The role of visual speech cues in sound change: A study of the cot-caught contrast among Michigan speakers

Jonathan Havenhill and Youngah Do
Georgetown University and The University of Hong Kong

Overview

- Interspeaker articulatory variation has been argued to be a driving force in sound change (Baker et al., 2011).
- The role of visual cues in speech perception is also well known (McGurk & MacDonald, 1976).
- We investigate the interaction of visual speech cues and articulatory variation in sound change (McGuire & Babel, 2012) through a perceptual study of the Northern Cities Vowel Shift.
- We find that tokens of CAUGHT produced without lip rounding are significantly more likely to be misperceived as COT, but with individual variation.
- Articulatory patterns in which speakers produce both COT and CAUGHT with unround lips have the potential for merger.

Background

- Majors & Gordon (2008): rounding is preserved in CAUGHT; fronting accomplished through a change in tongue position.
- Havenhill (2015): Detroit speakers vary in how fronted CAUGHT is produced:

  **Pattern A**
  - Lip rounding: COT
  - Lip spread: CAUGHT

  **Pattern B**
  - Lip rounding: COT
  - Lip spread: CAUGHT

- Pattern A exhibits a lip rounding contrast, while Pattern B does not. Both patterns result in the same degree of acoustic contrast. Do these strategies differ in how they are visually perceived?
- Hypothesis: Pattern B is susceptible to misperception; CAUGHT produced without lip rounding will be perceived as COT.

Methods

- Ten (7 men, 3 women) 21-41 year-old (M = 26) participants, raised in Michigan.
- Participants exposed to congruous and incongruous audiovisual nonce word stimuli, with pink noise at 12dB SNR.
  - Target vowels: COT and CAUGHT.
  - Filler vowels: /i u e o/.
- Identification task: select a real English word which rhymes with the nonce word.
  - To avoid overlapping orthography for CAUGHT and COT.
  - Unknown reliability of AXB paradigm for audiovisual stimuli.
- Example: Stimulus: [zaot] → Please select rhyming word: bot bought

Results

- Significant overall effect of visual lip rounding; CAUGHT is misperceived as COT when presented with unround lips.

<table>
<thead>
<tr>
<th>Perception results for all participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est. Std.</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Vowel (cot)</td>
</tr>
<tr>
<td>Congruity (match/mismatch)</td>
</tr>
</tbody>
</table>

- Mixed-effects logistic regression with 2ne4. Fixed effects of vowel audio and video congruency. Random effects of subject and item:

<table>
<thead>
<tr>
<th>Perception results for Participant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel audio</td>
</tr>
<tr>
<td>Lip rounding</td>
</tr>
</tbody>
</table>

- Participant 4 correctly identifies all items with congruous lip rounding. However, a loss of lip rounding on CAUGHT has a stronger effect than the addition of lip rounding to COT.

Discussion

- There is a significant overall effect of visual incongruity whereby CAUGHT is perceived as COT when produced without rounding.
- Some potential implications:
  - Articulatory variation may weaken the perceptibility of the COT-CAUGHT contrast, leading to merger.
  - Visual lip rounding strengthens a weak acoustic contrast; pressure to maintain the contrast may favor variants that retain rounding.
  - Participants varied in the degree and direction of misperception.
  - Future research: how does listeners’ use of lip rounding for CAUGHT in their own speech influence perception?

Possible confounding factors:

- Overall correct perception is somewhat low in all conditions; identification of COT and CAUGHT may depend heavily on lexical knowledge.
- Lexical variation in rhyming words: choices given may not have been contrastive for all participants.

Suggestions/Questions

<table>
<thead>
<tr>
<th>Perception results for Participant 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel audio</td>
</tr>
<tr>
<td>Lip rounding</td>
</tr>
</tbody>
</table>

- Participant 9 exhibits the opposite pattern: COT is perceived as CAUGHT when produced with lip rounding, but CAUGHT is still correctly identified without rounding.

- There is a significant overall effect of visual incongruity whereby CAUGHT is perceived as COT when produced without rounding.

Acknowledgments

Many thanks to Pro Beker, W-E. Stager, and the staff of the University of Michigan Department of Linguistics for providing access to one of their phonetics labs. Thanks to the members of the Georgetown University Experimental Phonology seminar for helpful comments and suggestions.

References


Some potential implications:

- Articulatory variation may weaken the perceptibility of the COT-CAUGHT contrast, leading to merger.

Notes:

- Perception results for all participants
- Perception results for Participant 4
- Perception results for Participant 9

Contact Information

Jon: jeh241@georgetown.edu | jhavenhill.com
Youngah: youngah@hku.hk | youngahdo.net