Breathy phonation in Gujarati: an acoustic and electroglottographic study

Sameer ud Dowla Khan, UCLA
sameeruddowlakhan@gmail.com Poster 5aSC25

Background

Breathy phonation
Breathy voice typically involves vibrating vocal folds spread further apart than in modal voice (L71, GL01):

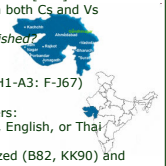
Open glottis — **Voiced** — **Closed glottis**

Voiceless — **Breathy** — Lax — **Modal** — Tense — **Creaky** — **?**

Breathy phonation in Gujarati
Gujarati vowels contrast breathy vs. modal (F-J67):

Modal [bar] 'twelve' vs. breathy [ba̤r] 'outside'

Breathy aspiration is also contrastive on consonants [bʰar] 'burden'
One of the few lgs. to contrast breathy voice on both Cs and Vs



How can breathy and modal vowels be distinguished?
Acoustic studies of Gujarati breathy vowels:
- Larger open quotient (H1-H2: F-J67, B82)
- Sharply falling spectral tilt (H1-A1, H1-A2, H1-A3: F-J67)
- Lower intensity (RMS energy: F-J67)

Perception studies show that Gujarati speakers:
- Are more sensitive to H1-H2 than Mandarin, English, or Thai speakers (KGK10)
- Use only H1-H2 when categorizing synthesized (B82, KK90) and foreign Vs (E06)

But what's going on at the source?
No previous electroglottographic (EGG) study of Gujarati

Research questions

1. **What acoustic cues distinguish breathy vs. modal vowels?**
   Expected: H1-H2, H1-A3, RMS energy
   Additional cues?
2. **What EGG cues distinguish breathy vs. modal vowels?**
   First EGG study of Gujarati
3. **Are some effects restricted to portions of the vowel?**
   Beginning, middle, end?
4. **How closely do acoustic properties match EGG properties**?
   H1-H2 often correlated with CQ, H1-A3 with PIC

Methods

Speakers
10 native Gujarati speakers; all but two in US <1 yr:
Spkrs 1, 3, 10: males, ages 21–25
Spkrs 2, 4, 6, 7, 8, 9: females, ages 21–30
Spkr 5: female, age 50

Stimuli
18 words in (near)-minimal pairs of [a] vs. [a̤]

બહાર
outside

Recording
Audio and EGG signals recorded simultaneously
Spkrs shown flashcard with stimulus in orthography, plus English gloss
Produced a sentence of their own creation, starting w/ stim, as many times as possible in 10s, producing **fast, fluent, naturalistic speech**

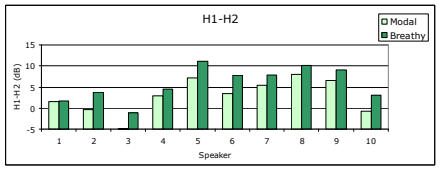
Analysis
"VoiceSauce" (SKV09) took >100 acoustic measurements per target V
"EGGWorks" took similar measurements from the EGG signal
Averaged in every 1/9 of the target vowel duration ("timeslices")

Measures
For each speaker, values were averaged across "context"
Grouped by preceding C type: voiced stop, voiceless, or sonorant
Regressions compared means of [a] vs. [a̤] with respect to:
- **Four acoustic measures:**
  - H1-H2 (uncorrected)
  - H1-A3 (uncorrected)
  - RMS energy
  - Cepstral peak prominence (CPP)
- **Two EGG measures:**
  - Closing quotient (CQ)
  - Peak increase in closure (PIC)

Acoustic results: H1-H2

Background
H1-H2 is the difference in amplitude of first two harmonics
Acoustic correlate of **open quotient** (OQ) (H&C99, HSKCS01, HPGG95, F95, F97, SH95, H97)
Breathy vowels have higher H1-H2 cross-linguistically
!Xóõ (B82); Jalapa Mazatec, Chong (B01); Chanthaburi Khmer (WJ02); Chong, Fuzhou, Green Mong, White Hmong, 3 Zapotec lgs., !Xóõ, (E06); Tsonga (L&A-B85), Wa (W02)

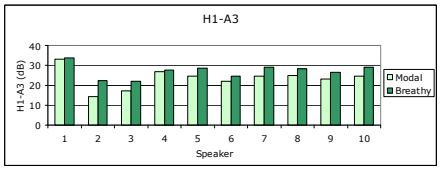
Current findings
Breathy vowels showed a higher H1-H2 than modals
β=2.88dB, s.e.=1.01, t=2.85, p<0.01
Supports that **breathy Vs have a more open glottis than modals**
Supports previous data showing the **salience of H1-H2 in Gujarati**



Acoustic results: H1-A3

Background
H1-A3 is a measure of **spectral tilt** (S&H95, H97)
Breathy Vs differ in their H1-A3 value cross-linguistically
Higher H1-A3 in Mon, 3 Zapotec lgs. (E06); Chong (DC09)
Lower H1-A3 in Chong, Wh.Hmong, Gr.Mong, Tamang, !Xóõ (E06)

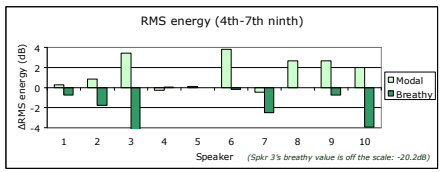
Current findings
Breathy vowels showed higher H1-A3 than modals
β=4.01, s.e.=1.21, t=-3.32, p<0.01



Acoustic results: RMS energy

Background
RMS energy is correlated with **overall intensity**
Breathy Vs differ in their RMS energy levels cross-linguistically
Higher energy in Javanese (WGJ94), Chanthaburi Khmer (WJ02).
Lower energy in Kui and Chong (T88)

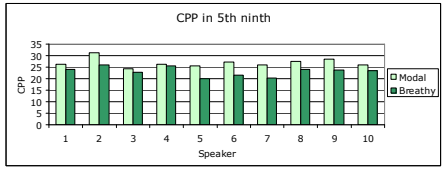
Current findings
Overall RMS energy does not distinguish breathy vs. modal
Contrary to previous work on Gujarati (F-J67)
However, there is a **dynamic difference in RMS energy**
Energy generally rises during breathy vowels
Energy generally falls during modal vowels
Greater energy difference between 4th-7th ninths in breathy Vs
β=5.90dB, s.e.=2.04, t=2.88, p<0.01



Acoustic results: Cepstral Peak Prominence (CPP)

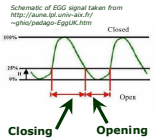
Background
CPP is a measure of **periodicity**
Breathy Vs have lower CPP (less periodicity) cross-linguistically
Chong, Fuzhou, (H)mong, Mon, 3 Zapotec lgs., Tamang, !Xóõ, (E06)

Current findings
Overall CPP does not distinguish breathy vs. modal
However, there is a difference at the midpoint of the vowel
CPP dips to a local minimum halfway through breathy Vs
CPP remains high in modal Vs
Breathy Vs are less periodic than modal Vs in the 5th ninth
β=3.31, s.e.=0.67, t=-5.83, p<0.01
Waveform at the midpoint of a breathy V has **more irregular cycles**



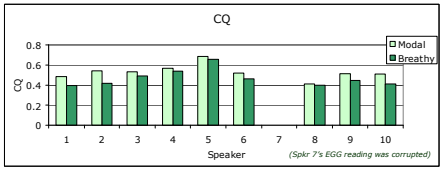
EGG results: Closing quotient (CQ)

Background
Glottal cycle = opening phase + closing phase
CQ = proportion of cycle in closing phase
Lower CQ = more open = more airflow
Breathy has lower CQ in SADV Zapotec (E04)



CQ = Closed / (Closed + Open)

Current findings
Breathy vowels had lower CQ than modals
β=0.06, s.e.=0.02, t=-2.93, p<0.01
Suggests wider glottis, more airflow in breathy



EGG results: Peak increase in contact (PIC)

Background
PIC = peak positive value in the first derivative of EGG signal (DEGG)
Equivalent to DECPA (M04)
Correlated with H1-A3 and speed of vocal fold closure (EPY09)
Breathy Vs have higher PIC in White Hmong (EPY09), Yi (K10)
Related measure shows similar results in SADV Zapotec males (E05)

Current findings
PIC does not distinguish breathy vs. modal
Unexpected, given that **correlated H1-A3 does distinguish them**

Summary of findings

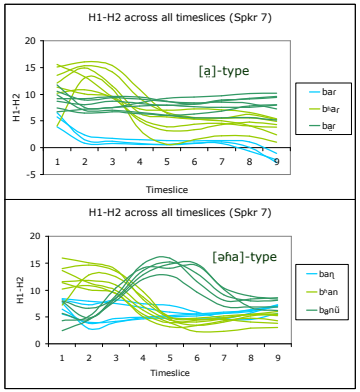
1. Two spectral measures distinguish breathy & modal
   H1-H2 and H1-A3 are both higher in breathy Vs
2. One EGG measure also distinguishes the two:
   Closing quotient (CQ) is lower in breathy Vs
3. Other cues were restricted to the middle of the V:
   Greater intensity rise across middle of breathy Vs
   Less periodicity at midpoint of breathy Vs
4. Unexpectedly, **PIC was not correlated with H1-A3**

Further questions

Breathy consonant release
How is [ba̤r] 'outside' distinguished from [bʰar] 'burden' acoustically?
Prediction: timing of cues will distinguish the two (see fig. below)

Non-low vowels
Is breathiness marked with the same cues despite effects of V quality?

Variant productions
Gujarati breathy Vs are variably pronounced [V̤]~[VɦV]~[V] (D67)
Continuum of production types seen in current data (see fig. below)
Is the distribution of these productions predictable?
Are the same cues significant across production types?



Acknowledgments

This study is part of a collaborative NSF-supported project entitled "Production and Perception of Linguistic Voce Quality" (Keating, et al., beg. 2007), exploring aspects of phonation in various languages.

Special thanks go to Ramesh Goud, Ranjan Goud, Patricia Keating, Claire Moore-Cantwell, Kevin Terrance Petway II, Henry Tehrani, Kristine Yu, and the ten subjects of the study.

References

Bickley C (1982). 'Acoustic analysis and perception of breathy vowels'. *Speech communication group working papers*: 71–82. Cambridge: MIT.

Blankenship B (2001). 'The timing of nonmodal phonation in vowels'. *JPhon* **30** (2): 163-191.

Dave, R (1967). 'A formant analysis of the clear, nasalized, and murmured vowels in Gujarati'. *Indian Ling* **28**: 1-30.

DiCanio, CT (2009). 'The phonetics of register in Takhian Thong Chong'. *JIPA* **39**: 162-188.

Esposito CM (2004). 'Santa Ana del Valle Zapotec Phonation'. *UCLA WPP* **103**.

Esposito CM (2005). 'An Acoustic and Electroglottographic Study of Phonation in Santa Ana del Valle Zapotec'. LSA79.

Esposito CM (2006). 'The Effects of Linguistic Experience on the Perception of Phonation'. PhD diss., UCLA.

Esposito CM, Ptacek J, & Yang S (2009). 'An acoustic and electroglottographic study of Wh. Hmong phonation'. ASA158.

Fant G (1995). 'The LF-model revisited. Transformations and frequency domain analysis'. STL-QPSR **2-3**: 119-156.

Fant G (1997). 'The voice source in connected speech'. *Speech Communication* **22**: 125-139.

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

Gordon M & Ladefoged P (2001). 'Phonation types: a cross-linguistic overview'. *JPhon* **29**: 383–406.

Hanson HM (1997). 'Glottal characteristics of female speakers: Acoustic correlates'. *Acoustics*.

Hanson HM & Chuang E (1999). 'Glottal characteristics of male speakers: acoustic correlates and comparison with female data'. *JASA* **106**: 1697-1714.

Hanson HM, Stevens KN, Kuo J, Chen M, & Slifka J (2001). 'Towards models of phonation'. *JPhon* **29**: 451-480.

Holmberg EB, Hillman RE, Perkell JS, Guiod PC, & Goldman SL (1995). 'Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice'. *Speech Hearing Research* **38**: 1212-1223.

Klatt DH & Klatt C (1990). 'Analysis, synthesis, and perception of voice quality variations among female and male talkers'. *JASA* **87** (2): 820–857.

Kreiman J, Gerratt BR, & Khan SD (2010). 'Effects of native language on perception of voice quality'. Ms.

Kreiman J, Iseli M, Neubauer J, Shue YL, Gerratt BR, & Alwan A (2008). 'The relationship between open quotient and H1*-H2*'. *JASA* **124** (4): 2495.

Kuang JJ (2010). 'An acoustic and electroglottographic study of phonation contrast in Yi'. Poster at ASA 159.

Michaud A (2004). 'A measurement from Electroglottography: DECPA, and its Application in Prosody'. *Proceedings of Speech Prosody 2004*.

Ladefoged P (1971). *Preliminaries to linguistic phonetics*. U Chicago.

Ladefoged P & Antoñanzas-Barroso N (1985). 'Computer measures of breathy phonation'. *UCLA WPP* **61**: 79–86.

Stevens, KN & Hanson HM (1995). 'Classification of glottal vibration from acoustic measurements'. In Fujimura, O & Hirano M (eds.), *Vocal Fold Physiology: Voice quality control*. Singular Publishing Group: 147-170.

Shue YL, Keating P, & Vicenik C (2009). 'VoiceSauce: A program for voice analysis'. Poster at ASA158.

Thongkum T (1988). 'Phonation types in Mon-Khmer languages'. In: Fujimura, O., (ed.). *Vocal fold physiology: voice production, mechanisms and functions*, Raven Press, New York: 319–334.

Watkins J (2002). 'The phonetics of Wa: Experimental phonetics, phonology, orthography, and sociolinguistics'. *Pacific Linguistics* **531**. Canberra: ANU.

Wayland R, Gargash S, & Jongman A (1994). 'Acoustic and perceptual investigation of breathy voice'. *JASA* **97**: 3364.

Wayland R & Jongman A (2002). 'Acoustic correlates of breathy and clear vowels: the case of Khmer'. *JPhon* **31** (2).