Electroglottographic and acoustic measures of phonation across languages

Patricia Keating and Jianjing Kuang
UCLA Linguistics Department
Phonation contrasts in languages of the world

- Many of the world’s languages use phonation contrastively on vowels and/or consonants – a different phonation makes a different word.
- Common especially in SE Asia, the Americas, India.
- Audio example next slide.
Jalapa Mazatec (Mexico)
Low tone, male speakers

\[ /\textit{jæ}^1 / \text{(modal)} - \text{(Engl. } \textit{boil} \text{ (noun)}) \]

\[ /\textit{jæ}^1 / \text{(creaky)} - \text{(Engl. } \textit{manure}) \]

\[ /\textit{jæ}^1 / \text{(breathy)} - \text{(Engl. } \textit{boil} \text{ (verb)}) \]
Relation of phonation to lexical tone in languages

- Some languages with phonation contrasts do not have lexical tone (pitch) contrasts.
- Some languages have both phonation and tone contrasts, independently, such that different tones and phonations can co-occur.
- Some languages use phonation as part of the tonal system: certain tones have their own correlated phonations.
How are phonation contrasts produced?

- Not really clear yet - direct observation of such laryngeal activity is very limited to date, often not practical

- **Electroglottography (EGG)** is a non-invasive, though indirect, way of comparing glottal differences among contrastive phonations – EGG indirectly indexes vocal fold contact
This talk

- Relate EGG to acoustics in two phonation languages
- Suggest advantages of studying phonation in languages where it’s contrastive:
  - Speakers share goals, i.e. the language’s phonological categories [Ladefoged]
  - Likely to see a wide range of values on phonation measures, so any relations among them are likely to be clear
Languages we have EGG recordings from

- **Hmong** (White Hmong, Laos) [with Christina Esposito]
  - 1 lexical tone is Breathy, 1 Creaky, others modal
- **Yi** (Yunnan province, China, Southern dialect)
  - Lax vs. Tense voice, crossed with Low and Mid lexical tones
- **Bo** (Yunnan province, China)
- **Hani** (Yunnan province, China)
- **Black Miao** (Guizhou province, China)
- **Gujarati** (Standard Gujarati) [with Sameer Khan]
- **Mandarin** (Standard Beijing) [with Kristine Yu]
- **Zapotec languages** (Santiago Matatlán, San Juan Guelavia, Santa Ana del Valle) [with Christina Esposito]
Yi fieldwork in Yunnan
Hmong fieldwork in Minnesota
[by Christina Esposito]
Hmong EGG example, Creaky vs. Breathy

Creaky: \(p\_21\), “see”

Breathy: \(p\_43\), “grandmother”

1 female speaker
1 rep each word

more contact

less contact
UCLA analysis tools

- **EggWorks**  
  (Tehrani 2009)

- **VoiceSauce**  
  (Shue 2010, Shue et al. 2011)

- Free by downloading
EGG measures

“relative contact duration”: Contact Quotient CQ
4 methods from EGG signal

peak velocities of contact Increase and Decrease from dEGG

“contact symmetry”: closing duration / opening duration from EGG signal
**EGG results: Quotients Contact (left) and Skew (right)**

**Hmong:** 8 male speakers
- CQ and SQ pattern similarly (inversely), distinguish Breathy from Creaky, Modal phonations

**Yi:** 3 male speakers
- CQ, SQ can distinguish Lax vs. Tense phonations

(5 time intervals)
Peak Increase (left) and Decrease (right) in Contact

**Hmong**: 8 male speakers
- PIC and PDC pattern similarly (inversely), distinguish all phonations, especially at vowel-end

**Yi**: 3 male speakers
- PIC and PDC distinguish Lax vs. Tense phonations (inversely)
Contact Quotient inversely related to rates of change

- Greatest rates of change are in Breathy voice, which has lowest CQ values
- Moderate correlations across speakers, better within speakers
- Possibly related to amplitude change within a pulse: “the further the faster” (next slide)

1 Hmong speaker: PkIncrCont vs. CQ

\[ R^2 = .80 \]
Hmong Breathy vs. Creaky example (EGG=black, dEGG=blue)

Breathy: faster larger

Creaky: slower smaller
Lower ContactQ with faster contacting

- Breathy phonation has lower ContactQ and greater rates of change, but also more gradual closing as seen in high speed imaging of glottal area (e.g. Shue 2010)

- Thus it appears that peak rate of contacting from EGG is not the same as abruptness of closing in glottal pulse
In Hmong, F0 cannot predict any EGG parameters above $R^2 = .08$, either across phonations or just in Modal.

In Yi, F0 accounts for ~20% variance in PeakIncreaseCont, PeakDecreaseCont, and contact rise time: higher F0 has faster, shorter increase in contact and slower decrease in contact.

Especially in Lax phonation.
Functional Data Analysis of Yi glottal pulse shapes

- An alternative to traditional measures (Ramsay & Silverman 1997/2002; Mooshammer 2010) = functional version of principal component analysis (FPCA) using the R package FDA version 1.2.4
- Pairs of pulses extracted from Yi vowels with Tense and Lax phonation types and with Low and Mid tones (3 males)
- Pulses time-normalized 0-1000 and amplitude-normalized 0-1 (next slide)
Pulses before and after amplitude normalization
1\textsuperscript{st} two principal components for Yi tense/lax pulses (87% of variance)

- contacting phase:
  - varies mostly with phonation type, not with tone

- maximum contacting:
  - varies with phonation type, but mostly for Low tone

(3\textsuperscript{rd} principal component varies with tone, not phonation type; 4\textsuperscript{th} is minor, more about individual speaker differences)
Relation (r) of 4 PCs to standard EGG measures

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ContactQ_Threshold</td>
<td>.9</td>
<td>.09</td>
<td>.13</td>
<td>.33</td>
</tr>
<tr>
<td>ContactQ_Hybrid</td>
<td>.81</td>
<td>.24</td>
<td>0</td>
<td>.01</td>
</tr>
<tr>
<td>PkIncreaseContact</td>
<td>-.77</td>
<td>.03</td>
<td>-.12</td>
<td>-.19</td>
</tr>
<tr>
<td>PkDecreaseContact</td>
<td>.91</td>
<td>-.11</td>
<td>.10</td>
<td>-.16</td>
</tr>
<tr>
<td>SkewQ</td>
<td>.06</td>
<td>-.23</td>
<td>.28</td>
<td>.66</td>
</tr>
</tbody>
</table>

*weaker*
Summary of EGG

- EGG measures generally distinguish the phonation types; are not strongly related to F0
- Peak Decrease in Contact (neg peak in dEGG), not a standard measure, is very distinctive here
- Peak changes in contact perhaps related to pulses as “the further the faster”
- Most variation in Yi EGG pulse shape is related to the phonation types, and mostly in terms of the shape of the contact increase and peak
- In Yi, EGG pulse shape is most strongly related to Contact Quotient and to Peak Decrease in Contact
Acoustic correlates

- Many acoustic measures distinguish 2 or even 3 phonation types

- H1*-H2*, shown here, does so across languages: H1*-A2* is another very distinctive measure
Relations of EGG and acoustic measures

Questions of interest: re H1*-H2*

- Given uncertain relation of OQ (in flow or area) to H1-H2 – how does CQ pattern?
- Given the robustness of H1-H2 as a phonation type measure, what does it reflect physiologically?
From CQ to H1*-H2*:
languages differ

**HMONG:** $R^2 = 0.56$

**YI:** $R^2 = 0.20$

(R^2 increases to 0.30 when only CQs from 0.4 to 0.6 are included)
This relation of H1*-H2* to CQ in Hmong can be very strong for individual speakers: here, 1 male, a larger dataset

\[ R^2 = .76 \]
From Peak Decrease in Contact to H1*-H2*

HMONG: $R^2 = .40$

YI: $R^2 = .27$
From Peak Increase in Contact to H1*-H2*

**HMONG:** $R^2 = .22$

**YI:** $R^2 = .07$
From Skew Quotient to H1*-H2*

HMONG: $R^2 = .17$

YI: $R^2 = .18$
From FDA Principal Components to acoustic measures (in Yi)

- 1\textsuperscript{st} principal component is most strongly related to $H1^*-H2^*$ ($r=-.7$)
- 2\textsuperscript{nd} principal component is less strongly related to $H1^*-H2^*$ ($r=-.48$); also to bandwidth of F2 ($r=-.5$)
Conclusions - 1

What do we learn from EGG about these languages’ phonation categories?

- In Yi, **Contact Quotient** is the most distinctive EGG measure, both directly and by its strong relation to those principal components of pulse shapes that relate to phonation.

- In Hmong, the two rate-of-change EGG measures (**Peak Increase in Contact**, **Peak Decrease in Contact**) are most distinctive.
Conclusions - 2

What do we learn from EGG about H1-H2, especially re Contact Quotient?

- H1*-H2* is correlated at least modestly with all the EGG measures (even ones we didn’t present here), and with PC1 and PC2 of Yi EGG pulse shape, suggesting it’s related to many aspects of pulse shape and timing.

- In Hmong, H1*-H2* is most strongly related to CQ. In Yi, PC1 of pulse shape is related to both CQ and H1*-H2*, but these measures are not strongly related to each other.
What do we learn from Functional Data Analysis of EGG pulse shape in Yi?

- (Not so important for Hmong, where some standard EGG measures already work well)

- But in Yi, no EGG measures account for much variance in acoustic measures – standard EGG doesn’t tell us much

- Yet in Yi, PC1 and PC2 are related to H1*-H2*, and to the phonation contrast – here, the shape of the contacting part of pulse is crucial, which only FDA could tell us.
Acknowledgments

- NSF grant BCS-0720304, and co-PIs Abeer Alwan, Jody Kreiman
- NSF grant IIS-1018863, PI Alwan
- Y.-L. Shue for VoiceSauce
- H. Tehrani for EggWorks
- Collaborators Christina Esposito, Marc Garellek, Sameer Khan
Extra slide – all 4 Yi PCs

PCA function 1 (Percentage of variability 66.6)

PCA function 2 (Percentage of variability 19.2)

PCA function 3 (Percentage of variability 6.2)

PCA function 4 (Percentage of variability 2)
Extra slide: Yi audio

- bə²¹
- bə²¹
- bə²¹