*INTERSPEECH 2017*
August 20–24, 2017, Stockholm, Sweden

# Acoustic and electroglottographic study of breathy and modal vowels as produced by heritage and native Gujarati speakers

*Kiranpreet Nara*

University of Toronto, Canada
kiranpreet.nara@mail.utoronto.ca

## Abstract

While all languages of the world use modal phonation, many also rely on other phonation types such as breathy or creaky voice. For example, Gujarati, an Indo-Aryan language, makes a distinction between breathy and modal phonation among consonants and vowels: /bʱaɾ/ 'burden', /baɾ/ 'twelve', and /ba̤ɾ/ 'outside' [1, 2]. This study, which is a replication and an extension of Khan [3], aims to determine the acoustic and articulatory parameters that distinguish breathy and modal vowels. The participants of this study are heritage and native Gujarati speakers.

The materials consisted of 40 target words with the modal and breathy pairs of the three vowel qualities: /a/ vs /a̤/, /e/ vs /e̤/, and /o/ vs /o̤/. The participants uttered the words in the context of a sentence. Acoustic measurements such as H1-H2, H1-A1, harmonic-to-noise ratio and articulatory measurements such as contact quotient were calculated throughout the vowel duration.

The results of the Smoothing Spline ANOVA analyses indicated that measures such as H1-A1, harmonic to noise ratio, and contact quotient distinguished modal and breathy vowels for native speakers. Heritage speakers also had a contrast between breathy and modal vowels, however the contrast is not as robust as that of native speakers.

**Index Terms**: Gujarati, breathy vowels, phonation, heritage speakers, electroglottography

## 1. Introduction

Voice qualities such as creaky or breathy phonation are used in addition to modal phonation in many languages of the world. For example, Jalapa Mazatec, an Otomanguean language, has a breathy, modal, and creaky phonation contrast among its vowels: /ja̤/ 'he wears', /já/ 'tree', and /ja̰/ 'he carries' (Ladefoged & Maddieson, 1996), while Gujarati, an Indo-Aryan language that is spoken in western India, phonemically contrasts breathy and modal phonation in both consonants and vowels: /bʱaɾ/ 'burden', /ba̤ɾ/ 'outside' and /baɾ/ 'twelve' [1, 2]. The focus of the current paper is to study the various acoustic and electroglottographic (EGG) characteristics that differentiate breathy and modal vowels in Gujarati. While a similar study has been conducted by Khan [3], a novel feature of the present research is studying two different groups of Gujarati speakers: native speakers and heritage speakers. Native Gujarati speakers are those who were born in India, learned Gujarati as their first language, and arrived in Canada between 2013-2016. Heritage speakers are those people who were either born in Canada or arrived in Canada before the age of seven. The heritage speakers learned Gujarati as their first language but their main language of communication is English.

Heritage language speakers are bilinguals who have had a limited access to their first language and because of this, they might not acquire certain features of their first language [4]. For example, it has been shown that heritage speakers of Western Armenian produce more English-like vowel qualities when speaking the language compared to native Western Armenian speakers [5]. Other researchers, however, have found no difference between heritage and native speakers in terms of production. For example, native and heritage Spanish speakers have shown no difference in their voice onset time productions [6]. The present experiment will help to determine how similar or different heritage Gujarati speakers are from native speakers in terms of their breathy and modal vowel productions.

There are eight modal vowels in standard Gujarati, which is spoken from the city of Baroda to the city of Ahmedabad in Gujarat, India: /i e ɛ a ɔ o u ə/. Each of these vowels has a breathy counterpart, however, the breathy counterparts of /e/ and /o/ are /ɛ̤/ and /ɔ̤/, respectively [1, 3]. Gujarati breathy vowels arose due to phonological processes occurring in certain environments involving breathy consonants and modal vowels [1, 2, 7]. An environment in which the most stable breathy vowels, /ɛ̤ a̤ ɔ̤ ə̤/, emerged is /əɦV/ [2, 8]. For example, the historical /bəɦen/ 'sister' is now pronounced /bɛ̤n/.

### 1.1. Characteristics of breathy voice

Breathy voice is defined as having a laryngeal setting in which the vocal folds do not come into as much contact during the closed phase of vocal fold vibration as they do during modal voice [9, 10]. This is due to the vocal folds not coming into complete contact during the closed phase of the vibratory cycle during breathy phonation [11, 12].

#### *1.1.1. Acoustic parameters*

Various parameters distinguish breathy and modal phonation. One such measure is subtracting the amplitude of the second harmonic (H2) from the amplitude of the first harmonic: H1-H2. The difference between H1 and H2 is higher for breathy phonation than for modal phonation [7, 13, 14], therefore, the measure H1-H2 should be higher for breathy vowels than modal vowels. H1-H2 has been shown to distinguish breathy and modal phonations in Green Mong [15], Gujarati [7], and Jalapa Mazatec [16]. Spectral tilt, which is the degree to which intensity drops off with increasing frequency, is more steeply negative for breathy vowels than for modal vowels [17]. The measures used to determine spectral tilt are H1-A1, H1-A2, and H1-A3 [18]. These three measures differentiate phonation types in language such as Jalapa Mazatec [19], Krathing Chong [16], !Xóõ [14], and Gujarati [3]. Phonation type also affects the measure H2-H4 in languages such as Chong [20] and Gujarati [3].



http://dx.doi.org/10.21437/Interspeech.2017-1774

Breathy phonation is associated with an increase in aperiodic noise at higher frequencies [13]. Measures of noise and periodicity are harmonic to noise ratio (HNR) and cepstral peak prominence (CPP), respectively. These two measures should be lower for breathy phonation than for modal phonation. For example, HNR has been found to be lower in breathy phonation in Javanese [21] and lower CPP for breathy phonation has been found in Krathing Chong [16] and Southern Yi [22]. Breathiness is also associated with lower fundamental frequency (F0) than modal voice [23, 12] in languages such as Nyah Kur and Kui [24], and Green Mong [15].

#### *1.1.2. Electroglottographic parameters*

EGG is a non-invasive method for measuring the contact between the vocal folds, and the contact quotient (CQ) is therefore one of the several possible measures afforded by EGG. CQ is the portion of time for which the vocal folds are in maximum contact during one period of the glottal cycle [25]. For example, breathy vowels tend to have a smaller contact quotient than modal vowels and this has been observed in languages such as Yi [22], White Hmong [20], and Gujarati [3]. Breathy phonation also has faster speed of vocal fold closing and this is measured using Derivative-EGG Closure Peak Amplitude (DECPA) [26, 3]. This measure is also called Peak Increase in Contact (PIC) in [26]. The faster the speed of vocal folds during glottal closure, the higher the DECPA value, and breathy phonation should have a smaller DECPA value than modal phonation. DECPA has been observed to differentiate modal from nonmodal phonation in languages such as Southern Yi [22] and White Hmong [20].

### 1.2. Gujarati literature

Previous work on Gujarati phonation has determined that breathy phonation has higher random acoustic energy at higher frequencies. Fischer-Jørgensen [7] has shown that the following parameters distinguish breathy and modal phonation in Gujarati: H1-H2, H1-A1, H1-A3, and lower CQ. Khan [3] has corroborated these findings and included H2-H4 and H1-A2, into the list of measures that are significantly different for the two phonation types [3].

### 1.3. Questions

This study aims to answer the following questions:

1. What acoustic and EGG parameters distinguish breathy vowels from modal vowels? Are they the same parameters identified in [3] and [7]?

2. Are the differences between the two phonation types localized to specific parts of the vowel, as determined in [3]?

3. Will the heritage Gujarati speakers' productions be like those of native speakers where the breathy and modal vowels are clearly distinguished or will the heritage speakers produce breathy and modal vowels that are more alike one another.

## 2. Methodology

### 2.1. Participants

Thirteen native speakers and six heritage speakers took part in the study. All speakers were recruited from Toronto, Canada. All native and heritage speakers were literate in Gujarati orthography though the heritage speakers preferred to translate the target word from English. The mean age of native speakers was 36.4 years and that of heritage speakers was 24 years.

### 2.2. Materials

The materials consisted of 40 words which contained 9 breathy and modal word pairs for the vowel quality /a/, 9 pairs for /e/, and 2 pairs for /o/. Table 1 presents examples of words pairs for the three vowel qualities, providing the IPA, gloss, and Gujarati orthography of the words. The target vowel was the first vowel in the word following an onset consonant. The remaining vowels and all consonants in the target word were unaspirated to avoid having more than one breathy segment in the word.

Table 1: *Examples of target stimuli.*

| Vowel Quality | Breathy | Modal |
|---|---|---|
| /a/ | [mə̤n̤] 'great'<br>મહાન | [man̤] 'respect'<br>માન |
| /e/ | [kɛ̤ʊu] 'say, speak'<br>કહેવું | [keʊu] 'of what kind'<br>કેવું |
| /o/ | [mɔ̤ɾ] 'stamp'<br>મહોર | [moɾ] 'peacock'<br>મોર |

### 2.3. Equipment and procedure

Articulatory data were collected using a (Glottal Enterprises electroglottograph (Model EG2-PCX2). Two 35 millimetre dual-channel electrodes measured the electrical impedance and the Behringer ECM-8000 omnidirectional microphone was used to record the audio signal at the sampling rate of 44000 Hz. The microphone and the EGG were connected using an XLR cable and a USB cord was used to connect the EGG to the computer.

The participants completed a questionnaire asking for their own and their family's language background. Following that, the participants were seated in a sound attenuating booth while EGG electrodes were secured to their neck. A microphone was placed about six to eight centimetres away from the mouth of the participant. The target words, written in Gujarati orthography as well as their English translation, were presented one at a time on a computer screen. The participants were asked to come up with a sentence containing the target word and say the sentence five times. This was done to keep the participant focused on the meaning of the words instead of the spelling so that the target words sound as they would in natural speech. The orthography might lead to the participants producing the target words without breathiness to sound more 'educated' [3]. The experiment took place in the Phonetics Laboratory at the University of Toronto.

### 2.4. Analysis

The acoustic and EGG recordings were made and annotated in Praat [27]. VoiceSauce [28], a free software that makes automated measurements from audio files, was used to analyze the following acoustic parameters: H1-H2, H2-H4, H1-A1, H1-A2, H1-A3, fundamental frequency (F0), Harmonic-to-Noise ratio (HNR), and Cepstral Peak Prominence (CPP). EGGWorks [29], also a free software, was used to analyze the EGG measures such as contact quotient and DECPA. Each vowel was divided into nine sections to plot the measurements across time.

Repeated measures (RM) ANOVA analyses were conducted at each of the nine intervals for each vowel to stay consistent with and to compare the results of the present experiment to those of Khan's [3]. In addition, Smoothing

Spline ANOVA (SS-ANOVA) analyses were conducted to examine global differences between the modal and breathy vowels over time. These two analyses were used because RM ANOVA, in addition to staying consistent with Khan [3], informs on the difference between the two phonation types at each interval whereas the SS-ANOVA presents an overall picture of the differences across the entire vowel duration. This paper only presents the analyses of the SS-ANOVAs. All statistical analyses were done using R (R core team, 2016).

## 3. Results

The smoothing splines function in the SS-ANOVA connects discrete data points and creates curves or 'splines'. SS-ANOVA determines if two or more splines are significantly different from one another (Davidson, 2006). SS-ANOVA has been used to analyze differences between tongue shapes (Davidson, 2006) as well as to analyze formant contours of vowels (De deker, 2006). SS-ANOVA also calculates 95% Bayesian confidence intervals, which are represented as curves above and below the main spline. Two splines are deemed significantly different from one another if the curves of their confidence intervals do not intersect (Davidson).

The results of SS-ANOVA, summarized in Table 2, showed that the native speakers had more parameters differentiating breathy and modal vowels than heritage speakers. The differences remained for longer durations for native speakers and were localized to specific parts of the vowel for both groups.

Table 2: *Results of SS-ANOVAs showing intervals with significant differences as well as the direction of difference (B = Breathy; M = Modal).*

| Measure | Native speakers | Heritage speakers |
|---|---|---|
| H1-H2 | - | - |
| H2-H4 | 3-9 (B > M) | - |
| H1-A1 | 2-7 (B > M) | - |
| H1-A2 | 2-8 (B > M) | 4-7 (B > M) |
| H1-A3 | - | - |
| HNR (0-3500 Hz) | 4-8 (B < M) | - |
| CPP | 3-8 (B < M) | 4-7 (B < M) |
| F0 | - | - |
| CQ | 1-7 (B < M) | 3-6 (B < M) |
| DECPA | - | - |

### 3.1. Acoustic results

For native speakers, breathy and modal phonation were significantly different for H2-H4 from intervals 3 to 9, with the magnitude of the difference increasing from interval 3 to interval 9. H2-H4 did not differ between breathy and modal vowels for heritage speakers. H1-A1 was significantly different from intervals 2 to 7 for the two phonation types for native speakers, but there was no difference for heritage speakers. H1-A2 was significantly different from interval 2 to interval 8 for native speakers. However, for heritage speakers, the significant difference lay from intervals 4 to 7. The magnitude of difference for heritage speakers was much less than that of native speakers. Parameters H1-H2 and H1-A3 did not distinguish breathy and modal phonation for either native or heritage speakers.

HNR, measured from 0 Hz to 3500 Hz, was significantly different from intervals 4 to 8 for native speakers while there was no difference for heritage speakers. CPP distinguished breathy and modal vowels from intervals 3 to 8 for native speakers and from intervals 4 to 7 for heritage speakers. The magnitude of the difference was higher for native speakers. Finally, F0 did not distinguish phonation type for either speaker group.

### 3.2. EGG results

CQ of breathy vowels was significantly lower than that of modal vowels from intervals 2 to 7 for native speakers and from intervals 4 to 5 for heritage speakers (see Figures 1 and 2). The difference between breathy and modal vowels was much higher in magnitude for native speakers. Finally, DECPA did not distinguish phonation types for either speaker group.

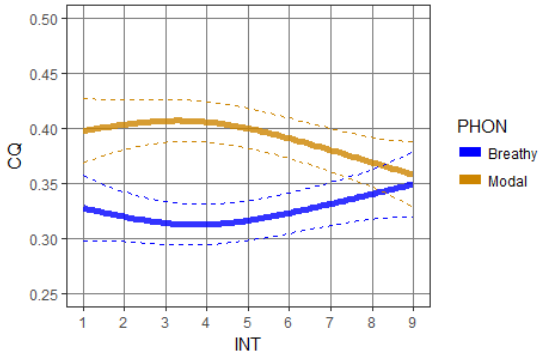


Figure 1: *Contact quotient of native speakers.*

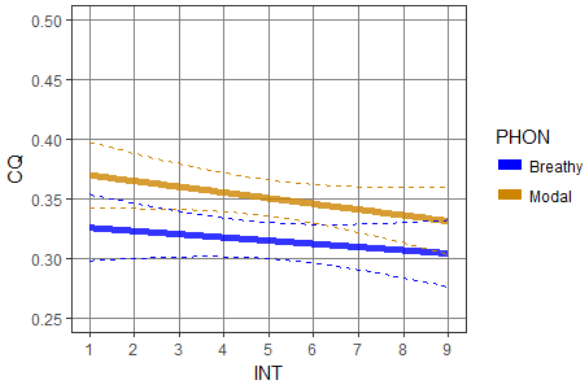


Figure 2: *Contact quotient of heritage speakers.*

## 4. Discussion

The results of the present experiment indicated that the acoustic and electroglottographic parameters that differentiate breathy and modal vowels in native Gujarati speakers were H2-H4, H1-A1, H1-A2, harmonic to noise ratio, cepstral peak prominence, and contact quotient.

H1-H2, which many researchers have found to be an important measure in distinguishing breathy and modal phonations in languages such as Green Mong [15] and Gujarati [3, 7], did not distinguish phonation type in the present study. While it varied in the expected direction, with breathy vowels having higher H1-H2 than modal vowels, the difference was not statistically significant.

Measures such as H1-A1 and H1-A2 indicated the intensity of breathy vowels at higher frequencies relative to the

amplitude of the first harmonic was much lower than that of modal vowels, as also found in [18, 19]. Khan [3] also showed that H1-A3 distinguished breathy and modal vowels, however these findings were not replicated in the present study. HNR and CPP were lower for breathy vowels than for modal vowels and this indicated that native speakers' breathy vowels were noisier and more aperiodic than modal vowels. Similar results for HNR and CPP were seen in Khan's [3] Gujarati study as well as for Javanese [21] and Krathing Chong [16].

F0 did not show a significant difference between breathy and modal vowels in either speaker group, corroborating Khan's [3] results. The results of the contact quotient for the present experiment were like those obtained for Gujarati [3, 7], as well as for Yi [22], and White Hmong [20].

The results also showed that the differences between modal and breathy vowels were localized to specific parts of the vowel for many parameters as also determined by Khan [3]. The localization happened in the middle or the latter two thirds of the vowel for many measures such as H2-H4, H1-A2, H1-A2, HNR, and CPP. For CQ, phonation types were distinguished within the first two-thirds of the vowel duration.

For heritage speakers, not only were there fewer parameters that significantly distinguished breathy and modal vowels, H1-A2, CPP, and CQ, but the difference occurred for a shorter duration and smaller magnitude. The results indicated that while heritage Gujarati speakers acquired the contrast between modal and breathy vowels, the difference was reduced in terms of duration and magnitude compared to native speakers. The results of heritage speakers in the current study are like those of the study of Western Armenian where heritage speakers produced more English-like vowel qualities.

## 5. Conclusions

The acoustic and EGG parameters that distinguish breathy and modal phonation are different for native and heritage speakers. The differences tend to be localized to specific parts of the vowel. Finally, heritage speakers have similar pronunciations of breathy and modal vowels compared to native speakers, indicating limited acquisition of heritage language sound production.

Next steps for this project are studying the phonation productions of different proficiencies of heritage speakers as well as how the perception of phonation differs between heritage speakers and native speakers.

## 6. Acknowledgements

I would like to thank Dr. Alexei Kochetov for his guidance throughout this project. The use of EGG in this study was supported by a SSHRC Grant (#435-2015-2013) to Dr. Kochetov. This study was also supported by a SSHRC Institutional Grant (SIG), 2016-2017.

## 7. References

[1] G. Cardona, *A Gujarati reference grammar.* Philadelphia: The University of Pennsylvania Press, 1965.

[2] P. B. Pandit, "Nasalisation, aspiration and murmur in Gujarati," *Indian Linguistics,* vol. 17, pp. 165-172, 1957.

[3] S. D. Khan, "The phonetics of contrastive phonation in Gujarati," *Journal of Phonetics*, vol. 40, no. 6, pp. 780-795, 2012.

[4] M. Polinsky, "Cross-linguistic parallels in language loss," *Southwest Journal of Linguistics,* vol. 14, pp. 87-123, 1995.

[5] L. Godson, "Vowel production in the speech of Western Armenian heritage speakers," *Heritage Language Journal*, vol. 2, pp. 1-26, 2004.

[6] T. K. F. Au, L. M. Knightly, S. A. Jun, and J. S. Oh, "Overhearing a language during childhood," *Psychological Science*, vol. 13 no. 3, pp. 238-243, 2002.

[7] E. Fischer-Jørgensen, "Phonetic analysis of breathy (murmured) vowels in Gujarati," *Indian Linguistics*, vol. 28, pp. 71-139, 1967.

[8] G. Cardona and B. Suthar, *Gujarati: The Indo-Aryan languages*, ed. by G. Cardona and D. Jain, pp. 659-697, 2003.

[9] G. Fairbanks (Ed.), *Voice and articulation: drillbook*. Harper & Brothers, 1940.

[10] W. R. Zemlin, Speech and Hearing Science, Anatomy and Physiology, 1968.

[11] J. C. Catford, *Fundamental problems in phonetics*, Midland Books, 1977.

[12] J. Laver, "The phonetic description of voice quality," *Cambridge Studies in Linguistics London*, vol. 31, pp. 1-186, 1980.

[13] D. H. Klatt, and L. C. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *The Journal of the Acoustical Society of America*, vol. 87, no. 2, pp. 820-857, 1990.

[14] P. Ladefoged, "The linguistic use of different phonation types," *Vocal fold physiology: Contemporary research and clinical issues*, pp. 351-360, 1983.

[15] J. E. Andruski, and M. Ratliff, "Phonation types in production of phonological tone: The case of Green Mong," *Journal of the*

[16] *International Phonetic Association*, vol. 30, no.1-2, pp. 37-61, 1983.

[17] B. Blankenship, "The timing of nonmodal phonation in vowels," *Journal of Phonetics* vol. 30, pp. 163–191, 2002.

[18] M. Gordon, and P. Ladefoged, "Phonation types: a cross-linguistic overview," *Journal of Phonetics*, vol. 29, no. 4, pp. 383-406, 2001.

[19] N. Henrich, C. d'Alessandro, B. Doval, and M. Castellengo, "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation," *The Journal of the Acoustical Society of America* vol. 115, no. 3, pp. 1321-1332, 2004.

[20] M. Garellek, and P. Keating, "The acoustic consequences of phonation and tone interactions in Jalapa Mazatec," *Journal of the International Phonetic Association*, vol. 41, no. 2, pp. 185-205, 2011.

[21] C. M. Esposito, "An acoustic and electroglottographic study of White Hmong tone and phonation," *Journal of Phonetics*, vol. 40, no. 3, pp. 466-476, 2012.

[22] R. Wayland, S. Gargash, and A. Jongman, "Acoustic and perceptual investigation of breathy voice," *Journal of the Acoustical Society of America*, vol. 97, no. 5, pp. 3364, 1994.

[23] J. Kuang, "Production and perception of the phonation contrast in Yi," *Ms., University of California, Los Angeles*, 2011.

[24] J. M. Hombert, J. J. Ohala, and W. G. Ewan, "Phonetic explanations for the development of tones," *Language*, pp. 37-58, 1979.

[25] T. Luang-Thongkum, "Phonation types in Mon-Khmer languages," *UCLA Working Papers in Phonetics*, vol. 67, pp. 29–48, 1987.

[26] R. J. Baken, "Electroglottography," *Journal of Voice*, vol. 6, no. 2, pp. 98-110, 1992.

[27] P. Keating, C. Esposito, M. Garellek, S. Khan, and J. Kuang, "Phonation contrasts across languages," in *Poster presented at the 12th Conference on Laboratory Phonology*, 2010.

[28] P. Boersma, and D. Weenink, *Praat: doing phonetics by computer [Computer program]*,Version 5.3.51, Retrieved 2 June 2013 from <http://www.praat.org/>, 2010.

[29] Y.-L. Shue, P. Keating, and C. Vicenik, "VoiceSauce: A program for voice analysis," *Journal of the Acoustical Society America*, vol. 124, no. 4, pp. 2221, 2009.

[30] H. Tehrani, *EGGWorks: A program for automated analysis of EGG signals*. Retrieved July 10, 2016 from <http://www.linguistics.ucla.edu/faciliti/facilities/physiology/EggWorksSetup.exe>, 2009.

[31] A. S. Abramson, P. W. Nye, and T. Luang-Thongkum, “Voice register in Khmu’: Experiments in production and perception,” *Phonetica*, vol. 64, pp. 80–104, 2007.