An acoustic and electroglottographic study of White Hmong phonation

Christina M. Esposito
Macalester College

Background: Hmong

- Hmong-Mien language family
- Spoken in and around northern Vietnam, Thailand, and Laos
- Large population of speakers in Minnesota and California
- Several dialects (White Hmong, Green Mong, Blue Mong, etc.) with various levels of mutual intelligibility
- Total number of speakers worldwide = 4 million

Background: Tones and Phonations

- Seven tones and three phonations:

<table>
<thead>
<tr>
<th>Traditional description</th>
<th>Ratliff (1993)</th>
<th>Example in IPA</th>
<th>Example in White Hmong orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>55</td>
<td>[pɔ]</td>
<td>pob’ball’</td>
</tr>
<tr>
<td>Mid</td>
<td>33</td>
<td>[pɔ]</td>
<td>po’spleen’</td>
</tr>
<tr>
<td>Low</td>
<td>22</td>
<td>[pɔ]</td>
<td>pos’thorn’</td>
</tr>
<tr>
<td>High-falling</td>
<td>52</td>
<td>[pɔ]</td>
<td>poj ‘female’</td>
</tr>
<tr>
<td>Mid-rising</td>
<td>24</td>
<td>[pɔ]</td>
<td>pov ‘to throw’</td>
</tr>
<tr>
<td>Low-falling creaky</td>
<td>21 (creaky)</td>
<td>[pɔ]</td>
<td>pom ‘to see’</td>
</tr>
<tr>
<td>Mid-low breathy</td>
<td>42 (breathy)</td>
<td>[pɔ]</td>
<td>pog’grandmother’</td>
</tr>
</tbody>
</table>

* = difference between the traditional description and Ratliff’s description
*There is an eighth tone, which is a syntactic variant of the low-falling creaky.
*All tones are produced with modal phonation unless otherwise noted.

Previous Research: Huffman (1987)

- Examined tones with either modal or breathy phonation in a variety of Hmong dialects:
  - Measures of spectral tilt and flow pulse symmetry were not reliable measures of phonation
  - Measures of spectral balance (specifically H1-H2) and closed-phase duration (measured from glottal flow) successfully distinguished modal from breathy phonation

Previous Research: Ratliff (1992)

- Studied tonal morphology in White Hmong
  - -g tone (42, breathy)
    - For male speakers, the –g tone is more accurately described as 31.
    - For female speakers, the –g tone is more accurately described as 53 (a high fall).
  - -m tone (21, creaky)
    - Shorter than the other tones/phonations

Background: Hmong orthography

- Tone in the Hmong orthography is indicated by one of the following symbols at the end of a syllable:

<table>
<thead>
<tr>
<th>Tone description (Traditional/Ratliff)</th>
<th>Orthographic tone symbol</th>
<th>Example in White Hmong orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>High/55</td>
<td>h</td>
<td>pob’ball’</td>
</tr>
<tr>
<td>Mid/33</td>
<td>ø</td>
<td>po’spleen’</td>
</tr>
<tr>
<td>Low/22</td>
<td>s</td>
<td>pos’thorn’</td>
</tr>
<tr>
<td>High-falling/52</td>
<td>j</td>
<td>poj ‘female’</td>
</tr>
<tr>
<td>Mid-rising/24</td>
<td>v</td>
<td>pov ‘to throw’</td>
</tr>
<tr>
<td>Low-falling/creaky:21</td>
<td>m</td>
<td>pom ‘to see’</td>
</tr>
<tr>
<td>Mid-low breathy/42</td>
<td>g</td>
<td>pog’grandmother’</td>
</tr>
</tbody>
</table>
Previous Research: Andruski and Ratliff (2000)

- Examined three tones with similar f0 contours, but different phonations (modal, breathy, creaky) in Green Mong
- Results:
  - Tones are distinguished by their f0 values.
  - H1-H2 was successful in distinguishing breathy from non-breathy phonation.
    - There was an effect of vowel quality. H1-H2 values were lower for breathy productions of /i/ and /u/ than they were for /a/.
  - No significant main effect or interaction for gender.

Results:

- Tones are distinguished by their f0 values.
- H1-H2 was successful in distinguishing breathy from non-breathy phonation.
  - There was an effect of vowel quality. H1-H2 values were lower for breathy productions of /i/ and /u/ than they were for /a/.
- No significant main effect or interaction for gender.

Current Study

- Examines the phonations and tones in White Hmong
  1) What is the most accurate description of tone?
  2) What are the acoustic and electroglossotographic correlates of phonation?
  3) How does phonation co-vary with f0?
  4) Are there durational differences between the three phonations?
  5) Are there gender differences in the production of phonation?

Methods: Speakers

- Speakers
  - Thirty-three native speakers of White Hmong were recorded.
  - Speakers:
    - were born in either Laos (17 speakers), Thailand (13) or the US (3)
    - ranged from 18-48 years of age
    - currently reside in the Twin Cities, MN
    - Spoke English in addition to White Hmong
    - The age of English onset was 5-35 years of age
    - Three speakers also spoke Green Mong natively in addition to White Hmong and English.

Methods: Speech Materials

- Speakers produced words with all seven tones/three phonation types for a total of 73 words.
  - All words were monosyllabic and contained one of six White Hmong vowels [i, e, ɨ, a, u, ɔ]
- Words were uttered in the frame [ɾɔ hai ____ duɔ] “Say ____ again”

Procedure

- For twenty-one speakers, audio recordings were made.
- For twelve speakers, simultaneous electroglossotographic (EGG) and audio recordings were made.
- Acoustic and EGG measures were
  - taken automatically using VoiceSauce
  - and/or PeQuireX

Procedure: Acoustic Measures

- f0
- CPP (Cepstral peak prominence)
- H1* (Amplitude of the first harmonic)
- H2* (Amplitude of the second harmonic)
- H1* - H2* (Amplitude of the first harmonic minus the amplitude of the second harmonic)
- H1* - A1* (Amplitude of the first harmonic minus the amplitude of the first formant peak)
- H1* - A2* (Amplitude of the first harmonic minus the amplitude of the second formant peak)
- H1* - A3* (Amplitude of the first harmonic minus the amplitude of the third formant peak)
- H2* - H4* (Amplitude of the first harmonic minus the amplitude of the fourth harmonic)
- * = harmonic amplitudes are reported with corrections for formant frequencies and bandwidths
Procedure: EGG Measures

- **Closed quotient (CQ)**
  - reflects the portion of time the vocal folds are closed during each glottal cycle
  - \[ \text{Closed quotient} = \frac{T_c}{(T_c + T_o)} \]

- **Peak closing-velocity**
  - reflects the speed of vocal fold closure

Results: Tones, Males

Results showed that:
1. The high/55 –b tone is more accurately described as rising.
2. While traditionally described as mid-low, the –g tone is more accurately described as falling, as Ratliff (1992) noted.

Results: Tones, Females

Results showed that:
1. The high/55 –b tone is more accurately described as rising.
2. While traditionally described as mid-low, the –g tone is more accurately described as high falling for the female speakers.

Results: Acoustic Measures

Results showed that none of the acoustic measures tested distinguished all three phonation types on a given time point.

Several measures distinguished two categories.

<table>
<thead>
<tr>
<th>Point 1 (Beginning)</th>
<th>Modal Breathy Creaky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinguished Modal from</td>
<td>HI*</td>
</tr>
<tr>
<td>Distinguished Breathy from</td>
<td>HI*</td>
</tr>
<tr>
<td>Distinguished Creaky from</td>
<td>HI*</td>
</tr>
</tbody>
</table>

- HI* and H2* and H2* distinguished breathy from non-breathy (modal and creaky) phonation, but not modal from creaky phonation.

Results: Acoustic Measures

- CPP distinguished breathy from creaky phonation on point 5.
- HI* distinguished breathy from non-breathy phonation on point 5.
Results: Acoustic Measures

<table>
<thead>
<tr>
<th>Point 9 (End)</th>
<th>Modal</th>
<th>Breathy</th>
<th>Creaky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinguished</td>
<td>Modal Breathy Creaky</td>
<td>III*</td>
<td>...</td>
</tr>
<tr>
<td>Distinguished</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- H1* distinguished creaky from non-creaky phonation on point 9.
- H1*-A1*, H1*-A2*, H1*-A3*, and H2*-H4* did not distinguish any of the phonation types.

Results: EGG Measures

- Closed quotient distinguished breathy from non-breathy (modal and creaky) phonation at all three time points.

One unexpected finding: breathy phonation is faster than modal and creaky phonation. This is consistent with breathy phonation in Gujarati (Khan 2009).

Results: How measures co-vary with f0

- The breathy (-g) tone was significantly correlated with:
  - H1* (R² = .77) and H1*-H2* (R² = .79)

Results: Duration

- Unlike Ratliff (1992), these results show that the shortest tone is the –b tone and not the –m tone.
Results: Gender difference

Females were significantly breathier than males on productions of breathy phonation at all three time points.

Graphs of $H_1^*$, $H_1^*-H_2^*$, CQ, and PCV for breathy vowels

Conclusion

1) What is the most accurate description of tone?

<table>
<thead>
<tr>
<th>Traditional description</th>
<th>Ratliff (1992)</th>
<th>Current Study</th>
<th>Orthographic tone symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>55</td>
<td>High-rising</td>
<td>h</td>
</tr>
<tr>
<td>Mid</td>
<td>43</td>
<td>Mid</td>
<td>i</td>
</tr>
<tr>
<td>Low</td>
<td>22</td>
<td>Low</td>
<td>l</td>
</tr>
<tr>
<td>High-falling</td>
<td>52</td>
<td>High-falling</td>
<td>j</td>
</tr>
<tr>
<td>Mid-rising</td>
<td>24</td>
<td>Mid-rising</td>
<td>v</td>
</tr>
<tr>
<td>Low-falling creaky</td>
<td>21</td>
<td>Low-falling creaky</td>
<td>m</td>
</tr>
<tr>
<td>Mid-low breathy</td>
<td>43</td>
<td>Mid-falling (for males)/High-falling (for females)</td>
<td>g</td>
</tr>
</tbody>
</table>

2) What are the acoustic and electroglottographic correlates of phonation?

Several measures distinguished two phonation categories.

- This suggests that phonation contrasts are realized across several phonetic dimensions in White Hmong.
- PCV and the amplitude of $H_1^*$ are the only measures that distinguished all three phonations.
- However, they did so at different time points.
- For both PCV and the amplitude of $H_1^*$:
  - On points 1 and 3, breathy phonation was distinguished from modal and creaky phonation.
  - On point 9, creaky phonation was distinguished from modal and breathy phonation.

3) How does phonation co-vary with f0?

- The breathy (-g) tone was significantly correlated with $H_1$ ($R^2 = .77$) and $H_1^*-H_2^*$ ($R^2 = .59$).
- The creaky (-m) tone was significantly correlated with $H_1-A3$ ($R^2 = -.91$) and PCV ($R^2 = .83$).

4) Are there durational differences between the three phonations?

- b was the shortest tone/phonation, not –m

5) Are there gender differences in the production of phonation?

- Females were significantly breathier than males on productions of breathy phonation at all three time points.

References