. See Modi (1987) for an alternative analysis in which tense ("tight") phonation is considered a feature specific to peninsular dialects.

[5] Not all dialects have the full inventory of vowels shown here; some dialects lack the low-mid modal vowels [ɛ ɔ] (Firth, 1957, pp. 231–232, Pandit, 1961, pp. 62–63).

[6] Outside Gujarati, the breathy vs. modal contrast across both consonants and vowels is seen only in Khoisan (Khoesan) languages (Traill, 1985) and White Hmong (Esposito, 2010c); see Esposito and Khan (2010, 2012) and Khan and Esposito (2011) for an analysis of the acoustic and EGG properties of breathy-voiced aspiration vs. breathy vowels in Gujarati and White Hmong.

[7] The two types of breathy segments (i.e. breathy-aspirated consonants and breathy vowels) do not cooccur in the same syllable (Pandit, 1957, p. 169), largely due to a cooccurrence restriction in pre-modern Indic languages (Grassmann's Law; see Whitney, 1889; Wackernagel, 1896) and a low number of borrowed or newly-coined words with breathy consonants in the modern lexicon.

[8] Pandit (1957) analyzes breathy vowels as underlying sequences of a modal vowel and [h], e.g. /bahɾ/→[ba̤ɾ] 'outside', a phonemic representation adopted by most subsequent studies.

[9] This slight lowering of breathy vowels in Gujarati is typologically exceptional; Denning (1989) documents a much larger number of languages that show *raising* of breathy vowels instead.

reconfirming Fischer-Jørgensen's conclusions; she did not report other acoustic findings. Bickley also used inverse filtering to reveal their more sinusoidal glottal waveform.[10] Thus, the general consensus from these four early acoustic studies of Gujarati is that breathy voicing can be characterized most reliably by a higher H1–H2, but also by a lower f0, higher H1–A1, H1–A2, and H1–A4, lower RMS energy, possibly lower vowel height, and occasional aspiration noise. Later acoustic studies of Gujarati breathy vowels include Khan and Thatte (2009) and Khan (2009, 2010), which are earlier versions of the current study, and Keating et al. (2010), Esposito and Khan (2010, 2012), and Khan and Esposito (2011), which are comparative studies of Gujarati and White Hmong; due to their close relation to the current work, their findings on Gujarati are not reviewed here in detail.

2.3.3. *Perception studies*

While previous studies have identified multiple acoustic properties of Gujarati breathy vowels, perception studies show a far narrower range of cues are actually used by Gujarati-speaking listeners when categorizing voice quality. In a smaller perceptual component of her study, Fischer-Jørgensen (1967) found that speakers categorized a vowel as breathy when the fundamental frequency was low but its amplitude (i.e. H1) was high relative to the rest of the spectrum, and when audible noise was present in the higher frequencies. Bickley (1982), however, found that while both a high H1–H2 value and aspiration noise were typical characteristics of breathy vowels produced by her four Gujarati speakers, only H1–H2 was used by those same speakers in categorizing synthesized vowels; noise had no significant effect on their perception of breathiness. Furthermore, Esposito (2010a) found that Gujarati-speaking listeners attended solely to H1–H2 when categorizing vowels excised from various languages (i.e. Chong, Fuzhou, Green Mong, White Hmong, Mon, Santa Ana Del Valle Zapotec, San Lucas Quiaviní Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ), effectively ignoring variations in CPP, H1–A1, H1–A2, H1–A3, H2–H4, A2–A3, and ((H1+H2)/2)–A1. Interestingly, the Gujarati-speaking listeners consistently used H1–H2 variation in their judgments regardless of whether H1–H2 could actually distinguish phonation categories in the recorded languages. Later, Kreiman, Gerratt, and Khan (2010) confirmed that Gujarati speakers are indeed more sensitive to small changes in H1–H2 than speakers of languages without a phonemic contrast in breathiness, presumably due to the important role it plays in voice quality categorization in Gujarati. The just-noticeable differences (JNDs) in H1–H2 for Gujarati speakers was 2.60 dB, while English and Thai speakers' JNDs were 3.67 dB and 3.35 dB, respectively, indicating that Gujarati speakers were able to perceive finer distinctions in H1–H2 than could speakers of other languages. From these perceptual studies, it is clear that Gujarati speakers are especially sensitive to H1–H2, and largely ignore other acoustic cues, in categorizing vowels into contrastive phonation types.

2.3.4. *Articulatory studies*

In comparison to the fair amount of attention given to the acoustics and perception of Gujarati breathy vowels, their articulatory properties have thus far not been examined in as much detail. To my knowledge, the few studies of Gujarati phonation that incorporate an articulatory component all examine data from no more than three subjects each. Examining breathy vowels produced by a subset of her subject pool, Fischer-Jørgensen (1967) reported a shorter closed phase and what she interpreted as possibly a wider glottis in an EGG analysis of two speakers, and greater airflow in an aerodynamic analysis of three speakers. In Modi (1987), tomographic data served as the empirical base for a proposal to split Gujarati into "murmur" dialects and "tight phonation" dialects. Breathy phonation in the "murmur" dialects—which include Standard Gujarati—was shown to be produced with a lowered and widened glottis in the word [kɛ̤ʋət̪] 'proverb' ([kHe̤vAt] in her transcription). However, the majority of Modi's (1987) study focused on refuting theoretical issues raised in previous studies, and the articulatory data were not further analyzed. With such limited views into the articulation of Gujarati breathy vowels, it is clear that there is still much to be studied in terms of what is physiologically occurring in the vocal tract during their production.[11]

2.4. *Motivation for the current study*

Considering the wide array of languages (including Gujarati) whose phonation contrasts have been shown in the literature to be reliably distinguished by H1–H2, and considering Gujarati speakers' sensitivity to H1–H2 when listening to other languages, it is safe to say that H1–H2 should be a reliable measurement distinguishing modal and breathy vowels in Gujarati. However, additional questions remain, due to important issues left unanswered in previous work.

The first issue is that, with the use of modern automated technology to examine a larger corpus of data, it is now possible to investigate many more acoustic and EGG measures mentioned in previous studies of other languages to determine which of them can also reliably distinguish the phonation categories of Gujarati. Second, spectral measures such as H1–H2 are no longer used without various adjustments to correct for the effects of surrounding formant frequencies and bandwidths (Hanson, 1995), but none of the early studies of Gujarati phonation accounted for these effects. Especially when comparing across vowel qualities, such corrections are imperative. Third, previous studies measured acoustic properties on the averages of entire vowel durations, thus overlooking the possibility that phonation is dynamic and that certain cues may be localized to certain portions of the vowel (Blankenship, 2002; Gordon & Ladefoged, 2001, pp. 11–13). Fourth, articulatory investigations of Gujarati breathy voice (including EGG studies) are particularly limited, leaving much to be explored.

Additionally, there is the issue of data quantity and quality; previous acoustic studies examined either a small number of speakers (e.g. Bickley, 1982; Dave, 1967) or different sets of data for different speakers (e.g. Fischer-Jørgensen, 1967), thus not carefully controlling for interspeaker variation. Furthermore, such studies investigated only words in isolation and in predetermined phrases produced in a laboratory setting despite the established caveats regarding the effects of spelling pronunciations and of the "educated" speech register; thus additional work is still needed to confirm the validity of previous work in a wider context.

Thus, the current study seeks to answer the following five questions, using a stricter methodology described in the subsequent section:

1. **Acoustics**: In addition to H1–H2, what acoustic measures reliably distinguish Gujarati phonation types?

[10] Also worthy of note is Rami, Kalinowski, Stuart, and Rastatter (1999), which examined voice onset times and burst frequencies in Gujarati stop consonants, finding that the properties of breathy aspiration were not immediately identifiable given their limited acoustic parameters; however, their study was not concerned with vocalic breathiness.

[11] A fiberscopic investigation of the articulation of breathy-voiced aspirated stops in one Gujarati speaker is covered in an educational video (Hayward, 1992). However, measurements are not presented, and breathy vowels are not discussed.

2. **Electroglottography**: Can electroglottographic measures reliably distinguish these phonation types?
3. **Localization**: Are the effects of certain cues localized to particular points in a vowel's duration?
4. **Vowel quality**: Are the acoustic cues for breathiness affected by vowel quality?
5. **Articulation**: What can the acoustic and electroglottographic results tell us about how Gujarati speakers produce these voice qualities?

3. Methods

3.1. Subjects

Ten native Gujarati-speaking subjects (3 M, 7 F) participated in the current study; all were college-educated and fully literate in the language. None had any prior academic background in linguistics. At the time of recording, all but two subjects were in their 20s or 30s and had spent the majority of their lives in India, having only recently (< 1 year prior) moved to the US at the time of recording. Of the remaining two speakers, one was in her 50s and had lived in the US for 26 years, and the other was in her 20s and had lived in the US for three years. All speakers are also fluent in English as well as various Indian languages, most commonly Hindi and Marathi.

3.2. Controlling for sociolinguistic effects

Because all subjects were college-educated adults and because of the formal atmosphere inherent in laboratory recordings, the choices of stimuli and task had to be controlled to avoid triggering both the neutralization of breathiness as seen in "educated" speech (e.g. [banũ] for [ba̤nũ] 'excuse'; see Masica, 1993, p. 120; Pandit, 1957, p. 170) and the over-enunciating disyllabic productions as seen in slow, formal, read speech (e.g. [bəɦanũ] for [ba̤nũ] 'excuse') as in Dave (1967, p. 4). In this way, the current study departs from previous studies in controlling for these sociolinguistic effects in three ways: highly academic or scientific words were excluded to avoid the "educated" speech style, a special recording procedure was used to minimize the effect of orthographic representation (see Section 3.5), and words of the relatively unstable [Vɦə], [#ɦV], and [VCʱV] sources were not examined; only those from the most stable source (i.e. [əɦV]→[V̤]) were used. Two native speaker consultants, who were aware of the purpose of the study and therefore did not serve as experimental subjects, were especially helpful in stimulus selection and task development.

3.3. Stimuli

All 26 stimulus items used in the current study had either target modal vowels or target breathy vowels of the most stable source type (i.e. [əɦV]→[V̤]), in words predicted by native speaker consultants to be familiar to the average speaker and not overly academic or literary. Stimuli were selected with the use of a Gujarati–English dictionary (Suthar, 2003).[12] Some are members of minimal pairs (e.g. [baɾ] 'twelve' vs. [ba̤ɾ] 'outside'), while others are members of near-minimal pairs (e.g. [keʋũ] 'of what kind' vs. [ke̤ʋũ] 'to say'), due to constraints on the Gujarati lexicon.[13] Because of the potential long-distance effects of breathy phonation (Modi, 1987, pp. 46–50), words with target modal vowels but breathy phonation in other segments were excluded. The full list of stimuli can be found in the appendix.

3.4. Recording

Recordings were made in 2008–2009 in a sound-proof booth in the phonetics laboratory at UCLA. In each recording session, both subject and investigator were seated in front of a desktop computer running the software program PCQuirerX. The subject both held the microphone and wore an EGG electrode collar around his or her neck. The microphone was attached to an external pre-amplifier, which was attached to the computer by a USB cable. The electrodes were attached to the EGG, which was in turn attached to the same external box to which the microphone input was attached. The visual readings of the audio and EGG signals were visible to both investigator and subject in the PCQuirerX display and the EGG box itself, so the subject could adjust the volume of his or her voice, and so the investigator could adjust the positioning of the EGG electrodes, if needed.

3.5. Task

Due to concerns described above, a task had to be devised that would minimize exposure to the orthographic representation of the target word and maximize speech rate, with the intention of inducing more casual speech. Thus, the investigator sat to the right of the subject, and briefly (< 2 s) displayed the first flash card—an example word, not recorded—with both the Gujarati word કુત્રો [kut̪ro] and its English translation 'dog' handwritten on it. The investigator explained that the subject should immediately think of a short sentence beginning with the word, such as કુત્રોભાગી ગયો [kut̪ro bʱagi gəjo] 'the dog ran away,' and say it out loud into the microphone as many times as possible during the timed recording. Once the subject understood the task, the investigator began displaying the flash cards containing target stimuli. For each word, the investigator created a recording in PCQuirerX with a fixed duration of 10 s, and the subject repeated the sentence he or she thought up as many times as could be accommodated in that time. Depending on the speaker's typical speech rate and on the length of the sentence created by the speaker, each recording typically contained between four and six fluent sentences. Between recordings, the investigator frequently asked the subject about the translations of some sentences to keep the subject focused on meaning rather than on spelling.

3.6. Annotation

To prepare each recording for analysis, a Praat text grid (Boersma & Weenink, 2010) was created for each audio file, with one interval tier in which the target vowel of each stimulus word was later segmented and transcribed, as shown in Figs. 1 and 2. Because each recording's audio file and EGG file were of the exact same length (due to the fixed 10-s recording window), the segment boundaries in the text grid could be accurately aligned

[12] Many of the stimuli selected for the current study were used in a pilot study conducted by Victoria Thatte. All but one stimulus was taken from Suthar (2003). The remaining stimulus (i.e. [ka̤n]) was taken from Ladefoged and Maddieson (1996).

[13] Due to the peculiarities of Gujarati historical and dialectal phonology, low-mid modal vowels [ɛ ɔ] contrast with the high-mid modal vowels [e o] in word-initial syllables (e.g. [mɔko] 'favorable opportunity' vs. [mokl̤ũ] 'spacious', 'open'), but only in certain dialects; many other dialects have only high-mid modal vowels [e o] (Cardona & Suthar, 2003; Firth, 1957). All dialects with contrastive breathiness have low-mid breathy vowels [ɛ̤ ɔ̤]. Thus, for many speakers, the mid-vowel (near-)minimal pairs in fact contrast high-mid modal vowels and low-mid breathy vowels (e.g. [mel] 'filth' vs. [mɛ̤l] 'palace' instead of standard [mɛl] 'filth' vs. [mɛ̤l] 'palace').

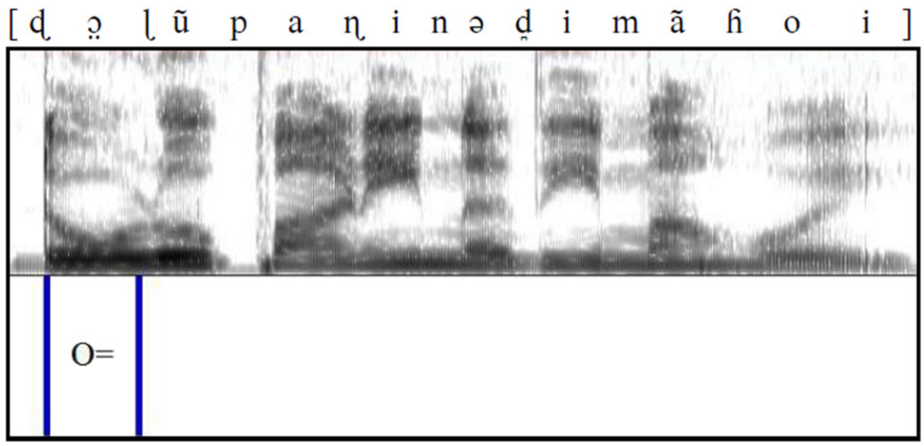

Fig. 1. The target vowel [ə̤] in the word [ɖə̤ɭũ] 'polluted (neut.)' is segmented and labeled ⟨O=⟩ in the Praat text grid of the sentence [ɖə̤ɭũ paɳi nəɖ̤i-mã ɦoi] 'there may be polluted water in the river', created and produced by Speaker 9 (female, 30, from Mumbai).

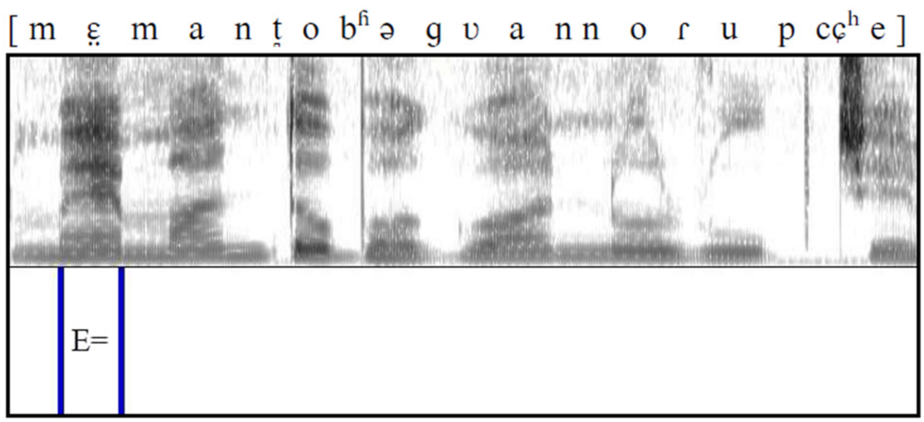

Fig. 2. The target vowel [ɛ̤] in the word [mɛ̤man] 'guest' is segmented and labeled ⟨E=⟩ in the Praat text grid of the sentence [mɛ̤man t̪o bʱəɡʋan-no ɾup cçʰe] 'but a guest is a manifestation of god', produced by Speaker 6 (female, 29, from Mumbai).

with the simultaneously-recorded EGG signal to identify the target vowels for the semi-automated acoustic and EGG analysis described below.

3.7. Analysis

The audio data was then analyzed using VoiceSauce (Shue, Keating, & Vicenik, 2009), a free software program developed to semi-automatically make multiple acoustic measures for each annotated input file. Each acoustic measure was calculated at every millisecond of target vowel duration, and then averaged within nine equal parts of the vowel, making it possible to see changes in acoustic values over these nine snapshots (henceforth "timepoints") of the vowel's duration. The acoustic parameters measured for the current study include *Amplitude Differences of various Harmonics and Formants* (i.e. H1–H2, H2–H4, H1–A1, H1–A2, H1–A3) corrected for the effects of surrounding formant frequencies and bandwidths (Hanson, 1995) using the algorithm developed in Iseli et al. (2007), *Fundamental Frequency (f0)* as calculated by the STRAIGHT algorithm of Kawahara, de Cheveigne, and Patterson (1998), and three measures that examine a variable window of analysis equal to five pitch pulses: *Root Mean Square (RMS) Energy* (a measure of overall intensity), *Cepstral Peak Prominence (CPP)* as defined in Hillenbrand et al. (1994), and *Harmonic-to-Noise Ratio (HNR)* as calculated in de Krom (1993), focusing on four regions of the spectrum: 0–500 Hz, 0–1500 Hz, 0–2500 Hz, and 0–3500 Hz.

In some cases, manual intervention was needed during the acoustic analysis. The formant tracking algorithms in VoiceSauce were at times unable to properly identify F1; this was most common when F1 was either weakened due to aspiration noise or difficult to distinguish from a strong H1, both properties being typical of breathy voice. In these cases, formant listing from Praat (Boersma & Weenink, 2010) were manually fed into VoiceSauce to recalculate the amplitudes of those formants and thus accurately correct the spectral measures.

The EGG signal was analyzed using EggWorks (Tehrani, 2009), a free software program developed to work with VoiceSauce, automatically making multiple EGG measures for each annotated input file. As with the acoustic measures, each EGG measure was calculated at every millisecond of target vowel duration, and then averaged within the nine timepoints of the vowel. There were two EGG parameters measured for the current study, the first being *Contact Quotient (CQ)*: the duration of the closed phase relative to the open phase of each glottal cycle, measured here using a hybrid method with the maximum value of the first derivative of each cycle of the EGG signal marking the onset of the closed phase and a 25% threshold marking its offset. Hybrid methods of calculating CQ—roughly, those in which the onset and offset of the closed phase are determined using different criteria—have been used in electroglottography since the 1980s. See Davies, Lindsey, Fuller, and Fourcia (1986), Orlikoff (1991), and Howard (1995) for a review of some of the first hybrid methods proposed. Use of the first derivative of the EGG to determine the onset of the closed phase has been supported by multiple imaging studies of glottal closure,

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001

e.g. Anastaplo and Karnell (1988), Baer, Löfqvist, and McGarr (1983), Childers, Naik, Larar, Krishnamurthy, and Moore (1983), Hess and Ludwigs (2000), Henrich et al. (2004) and Karnell (1989).

The second EGG parameter measured for the current study was *Derivative-EGG Closure Peak Amplitude (DECPA)*, measured as the peak positive value in the first derivative of the EGG signal (Michaud, 2004; Vũ-Ngọc et al., 2005, see also PIC in Keating et al., 2010).

4. Results

As can be observed in pairs such as the one presented in Fig. 3, breathy vowels are often easily distinguishable through qualitative spectrographic analysis from their modal counterparts. The most easily distinguishable breathy vowels have visibly weakened formant structure and increased aperiodic noise, especially above 2000 Hz. However, many pairs collected in the current study involved a distinction that could not be identified so clearly.

As a preliminary step in the quantitative analysis, all acoustic and EGG measurements from the multiple recordings of the same target word produced by the same speaker were averaged within each of the nine timepoints of equal duration, before undergoing statistical analysis. The results of comparing modal and breathy vowels along the various acoustic and EGG parameters are summarized in Table 1. The statistical results are drawn from two-way repeated measures ANOVAs with phonation and vowel quality as the two independent variables and subject as the random sampling variable. Data in the "main effect of phonation" column indicate whether there was a significant ($p < 0.01$) main effect of breathiness on the measure at any timepoint within the vowel's duration. If the measure also significantly ($p < 0.01$) distinguishes voice quality categories averaging across the entire vowel duration, two asterisks appear in the "main effect of phonation" column. Most often, a given measure significantly distinguished vowel quality categories at some but not all

Table 1
Results of each acoustic and electroglottographic measure's capacity to distinguish breathy and modal vowels.

Measure	Main effect of phonation	Localization	Interaction with vowel quality
Spectral measures			
H1–H2	V̤ is higher**	points 1, 4–7	point 4
H2–H4	V̤ is higher**	points 4–6	none
H1–A1	V̤ is higher**	all points	point 8
H1–A2	V̤ is higher**	all points	none
H1–A3	V̤ is higher**	points 1, 3–9	none
Periodicity measures			
HNR: 0–500 Hz	V̤ is lower	point 2	none
HNR: 0–1500 Hz	V̤ is lower	points 2, 5–6	point 1
HNR: 0–2500 Hz	V̤ is lower	points 2, 5–7	points 1, 9
HNR: 0–3500 Hz	V̤ is lower	points 2, 5–7	none
CPP	V̤ is lower	points 4–6	points 3–5
Other acoustic measures			
RMS energy	none	n/a	none
f0	none	n/a	points 4–9
EGG measures			
CQ	V̤ is lower**	points 2–9	none
DECPA	none	n/a	none

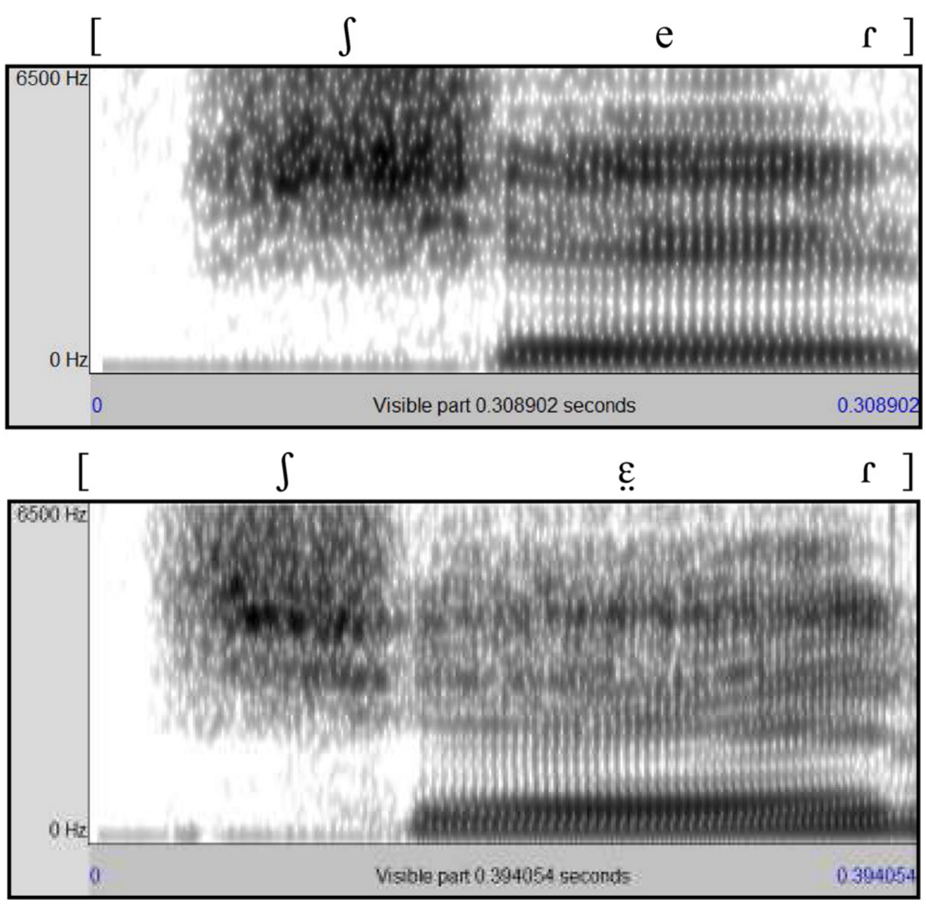

Fig. 3. Praat (Boersma & Weenink 2010) images of wide-band spectrograms of the words [ʃeɾɖi] 'sugarcane' (top, initial syllable only) and [ʃɛ̤ɾ] 'city' (bottom) produced by Speaker 9 (female, 30, from Mumbai), showing the contrast of [e] vs. [ɛ̤].

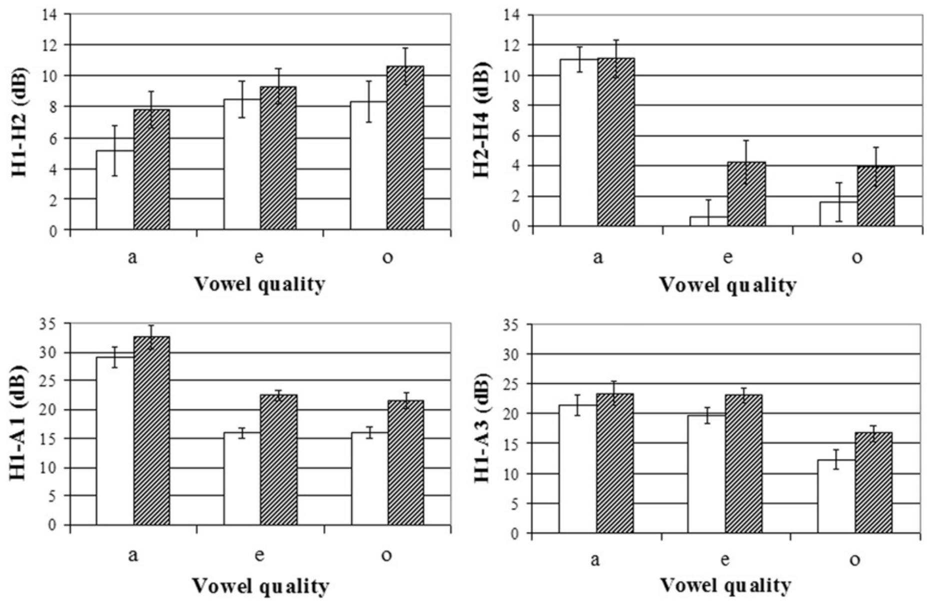

Fig. 4. Spectral measures: H1–H2 (top left), H2–H4 (top right), H1–A1 (bottom left), and H1–A3 (bottom right) values of modal (solid) and breathy (striped) vowels, averaged across the entire vowel duration and speaker, separated by vowel quality; error bars are standard error.

timepoints; timepoints with significant ($p<0.01$) main effects are listed under the "localization" column. Timepoints exhibiting significant ($p<0.01$) interactions of phonation and vowel quality are listed in the rightmost column.

4.1. *Acoustic results*

The results confirm that several acoustic cues, including some not mentioned in previous studies, reliably contrast modal and breathy vowels in Gujarati; many of these cues are localized to specific parts of the vowel, and several also interact with vowel quality. Results of spectral measures are presented in Section 4.1.1, results of periodicity measures in Section 4.1.2, and results of other acoustic measures in Section 4.1.3.

4.1.1. *Spectral measures*

As expected, there were significant effects of phonation on four familiar spectral measures, averaging across all nine timepoints of the vowel: H1–H2 [$F(1, 196)=23.37$; $p<0.01$], H1–A1 [$F(1, 196)=48.92$; $p<0.001$], H1–A2 [$F(1, 196)=50.09$; $p<0.001$], and H1–A3 [$F(1, 196)=23.44$; $p<0.01$]. A similar effect was seen on H2–H4 [$F(1, 196)=13.03$; $p<0.01$], which had not been previously studied in Gujarati. As expected, breathy vowels exhibited the higher value in all five spectral measures, shown for four of these measures in Fig. 4. Higher values indicate that the lowest-frequency harmonic has a high amplitude relative to the rest of the spectrum, a common property of breathier phonations cross-linguistically. Because the current study explores corrected spectral measures to control for the effects of formant frequencies and bandwidths on the different vowels examined, it was possible to increase the number of measures investigated beyond those of these earlier studies and also avoid the problems identified in Fischer-Jørgensen (1967), where some spectral tilt measures (in particular, H1–A3) were unreliable indicators of breathiness on certain vowels. There was no significant interaction between any spectral measures and vowel quality averaging across the entire vowel duration.[14]

The three measures of spectral tilt (i.e. H1–A1, H1–A2, H1–A3) not only showed a significant increase in breathy vowels when averaging across the nine timepoints, they also showed significant increases at each individual timepoint, with one exception: H1–A3 was higher during timepoint 2 of breathy vowels only at a weaker threshold of significance ($p<0.05$). The two harmonic measures (i.e. H1–H2, H2–H4), however, showed a localization effect: while both were significantly higher in breathy vowels than in modal vowels during the mid-region (i.e. timepoints 4–6), their difference at other timepoints was significant only at a weaker threshold of significance ($p<0.05$). Because of this difference in statistical significance, the timecourse of breathy vowels was investigated; in Fig. 5, the timecourses for H1–H2, H2–H4, H1–A1, and H1–A3 in breathy and modal vowels are shown. Note that breathiness increases in the mid-region of breathy vowels across all spectral measures, more notably in some measures than others. This localization may seem somewhat trivial in spectral measures, as the overall means are significant to begin with, but it becomes much more important in measures of periodicity.

4.1.2. *Periodicity measures*

Averaging across the entire vowel, there was no significant effect of phonation on the HNR within the 0–500 Hz range [$F(1, 196)=0.01$; $p=0.93$]. There was a marginally significant effect of phonation on the HNR of the 0–1500 Hz range [$F(1, 196)=5.79$; $p=0.05$], the 0–2500 Hz range [$F(1, 196)=9.51$; $p=0.02$], and the 0–3500 Hz range [$F(1, 196)=9.97$; $p=0.02$], as well as on CPP [$F(1, 196)=8.80$; $p=0.02$], with breathy vowels consistently exhibiting lower (less periodic) values. These last two measures are shown in Fig. 6, although they are not listed in Table 1, which reports only the effects found to be significant at $p<0.01$.

These results seem to suggest that the periodicity measures used are not reliable indicators of breathiness. This finding stands in contrast to results of early studies (e.g. Fischer-Jørgensen, 1967), in which periodicity (measured qualitatively) was thought to be

[14] There were, however, isolated interactions at two timepoints: H1–A1 at timepoint 8 [$F(2, 196)=6.53$; $p<0.01$] and H1–H2 at timepoint 4 [$F(2, 196)=7.20$; $p<0.01$]. Given the lack of a clear pattern in these two phonation–vowel quality interactions, they are not discussed in greater detail.

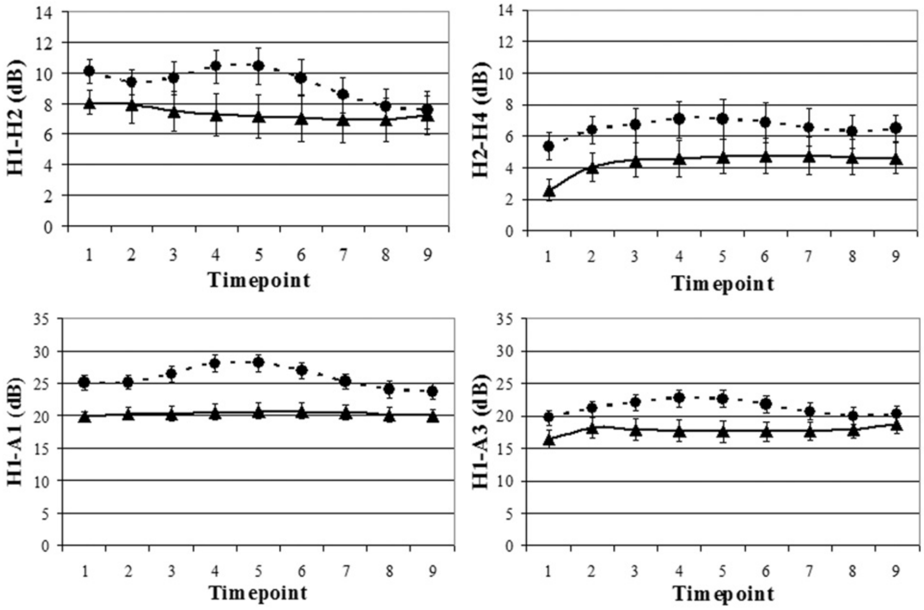

Fig. 5. Timecourse of H1–H2 (top left), H2–H4 (top right), H1–A1 (bottom left), and H1–A3 (bottom right) across the nine timepoints of modal (solid line, triangles) and breathy (dotted line, circles) vowels, averaged across vowel quality and speaker; error bars are standard error.

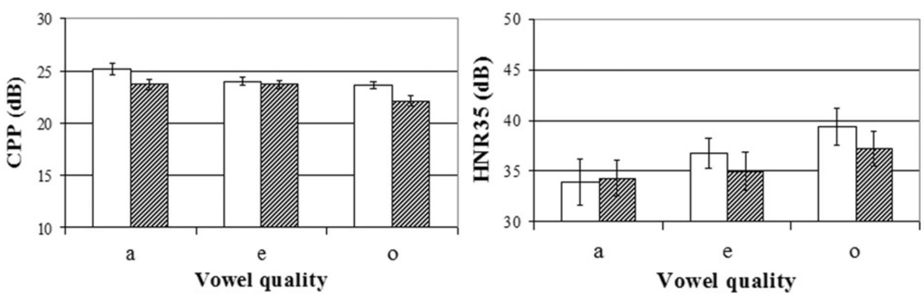

Fig. 6. Periodicity: CPP (left) and HNR of the 0–3500 Hz range (right) values of modal (solid) and breathy (striped) vowels, averaged across the entire vowel duration and speaker, separated by vowel quality.

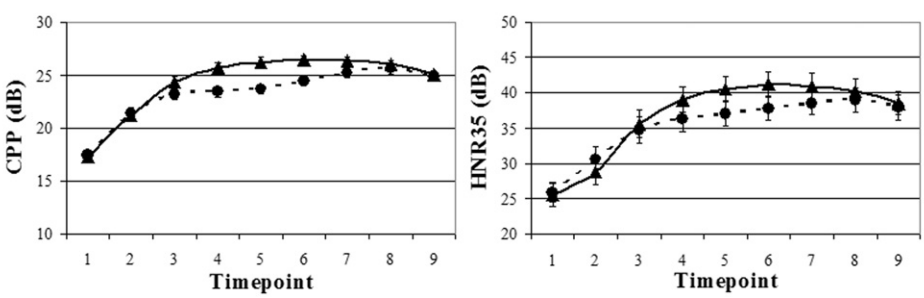

Fig. 7. Timecourse of CPP values (left) and HNR values within the 0–3500 Hz range (right) across the nine timepoints of modal (solid line, triangles) and breathy (dotted line, circles) vowels, averaged across vowel quality and speaker; error bars are standard error.

a fairly good cue to breathiness, although secondary to spectral measures such as H1–H2. Given this incongruity, the periodicity results were run through further scrutiny: an examination of the timecourse of periodicity values across the vowel's duration reveals that while breathy and modal vowels are not distinct at their onset and offset, they are distinct at their midpoint, as shown in Fig. 7.

Indeed, focusing on the middle third of the vowel (i.e. timepoints four, five, and six of nine) reveals statistically significant ($p<0.01$) differences between modal and breathy categories along the four measures of periodicity that show marginally significant ($0.01<p<0.05$) results in measurements averaging across the entire vowel. For example, at the fifth of nine timepoints, there are significant effects of phonation on HNR of the 0–1500 Hz [$F(1, 196)=17.78$; $p<0.01$], 0–2500 Hz [$F(1, 196)=24.00$; $p<0.01$], and 0–3500 Hz ranges [$F(1, 196)=25.23$; $p<0.01$], as well as CPP [$F(1, 196)=50.80$; $p<0.001$]. These last two measures are shown in Fig. 8. In the same position, there was a marginally significant effect of phonation on HNR of the 0–500 Hz range [$F(1, 196)=8.92$; $p=0.02$]. The midpoint of the vowel (here defined as timepoints 3, 4, and 5) is also the site of a significant interaction between phonation and vowel quality on CPP, e.g. at timepoint 4 [$F(2, 196)=12.216$; $p<0.001$]: this is the area of widest differentiation between

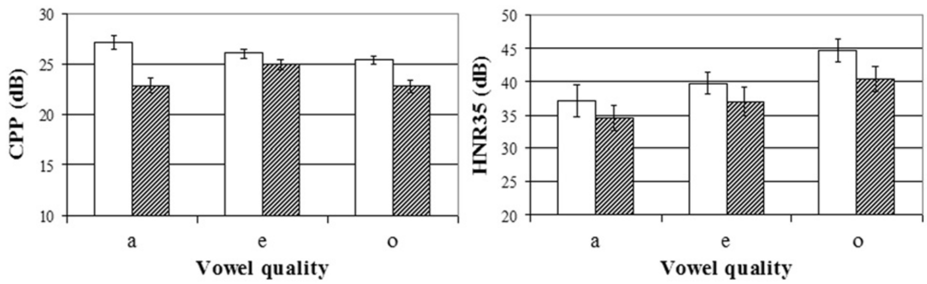

Fig. 8. Periodicity at midpoint: CPP (left) and HNR of the 0–3500 Hz (right) values of modal (solid) and breathy (striped) vowels at the fifth of nine timepoints, averaged across speakers, separated by vowel quality; error bars are standard error.

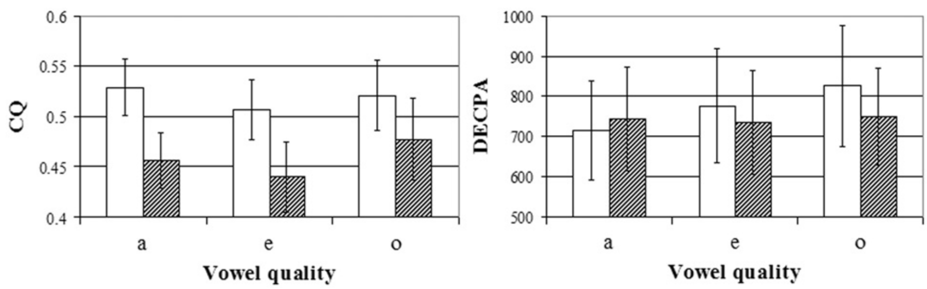

Fig. 9. Electroglottographic measures: CQ (left) and DECPA (right) values of modal (solid) and breathy (striped) vowels, averaged across the entire vowel duration and speaker, separated by vowel quality; error bars are standard error.

the members of the pairs [a a̤] and [o o̤], while the front vowel pair [e e̤] shows no such exaggeration in differentiation.[15]

4.1.3. Other acoustic measures

There was no significant effect of phonation on the mean RMS energy across the entire vowel [$F(1, 196)=2.54$; $p=0.15$], or at any particular timepoint. There was also no significant interaction with vowel quality. These results indicate that RMS energy is not a good measure of breathy phonation in Gujarati vowels produced in running speech. Similar to the RMS energy results, there was no significant effect of phonation on the mean f0 across the entire vowel [$F(1, 196)=0.16$; $p=0.70$], or at any particular timepoint. However, an unexpected pattern emerges when investigating each vowel quality separately: a significant interaction of phonation and vowel quality on mean f0 was found [$F(2, 196)=8.79$; $p=0.002$]. While [e̤] had a lower f0 value than its modal counterpart [e] (as expected, following the findings of Pandit, 1957; Fischer-Jørgensen, 1967 as well as predominant crosslinguistic patterns), breathy [a̤ o̤] had *higher* f0 values than their modal counterparts [a o]. It is unclear why the front vowel pair [e e̤] would work differently from the central and back vowel pairs [a a̤] and [o o̤]; this is particularly interesting considering it is indeed the front vowel pair [e e̤] that is poorly differentiated in CPP, the only other measure involving a significant interaction between phonation and vowel quality across adjacent timepoints. This interaction between phonation and vowel quality on f0 does not appear to be related to the interactions seen in Javanese vowels following slack-voiced stops (Thurgood, 2004), in which higher and lower vowels behave differently with respect to their interaction with phonation on f0. Because this latter study describes distinctions in vowel height, it does not provide a clear prediction for interactions of vowel frontness/backness with phonation on f0.

4.2. Electroglottographic results

Shifting focus to the EGG results (Fig. 9), the findings reveal a significant effect of phonation on the mean CQ across the entire vowel [$F(1, 195)=79.86$; $p<0.001$], with no significant interaction with vowel quality. Breathy vowels consistently had a lower mean CQ, indicating the glottis is more open in these vowels relative to modal vowels; this is consistent with cross-linguistic findings. DECPA, on the other hand, was not significantly affected by phonation type [$F(1, 195)=0.18$; $p=0.68$], and there was no significant interaction between phonation and vowel quality, at any specific timepoint or across the entire vowel duration. (Note that unlike CQ values, DECPA values are arranged on an arbitrary scale of unspecified units.)

The lack of a clear connection between phonation type and DECPA in Gujarati is unlike the situation in languages such as White Hmong (Esposito, 2010c; Esposito & Khan, 2010, 2012) and Southern Yi (Kuang, 2011), in which DECPA was found to be significantly higher (i.e. faster glottal closure) in breathier phonations. It may surprise the reader to know that DECPA was significantly higher in the breathier phonations of those languages, given that higher DECPA values reflect faster glottal closure while breathy phonations are traditionally thought of as having less abrupt glottal closure. As suggested in Keating et al. (2010, p. 93), the higher DECPA values are due to a principle of 'the further, the faster': as breathier phonation requires a greater degree of glottal opening, this can require the vocal folds to move more quickly between the (wider) open phase and closed phase. Esposito and Khan (2012) go on to suggest that as White Hmong is a lexical tone language, it is possible that this strategy of speeding up the transition is what makes it possible to lengthen the open phase (which allows for greater breathiness) without lengthening the entire pulse (thus maintaining the pitch). The same is presumably true for Southern Yi, another lexical tone language.

[15] The only other significant interactions between periodicity measures and vowel quality are restricted to one timepoint each, instead of spanning multiple adjacent timepoints, and are thus not taken to be meaningful. There is a significant interaction between vowel quality and phonation on HNR in the 0–1500 Hz range at timepoint 1 [$F(2, 196)=6.11$; $p<0.01$] and on HNR in the 0–2500 Hz range at timepoints 1 [$F(2, 196)=6.94$; $p<0.01$] and 9 [$F(2, 196)=6.85$; $p<0.01$].

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001

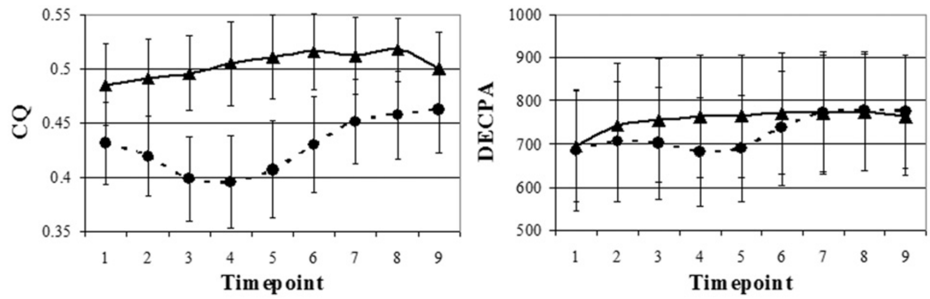

Fig. 10. Timecourse of mean CQ and DECPA values across the nine timepoints of modal (solid line, triangles) and breathy (dotted line, circles) vowels, averaged across speaker and vowel quality; error bars are standard error.

Although mean CQ values are significantly different across breathy and modal vowels, collapsing across timepoints, further investigation into their timecourse was conducted to see if there is any articulatory exaggeration at the voice source concentrated at the midpoint of breathy vowels to match the acoustic exaggeration in aperiodicity described above. As the graph in Fig. 10 illustrates, mean CQ values are indeed lowest (i.e. breathiest) around the third, fourth, and fifth of nine timepoints during breathy vowels, resembling the CPP effects seen during the fourth, fifth, and sixth timepoints and the HNR effects seen during the fifth and sixth timepoints. This behavior of exhibiting a significant difference between modal and breathy voice averaging across the entire vowel duration while also showing signs of exaggeration of the distinction in the mid-region strongly resembles that of the spectral measures, e.g. H2–H4. This region of lower CQ values indicates that the stronger breathiness seen at the vowel midpoint in the acoustic output is likely the result of an exaggerated articulation at the voice source, at least in terms of mean glottal opening. DECPA values, on the other hand, were not significantly different between breathy and modal vowels at any timepoint, although there was a non-significant exaggeration of the DECPA difference in the vowel's mid-region.

5. Discussion and conclusion

The current study reveals that Gujarati breathy vowels can be distinguished from their modal counterparts along numerous spectral measures. Corroborating the findings of previous studies, the current results find that breathy vowels are characterized by a significantly higher H1–H2, H1–A1, and H1–A2 averaging across their entire duration. Furthermore, due to the correction algorithms implemented in the current study to control for the effects of formants, a more accurate investigation of the spectrum was possible, revealing that H1–A3 was also significantly higher in breathy vowels. These spectral properties of breathiness hold true independent of vowel quality. Lastly, a relatively obscure spectral measure of voice quality was also found to distinguish breathy from modal vowels: H2–H4. Although most previous studies investigating the significance of the H2–H4 measure find that it is most closely associated with high pitch—including high lexical tone in Southern Yi (Kuang, 2011) and falsetto in English and Mandarin (Bishop & Keating, 2010)—the current study finds that this measure is also a reliable indicator of breathy voice. This is particularly striking in light of the fact that breathy vowels were not significantly distinct from their modal counterparts in terms of f0 in the current study (i.e. they did not have the high pitch normally associated with a high H2–H4 value), and in fact they have been associated with slightly *lower* f0 in early studies of the language. It appears that, despite its reliability in making various distinctions cross-linguistically, the significance of the H2–H4 measure remains unclear.

Four acoustic reflexes of noise were also found to be reliable measures of breathiness, and revealed a dynamic realization of breathy voice in Gujarati. CPP and the three highest measures of HNR examined in the current study were found to be lower in breathy vowels. The fact that the three highest measures of HNR investigated here (0–1500 Hz, 0–2500 Hz, and 0–3500 Hz) were significantly lower for breathy vowels, while the lowest measure of HNR (0–500 Hz) was not, indicates that the noise associated with breathy voice is not equally distributed across the spectrum; rather it is concentrated at frequencies above 500 Hz.

The findings of the various noise measures reveal a dynamic realization of breathy voice in Gujarati. The differences between breathy and modal voice in terms of noise are statistically significant only in the mid-region of the vowel—especially the fifth and sixth of the nine timepoints—indicating that this durational region is where breathy vowels are breathiest. Averaging across total vowel durations, none of the noise measures could distinguish breathy from modal vowels at the fairly standard threshold for statistical significance of $p < 0.01$. These results support at least one of two claims: (1) that while noise may be a characteristic of breathy vowels in Gujarati, its presence is not as reliable as other characteristics of the phonation type, and/or that (2) noise is the product of only the most exaggerated stretches of breathy voice, which occurs only vowel-medially. Both of these claims support previous acoustic and perceptual studies that found steeper spectral balance and spectral tilt to be more reliable cues to breathy phonation than the presence of aperiodic noise (Bickley, 1982; Esposito, 2010a; Fischer-Jørgensen, 1967), when averaging across total vowel duration.

Other acoustic measures, including RMS energy and f0, however, were entirely unsuccessful at distinguishing breathy vowels from their modal counterparts, at any timepoint. These results run counter to those established in early studies of Gujarati phonation, including Pandit (1957), Fischer-Jørgensen (1967), and Dave (1967), all of which reported a subtle and somewhat inconsistent lowering of f0 at the onset of breathy vowels. Because the experimental design of the current study differs from those of early studies in that it investigates more naturalistic speech, this finding may indicate that the slight differences in f0 (and possibly RMS energy) reported in early studies are restricted to highly enunciated laboratory-style speech, and/or are presumably overwhelmed by other factors in running speech, e.g., the effects of pragmatics and information structure on the intonation. As the f0 contour in Gujarati is largely a reflex of the intonational (as opposed to tonal or segmental) structure, the fact that the spelling pronunciation of breathy vowels involves a disyllabic pronunciation instead of a monosyllabic pronunciation would presumably alter the metrical structure of the word to the point that pitch accent placement would be affected, modifying the pitch and intensity contours.

In addition to the discovery of a large inventory of acoustic cues of Gujarati breathy voice, another important contribution of the current study is the EGG component, which serves as the first major articulatory investigation of breathy phonation in Gujarati. The

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001

EGG results determine that breathy vowels are characterized by a lower CQ throughout their duration than modal voice. This finding, combined with the spectral findings on H1–H2, strongly suggest that breathy voice in Gujarati is produced with a wider-than-modal glottis. Additionally, while the precise configuration of the glottis cannot be confirmed through acoustic and EGG measures alone, our current understanding of the H1–A1 results suggest that the arytenoids may be positioned in such a way to leave a gap for continuous airflow through the posterior region. What remains especially mysterious is the discrepancy regarding vocal fold closure; the current findings are not convergent in this regard. DECPA, considered the EGG measure of glottal speed, was *not* a reliable measure of breathiness, while the spectral measures considered the acoustic reflexes of the abruptness of vocal fold closure (i.e. H1–A2, H1–A3) showed statistically significant distinctions between breathy and modal vowels. It may be that DECPA is not a reliable indicator of vocal fold closure speed, or that the spectral measures are reflecting some other aspect of the glottal configuration. One limitation of EGG research, of course, is that while glottal impedance can be measured with a satisfactory degree of accuracy, one cannot isolate the precise location of the aperture (e.g. posterior, anterior). Visual observation of the pharynx and glottis will be necessary to measure glottal speed, as well as to confirm the exact configuration used to produce this voice quality. Given the wide variety of independently maneuverable mechanisms in the glottis, it is likely that more than one articulation is used within and across Gujarati speakers.

One notable strength of the current study is its methodology; unlike the scripted, purely laboratory-style speech recorded in previous studies, the current study examined more naturalistic running speech, finding that speakers produced consistent differences between contrastive voice quality categories even in this (relatively) more spontaneous speech style. By eliciting this speech style, the current study largely avoids the problems encountered in many previous studies of Gujarati, including effects of the "educated" speech register (neutralization of phonation contrasts to the modal category) and the read speech style (spelling pronunciation of breathy phonation as a sequence of modal vowels separated by /ɦ/).

As the first major analysis of corrected acoustic measures and electroglottographic measures in Gujarati, the current study supports the claim that breathy voice in Gujarati is produced with a wider-than-modal glottis, allowing for continuous, unimpeded airflow during voicing. This is evidenced by a large inventory of acoustic and articulatory parameters along which breathiness are reliably measured in the language, largely independent of vowel quality. The study also finds that breathy voice in Gujarati takes on a dynamic realization, with the most extreme breathiness produced at the midpoint; this is where four noise measures are statistically-significant markers of the nonmodal phonation, and where many other measures showed an exaggeration of the phonation distinction. Of course, despite these new findings, the current study naturally has some limitations that will need to be explored more fully in future work; the most noticeable of which is lack of perceptual data. While previous perceptual studies of Gujarati agree that H1–H2 is the most salient cue for Gujarati listeners, even when other cues were made available, these earlier studies did not cover the range of acoustic parameters presented in the current study, did not use algorithms for correction of spectral amplitudes, and did not manipulate any acoustic parameters by time. Consequently, it would be worthwhile to revisit this issue given these new findings.

Acknowledgments

I would like to thank (in alphabetical order) Christina M. Esposito, Ramesh Goud, Ranjan Goud, Patricia Keating, Claire Moore-Cantwell, Henry Tehrani, and Victoria Thatte for their valuable input; the ten subjects of the production study for their time and patience; and Kenneth de Jong, Michael Jessen, and two anonymous reviewers for their suggestions and insight. The study was funded by the NSF, as part of a collaborative project entitled "Production and Perception of Linguistic Voice Quality" (Investigators: Patricia Keating, Abeer Alwan, Jody Kreiman, & Christina Esposito. NSF Grant BCS-0720304: beginning 2007), which covered reimbursements for subject participation and the hiring of a research assistant to help with labeling and processing sound files.

Appendix A

See Table A1.

Appendix B

See Table B1.

Appendix C

See Table C1.

Table A1
Target vowels [a] vs. [a̤].

Order	Orthography	*Dictionary transcription*	IPA	Gloss
39	કાન	*kan*	kan	ear
15	કહાન	*(not listed)*	ka̤n	Krishna
10	બાર	*bar*	baɾ	twelve
02	બહાર	*bə.har*	ba̤ɾ	outside
42	બાણ	*baɳ*	baɳ	arrow
18	બહાનું	*bə.ha.nũ*	ba̤nũ	excuse
28	માલિક	*ma.lik*	malik	boss, god
14	મહારાજ	*mə.ha.raɟ*	ma̤ɾaɟʑ	priest, emperor

Table B1
Target vowels [e] vs. [ɛ̤].

Order	Orthography	*Dictionary transcription*	IPA	Gloss
24	કેવું	*ke.vũ*	keʋũ	of what kind
48	કહેવું	*kə.he.vũ*	kɛ̤ʋũ	to say
01	દેશી	*de.ʃi*	d̪eʃi	native, villager
44	દહેશત	*də.he.ʃət*	d̪ɛ̤ʃət̪	apprehension
56	પેલું	*pe.lũ*	pelũ	that
05	પહેલું	*pə.he.lũ*	pɛ̤lũ	first
70	બેરોજગાર	*be.roɟ.gar*	beɾoɟʑgaɾ	unemployed
67	બહેરાશ	*bə.he.raʃ*	bɛ̤ɾaʃ	deafness
57	મેં	*mɛ̃*	mɛ̃	I
51	મહેમાન	*mə.he.man*	mɛ̤man	guest
25	વેવાઇ	*ve.vai*	ʋeʋəj	father-in-law of one's child
13	વહેવાર	*və.he.var*	ʋɛ̤ʋaɾ	daily interaction
61	શેરડી	*ʃer.ɖi*	ʃeɾɖi	sugarcane
33	શહેર	*ʃə.her*	ʃɛ̤r	city

Table C1
Target vowels [o] vs. [o̤].

Order	Orthography	*Dictionary transcription*	IPA	Gloss
03	ડોળો	*ḍo.ḷo*	ɖoɭo	eyeball
30	ડહોળું	*ḍə.ho.ḷũ*	ɖo̤ɭũ	dirty, polluted
07	ડોલવું	*ḍol.vũ*	ɖolʋũ	to swing
22	ડહોળવું	*ḍə.hoḷ.vũ*	ɖo̤ɭʋũ	to stir in

Appendix D. Sound files for figures

⟨Fig1.wav⟩: Speaker 9, female, 30, from Mumbai
Target word ડહોળું [ɖo̤ɭũ] ‘polluted’
ડહોળું પાણી નદીમાં હોઈ
[ɖo̤ɭũ paɳi nəd̪i-mã hoi]
polluted water river-LOC there.is
‘There may be polluted water in the river’
⟨Fig2.wav⟩: Speaker 6, female, 29, from Mumbai
Target word મહેમાન [mɛ̤man] ‘guest’
મહેમાન તો ભગવાનનો રૂપ છે
[mɛ̤man t̪o bʱəgʋan-no ɾup cçʰe]
guest but god-GEN manifestation is
‘But a guest is a manifestation of god’
⟨Fig3a.wav⟩: Speaker 9, female, 30, from Mumbai
Target word શેરડી [ʃeɾɖi] ‘sugarcane’
શેરડીનો રસ મને બહુ ભાવે
[ʃeɾɖi-no ɾəs mə-ne bə̤u bʱaʋe]
sugarcane-GEN juice me-EXP much pleases
‘I like sugarcane juice a lot’
⟨Fig3b.wav⟩: Speaker 9, female, 30, from Mumbai
Target word શહેર [ʃɛ̤ɾ] ‘city’
શહેર મુંબઈ જેવું કોઈ નહી
[ʃɛ̤ɾ mumbəi ɟzeʋũ koi nə̤i]
city Mumbai of.which.kind any there.is.not
‘There’s no city like Mumbai’

Appendix E. Sound files illustrating a three-way breathiness contrast ([CV],[CV̤],[Cʱ V])

⟨Triplet5-modal.wav⟩: Speaker 5, female, 50, from Mumbai
Target word બાર [baɾ] ‘twelve’ cut from larger sentence
⟨Triplet5-breathy.wav⟩: Speaker 5, female, 50, from Mumbai
Target word બહાર [ba̤ɾ] ‘outside’ cut from larger sentence
⟨Triplet5-aspirated.wav⟩: Speaker 5, female, 50, from Mumbai
Target word ભાર [bʱaɾ] ‘burden’, ‘weight’ cut from larger sentence
⟨Triplet7-modal.wav⟩: Speaker 7, female, 23, from Mumbai
Target word બહાર [baɾ] ‘twelve’ cut from larger sentence
⟨Triplet7-breathy.wav⟩: Speaker 7, female, 23, from Mumbai
Target word ભાર [ba̤ɾ] ‘outside’ cut from larger sentence
⟨Triplet7-aspirated.wav⟩: Speaker 7, female, 23, from Mumbai
Target word બહાર [bʱaɾ] ‘burden’, ‘weight’ cut from larger sentence

Appendix F. Sound files illustrating a two-way breathiness contrast ([CV],[CV̤])

⟨Pair-modal.wav⟩: Speaker 10, male, 25, lived in various parts of India
Target word કેવું [keʋũ] ‘of what kind’ cut from larger sentence
⟨Pair-breathy.wav⟩: Speaker 10, male, 25, lived in various parts of India
Target word કહેવું [kɛ̤ʋũ] ‘to say’ cut from larger sentence

Appendix G. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.wocn.2012.07.001.

References

Abramson, A. S., Luangthongkum, T., & Nye, P. W. (2004). Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. *Phonetica*, *61*, 147–171.
Abramson, A. S., Nye, P. W., & Luangthongkum, T. (2007). Voice register in Khmu': Experiments in production and perception. *Phonetica*, *64*, 80–104.
Anastaplo, S., & Karnell, M. P. (1988). Synchronized videostroboscopic and electroglottographic examination of glottal opening. *Journal of the Acoustical Society of America*, *83*(5), 1883–1890.
Andruski, J. E., & Ratliff, M. (2000). Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association*, *30*(1/2), 37–61.
Baer, T., Löfqvist, A., & McGarr, N. S. (1983). Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques. *Journal of the Acoustical Society of America*, *73*(4), 1304–1308.
Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. New York: Singular, Thompson Learning.
Bickley, C. (1982). *Acoustic analysis and perception of breathy vowels. Speech Communication Group Working Papers* (pp. 73–93). Research Laboratory of Electronics, Cambridge, US: Massachusetts Institute of Technology.
Bishop, J., & Keating, P. (2010). Perception of pitch location within a speaker's own range: Fundamental frequency, voice quality and speaker sex. *UCLA Working Papers in Phonetics*, *108*, 113–140.
Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*, 163–191.
Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer*. Version 5.2.07. Free software. Available online: ⟨http://www.fon.hum.uva.nl/praat/⟩.
Cardona, G. (1965). *A Gujarati reference grammar*. Philadelphia: University of Pennsylvania Press.
Cardona, G., & Suthar, B. (2003). Gujarati. In: G. Cardona, & D. Jain (Eds.), *The Indo-Aryan languages* (pp. 659–697). Routledge.
Catford, J. C. (1977). *Fundamental problems in phonetics*. Bloomington: Indiana University Press.
Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, *90*(5), 2394–2410.
Childers, D. G., Naik, J. M., Larar, J. N., Krishnamurthy, A. K., & Moore, P. (1983). Electroglottography, speech and ultra-high speed cinematography. In: I. R. Titze, & R. C. Scherer (Eds.), *Vocal fold physiology: Biomechanics, acoustics and phonatory control* (pp. 202–220). Denver: Denver Center for the Performing Arts.
Choi, J.-Y., Hasegawa-Johnson, M., & Cole, J. (2005). Finding intonational boundaries using acoustic cues related to the voice source. *Journal of the Acoustical Society of America*, *118*(4), 2579–2587.
Dave, R. V. (1967). A formant analysis of the clear, nasalized, and murmured vowels in Gujarati. *Indian Linguistics*, *28*, 1–30.
Dave, R. V. (1977). *Studies in Gujarati phonology and phonetics*. Doctoral Dissertation. Cornell University.
Davies, P., Lindsey, G. A., Fuller, H., & Fourcia, A. J. (1986). Variation in glottal open and closed phase for speakers of English. *Proceedings of the Institute of Acoustics*, *8*, 539–546.
Denning, K. (1989). *The diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages*. Doctoral Dissertation. Stanford University.
DiCanio, C. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*(2), 162–188.
Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register, and stress: Laryngoscopic case studies. *Phonology*, *23*(2), 157–191.
Epstein, M. A. (2002). *Voice quality and prosody in English*. Doctoral Dissertation. Los Angeles: University of California.
Esling, J. H., & Harris, J. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In: W. J. Hardcastle, & J. Beck (Eds.), *A figure of speech: A Festschrift for John Laver* (pp. 347–383). Mahwah, NJ: Lawrence Erlbaum.
Esposito, C. M. (2010a). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, *38*(2), 303–316.
Esposito, C. M. (2010b). Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association*, *40*(2), 181–198.
Esposito, C. M. (2010c). *An acoustic and electroglottographic study of White Hmong tone and phonation*. Unpublished manuscript. Macalester College.

Esposito, C. M., & Khan, S. D. (2010). Breathiness contrasts across consonants and vowels: A comparative study of Gujarati and White Hmong. *Poster presented at the 2nd Pan-American/Iberian meeting on acoustics*. Cancún.
Esposito, C. M., & Khan, S. D. (2012). Contrastive breathiness across consonant sand vowels: Acomparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association*, *40*(2), 123–143.
Fant, G. (1972). Vocal tract wall effects, losses, and resonance bandwidths. *STL-quarterly progress and status report* (pp. 28–52).
Fant, G., Liljencrants, J., & Lin, Q. (1985). A four-parameter model of glottal flow. *KTH quarterly progress and status report*.
Firth, J. R. (1957). Phonetic observations on Gujarati. *Bulletin of the School of Oriental and African Studies*, *20*(1), 231–241.
Fischer-Jørgensen, E. (1967). Phonetic analysis of breathy (murmured) vowels. *Indian Linguistics*, *28*, 71–139.
Garellek, M., & Keating, P. (2010). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *UCLA Working Papers in Phonetics*, *108*, 143–163.
Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406.
Guion, S., Post, M., & Payne, D. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics*, *32*, 517–542.
Hanson, H. M. (1995). *Glottal characteristics of female speakers*. Doctoral Dissertation. Harvard University.
Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America*, *101*(1), 466–481.
Hanson, H. M., & Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America*, *106*(12), 1064–1077.
Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, *29*(4), 451–480.
Hayward, K. (1992). *Phonation types seen by fibreoptic laryngoscopy: Hausa, Korean & Gujarati (video)*. University of London, School of Oriental and African Studies.
Henrich, N. (2001). *Etude de la source glottique en voix parlée et chantée: Modélisation et estimation, mesures acoustiques et électroglottographiques, perception*. Doctoral Dissertation. Paris VI University.
Henrich, N., d'Alessandro, C., Doval, B., & Castellengo, M. (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, C. (2005). *A comparison of different methods to measure the EGG contact quotient*. Doctoral Dissertation. KTH Speech, Music, and Hearing.
Herbst, C., & Ternström, S. (2006). A comparison of different methods to measure the EGG contact quotient. *Logopedics Phoniatrics Vocology*, *31*(3), 126–138.
Hess, M. H., & Ludwigs, M. (2000). Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns. *Journal of Voice*, *14*(2), 255–271.
Hillenbrand, J. M., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, *37*, 769–778.
Holmberg, E., Hillman, R., Perkell, J., Guiod, P., & Goldman, S. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research*, *38*, 1212–1223.
Horii, Y. (1980). Vocal shimmer in sustained phonation. *Journal of Speech and Hearing Research*, *22*, 202–209.
Horii, Y. (1982). Jitter and shimmer differences among sustained vowel phonations. *Journal of Speech and Hearing Research*, *25*, 12–14.
Howard, D. M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice*, *9*(2), 163–172.
Huffman, M. K. (1985). Measures of phonation type in Hmong. *UCLA Working Papers in Phonetics*, *61*, 1–25.
Huffman, M. K. (1987). Measures of phonation type in Hmong. *Journal of the Acoustical Society of America*, *81*(2), 495–504.
Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America*, *121*, 2283–2295.
Jessen, M., & Roux, J. C. (2002). Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics*, *30*, 1–52.
Johnson, M. (1999). Tone and phonation in Western A-Hmao. *SOAS Working Papers in Linguistics*, *9*, 227–251.
Karnell, M. P. (1989). Synchronized videostroboscopy and electroglottography. *Journal of Voice*, *3*(1), 68–75.
Kawahara, H., de Cheveigne, A., & Patterson, R. D. (1998). An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised TEMPO in the STRAIGHT-suite. In *Proceedings of the 5th international conference on spoken language processing, Paper 0659*.
Keating, P., Esposito, C. M., Garellek, M., Khan, S. D., & Kuang, J. (2010). Phonation contrasts across languages. *UCLA Working Papers in Phonetics*, *108*, 188–202.
Khan, S. D. (2009). An acoustic and electroglottographic study of breathy phonation in Gujarati. *Poster presented at the 158th meeting of the Acoustical Society of America*. San Antonio.
Khan, S. D. (2010). Breathy phonation in Gujarati: An acoustic and electroglottographic study. *Poster presented at the 159th meeting of the Acoustical Society of America*. Baltimore.
Khan, S. D., & Esposito, C. M. (2011). Breathiness contrasts in consonants and vowels: A comparative study of Gujarati and White Hmong. *Poster presented at the 17th international congress of phonetic sciences*. Hong Kong.
Khan, S. D., & Thatte, V. (2009). The phonetics of contrastive phonation in Gujarati. *Poster presented at the 157th meeting of the Acoustical Society of America*. Portland.
Kirk, P. L., Ladefoged, J., & Ladefoged, P. (1993). Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In: A. Mattina, & T. Montler (Eds.), *American Indian linguistics and ethnography in honor of Laurence C*. Thompson, Missoula, MT: University of Montana Press.
Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*(2), 820–857.
Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. (2006). *Analysis and synthesis of pathological voice quality*. Available at 〈http://www.surgery.medsch.ucla.edu/glottalaffairs/files/GASoftwareManual2006.pdf〉.
Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research*, *50*, 595–610.
Kreiman, J., Gerratt, B. R., & Khan, S. D. (2010). Effects of native language on perception of voice quality. *Journal of Phonetics*, *38*(4), 588–593.
de Krom, G. (1993). A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *Journal of Speech and Hearing Research*, *36*, 224–266.
Kuang, J. (2011). *Production and perception of the phonation contrast in Yi*. Master's Thesis. Los Angeles: University of California.
Ladefoged, P. (1983). The linguistic use of different phonation types. In: D. Bless, & J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues* (pp. 351–360). San Diego: College-Hill Press.
Ladefoged, P., & Antoñanzas-Barroso, N. (1985). Computer measures of breathy phonation. *UCLA Working Papers in Phonetics*, *61*, 79–86.
Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Massachusetts: Blackwell Publishing.
Laver, J. (1980). *The phonetic description of voice quality*. Cambridge University Press.
Lewis, M. P. (2009). *Ethnologue: Languages of the World (16th edition)*. Dallas: SIL International. http://www.ethnologue.com/.
Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, *35*, 344–353.
Madiesson, I., & Hess, S. (1987). The effect of F0 on linguistics use of phonation types. *UCLA Working Papers in Phonetics*, *67*, 112–118.
Masica, C. P. (1993). *The Indo-Aryan languages*. Cambridge University Press.
Mazaudon, M., & Michaud, A. (2008). Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica*, *65*(4), 231–256.
Michaud, A. (2004). A Measurement from Electroglottography: DECPA, and its application in prosody. In *Proceedings of the speech prosody 2004* (pp. 633–636). Nara, Japan.
Miller, A. L. (2007). Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics*, *35*, 56–84.
Mistry, P. J. (1997). Gujarati phonology. In: A. S. Kaye (Ed.), *Phonologies of Asia and Africa*. Winona Lake: Eisenbrauns.
Modi, B. (1987). Rethinking of "murmur in Gujarati". *Indian Linguistics*, *47*, 39–55.
Nair, U. (1979). *Gujarati phonetic reader*. Mysore: Central Institute of Indian Languages.
Orlikoff, R. 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, P. B. (1957). Nasalization, Aspiration and Murmur in Gujarati. *Indian Linguistics*, *17*, 165–172.
Pandit, P. B. (1961). Historical Phonology of Gujarati Vowels. *Language*, *37*(1), 54–66.
Pandit, P. B. (1969). *Gujarati. Current trends in linguistics, Vol. 5: Linguistics in South Asia*. Mouton de Gruyter.
Pennington, M. (2005). *The phonetics and phonology of glottal manner features*. Doctoral Dissertation. Indiana University.
Pfitzinger, H. R. (2008). Segmental effects on the prosody of voice quality. *Proceedings of Acoustics*, *4*, 3159–3164.
Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, *11*(4), 478–504.
Rami, M. K., Kalinowski, J., Stuart, A., & Rastatter, M. P. (1999). Voice onset times and burst frequencies of four velar stop consonants in Gujarati. *Journal of the Acoustical Society of America*, *106*(6), 3736–3738.
Ren, N. (1992). *Phonation types and stop consonant distinctions: Shanghai Chinese*. Doctoral Dissertation. University of Connecticut.
Samely, U. (1991). *Kedang (Eastern Indonesia): Some aspects of its grammar*. Hamburg: Helmut Buske Verlag.
Scherer, R. C., Druker, D. G., & Titze, I. R. (1988). Electroglottography and direct measurement of vocal fold contact area. In: O. Fujimura (Ed.), *Vocal physiology: Voice production, mechanisms and functions* (pp. 279–291). New York: Raven.
Shue, Y.-L., Keating, P., & Vicenik, C. (2009). VoiceSauce: A program for voice analysis. *Journal of the Acoustical Society America*, *124*(4), 2221.
Stevens, K. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34*(4), 264–279.

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001

Suthar, B. (2003). *Gujarati-English Learner's dictionary (2nd draft)*. Philadelphia: A Nirman Foundation Project, University of Pennsylvania.

Tehrani, H. (2009). *EGGWorks: A program for automated analysis of EGG signals.* 〈http://www.linguistics.ucla.edu/faciliti/facilities/physiology/EggWorksSetup.exe〉.

(Luang-) Thongkum, T. (1987a). Another look at the register distinction in Mon. *UCLA Working Papers in Phonetics*, *67*, 132–165.

(Luang-) Thongkum, T. (1987b). Phonation types in Mon-Khmer languages. *UCLA Working Papers in Phonetics*, *67*, 29–48.

Thurgood, E. (2004). Phonation types in Javanese. *Oceanic Linguistics*, *43*(2), 277–295.

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

Turner, R. L. (1921). Gujarati phonology. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, *3*, 329–365.

Vũ-Ngọc, T., d'Alessandro, C., & Michaud, A. (2005). Using open quotient for the characterization of Vietnamese glottalized tones. In *Proceedings of the international conference. Eurospeech-Interspeech 2005: 9th European conference on speech communication and technology* (pp. 2885–2889). Lisbon.

Vyas, Y. D. (1977). *ગુજરાતી ભાષાનું વ્યાકરણ [Gujarātī Bhāṣānuṃ Vyākaraṇa: A grammar of the Gujarati language]*. Ahmedabad: Sāhitya Mudraālaya for the Gujarati University.

Wackernagel, J. (1896). *Altindische Grammatik, Teil I: Lautlehre*. Göttingen: Vandenhoeck und Ruprecht.

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

Wayland, R., Gargash, S., & Jongman, A. (1994). Acoustic and perceptual investigation of breathy voice. *Journal of the Acoustical Society of America*, *97*(5), 3364.

Wayland, R., & Jongman, A. (2003). Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics*, *31*(2), 181–201.

Whitney, W. D. (1889). *Sanskrit Grammar.* Republished 1964. Cambridge: Harvard University Press.

Please cite this article as: Khan, S.D., The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* (2012), http://dx.doi.org/10.1016/j.wocn.2012.07.001