Articulatory Correlates of Phonological Relationships

Kathleen Currie Hall, Hanna Smith, Kevin McMullin, Blake Allen, Noriko Yamane, Joash Gambarage

kathleen.hall@ubc.ca

LabPhon -- 27 July 2014
Research Question

• Do the phonological relationships a sound is a part of have an effect on how that sound is articulated?

Hypothesis

• Sounds that are "more contrastive" will be hyperarticulated relative to sounds that are "less contrastive."
Perception vs. Production

• **Perception:** Well-known that pairs of sounds that are contrastive are *perceived* as being more distinct than pairs of sounds that are non-contrastive (allophonic, neutralised), regardless of the acoustic reality. (Whalen, Best, & Irwin 1997; Huang 2001; Boomershine, Hall, Hume, & Johnson 2008, Kazanina, Phillips, & Idsardi 2006, Johnson & Babel 2010)

• **Production:** Assumption is that there is no effect of relationship on production (e.g., because acoustically identical pairs are perceived differently).

• Some inconclusive hints that this assumption isn’t true, and that degree of contrast matters for articulation. (Gick et al. 2006, Cristia & Seidl 2013, Baese-Berk & Goldrick 2009)
Quantifying Contrastiveness

• "Contrast" has typically been treated binarily: either 2 sounds are contrastive, or they are not.

• Contrast is primarily determined through lexical distinction and predictability of distribution.
  — Contrastive sounds can distinguish between words.
  — Contrastive sounds occur in the same contexts and thus are not predictable from their environment.
Quantifying Contrastiveness

- Both of these can be turned into gradient measures, which is useful for capturing "intermediate" phonological relations.

- **Functional load**: how many lexical items do sounds actually distinguish? (Martinet 1955; Hockett 1966; Surendran & Niyogi 2003; Wedel et al. 2013)

- **(Un)-predictability of distribution**: to what extent is the choice between two sounds uncertain across contexts? (entropy; Hall 2009, 2012)
Quantifying Contrastiveness

• For both measures, higher numbers indicate greater contrastiveness.

• Both functional load and unpredictability of distribution can be calculated from corpora (e.g., using Phonological CorpusTools (Mackie et al., yesterday).
Experimental Overview

• Conduct an ultrasound study to look for differences in productions between sounds that are being used more vs. less contrastively.

• **Dependent variable**: magnitude of the movement of the tongue throughout the course of each target segment.
  
  — NB: We are *not* looking at articulatory posture.

• **Test conditions**: High vowels in English, in closed vs. open syllables.
Optical Flow Analysis (OFA)

- A way of measuring apparent motion by comparing the difference in location of individual pixels from frame to frame.
  
  - See Horn & Schunck (1981), Barbosa et al. (2008), Fais et al. (2010; tomorrow!).

Adjacent frames from a video.

Resulting optical flow field mapping changes between frames & hence motion.

Fleet & Weiss (2006: 10)
Advantages of OFA

• Possible to extract information:
  – from the entire production, rather than static points;
  – about movement from different parts of the tongue;
  – about both velocity and magnitude of movement;
  – relatively quickly from many speakers (as fast as Praat segmenting).

• Direct comparison across speakers is possible, because measurements within a speaker can be easily normalized.
Target Sounds

• Tense (vs. lax) mid and high vowels
  — (Eventually, comparison between Canadian English and Canadian French...currently just English)

• Phonological relations:

<table>
<thead>
<tr>
<th></th>
<th>[i] / [i]</th>
<th>[ɛ] / [ɛ]</th>
<th>[ʊ] / [ʊ]</th>
<th>[o] / [ɔ]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Closed</strong></td>
<td>Contrastive</td>
<td>Contrastive</td>
<td>Contrastive</td>
<td>Contrastive</td>
</tr>
<tr>
<td>Syllables</td>
<td><em>(bead / bid)</em></td>
<td><em>(bayed / bed)</em></td>
<td><em>(who'd / hood)</em></td>
<td><em>(node / gnawed)</em></td>
</tr>
<tr>
<td><strong>Open</strong></td>
<td>Non-contrastive</td>
<td>Non-contrastive</td>
<td>Non-contrastive</td>
<td>Contrastive</td>
</tr>
<tr>
<td>Syllables</td>
<td>*(bee / <em>[bɪ]</em>)</td>
<td>*(bay / <em>[bɛ]</em>)</td>
<td>*(who / <em>[hʊ]</em>)</td>
<td><em>(know / gnaw)</em></td>
</tr>
</tbody>
</table>
Target Sounds

• Predictions:
  — Contrastive sounds show greater differences in articulatory measures than non-contrastive sounds

• Phonological relations:

<table>
<thead>
<tr>
<th></th>
<th>[i] / [i]</th>
<th>[ɛ] / [ɛ]</th>
<th>[ʊ] / [ʊ]</th>
<th>[o] / [ɔ]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Closed Syllables</strong></td>
<td>Contrastive <em>(bead / bid)</em></td>
<td>Contrastive <em>(bayed / bed)</em></td>
<td>Contrastive <em>(who'd / hood)</em></td>
<td>Contrastive <em>(node / gnawed)</em></td>
</tr>
<tr>
<td><strong>Open Syllables</strong></td>
<td>Non-contrastive *(bee / <em>[bɪ]</em>)</td>
<td>Non-contrastive *(bay / <em>[bɛ]</em>)</td>
<td>Non-contrastive *(who / <em>[hʊ]</em>)</td>
<td>Contrastive <em>(know / gnaw)</em></td>
</tr>
</tbody>
</table>
Stimuli

• 104 words total; 78 are targets (tense vowels):

<table>
<thead>
<tr>
<th></th>
<th>[i] / [ɪ]</th>
<th>[e] / [ɛ]</th>
<th>[u] / [ʊ]</th>
<th>[o] / [ɔ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Syllables</td>
<td>10 [i]</td>
<td>10 [e]</td>
<td>8 [u]</td>
<td>10 [o] / 5 [ɔ]</td>
</tr>
</tbody>
</table>

• All vowels occurred in stressed, word-final syllables.
• All but one word (delay) were monosyllabic.
Stimuli

- 104 words total; 78 are targets (tense vowels):

<table>
<thead>
<tr>
<th></th>
<th>[i] / [ɪ]</th>
<th>[ɛ] / [ɛ]</th>
<th>[ʊ] / [ʊ]</th>
<th>[ɔ] / [ɔ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Syllables</td>
<td>10 [i]</td>
<td>10 [ɛ]</td>
<td>8 [ʊ]</td>
<td>10 [ɔ] / 5 [ɔ]</td>
</tr>
</tbody>
</table>

- Task: Read wordlist, with words presented in random order one at a time on a computer screen.
Participants

• 14 female speakers who self-identify as native English speakers, between the ages of 18 and 26 (mean = 21.5).

• Recruited from a public "paid studies" list at the University of British Columbia and word of mouth.

• Paid $20 each for participation.
Recordings

• Aloka SSD-5000 ultrasound machine
• UST-9118 endovaginal 180° electronic curved array probe at approximately a 90° angle to the floor
• Two-dimensional mid-sagittal ultrasound video recordings of the tongue were recorded digitally, directly to an attached computer running iMovie, at a frame rate of 30 frames / sec.
• Sound simultaneously recorded with external microphone.
Sample Recordings

"face...plea...hoe...gill...me...pep"
Data Extraction

• Audio extracted from the video file.
• Praat TextGrids created to delimit the vowel regions of target words.
• Frames of interest calculated from Praat TextGrids.
Measurements

• Optical flow analysis using "FlowAnalyzer" software (Barbosa & Vatikiotis-Bateson, p.c.; see also Barbosa et al. 2008)

• Measures the magnitude of movement in the x- and y-dimensions, frame-by-frame, for all specified regions in the ultrasound recording (cf. Horn & Schunck 1981).

• Can be set to ignore certain regions.

• Also calculates overall Euclidean magnitude of movement, frame by frame.
NB: In the current study, the full video was used, with no selection of individual tongue areas.
Normalization

- Because each speaker's movements are measured in terms of magnitude, they can be normalized within speakers and then compared across speakers.

- A standard z-score normalization was used.

- No normalization is needed for duration in the current study, because measurements are being compared frame-to-frame.
  - Longer utterances have more points that go into the calculation, but individual points each cover the same time period.
Word-Embedding

• Words had target vowels of 3-53 frames in duration; the mean and median frame duration was 8 (all but 2 were between 3 and 16 frames).
• Movements in one frame are likely to co-vary with movements in adjacent frames.
• To minimize covariance, overall magnitudes of movement from all frames in a given word were averaged to give a single average "magnitude of movement" measure per word.
Results (14 Speakers, Wordlist)

Vowel: [i]

Vowel: [u]

Vowel: [e]

Vowel: [o]
Results (14 Speakers, Wordlist)

Vowel: [i]

Syllable Type

Mag. of Movement

-2 -1 0 1 2 3

closed open

t-test on magnitudes:
t(250.281) = 7.92
p < 0.001
Results (14 Speakers, Wordlist)

Vowel: [e]

![Box plot showing comparison between closed and open syllable types for vowel [e].]

- t-test on magnitudes: $t(275.82) = 4.56$
- $p < 0.001$
Results (14 Speakers, Wordlist)

Vowel: [o]

Mag. of Movement
-2 -1 0 1 2 3

Syllable Type
closed open

t-test on magnitudes:
t(231.69) = 6.04
p < 0.001
Results (14 Speakers, Wordlist)

Vowel: [u]

Mag. of Movement

Syllable Type

-2 -1 0 1 2 3

closed open

t-test on magnitudes:
t(241.76) = 0.51
p = 0.61
Measuring Contrastiveness

<table>
<thead>
<tr>
<th>Vowel Pair</th>
<th>Shows Effect?</th>
<th>Traditional Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i] / [ɪ]</td>
<td>✔️</td>
<td>✔️ ✔️ ✗</td>
</tr>
<tr>
<td>[e] / [ɛ]</td>
<td>✔️</td>
<td>✔️ ✔️ ✗</td>
</tr>
<tr>
<td>[o] / [ɔ]</td>
<td>✔️</td>
<td>✔️ ✔️ ✔️</td>
</tr>
<tr>
<td>[u] / [ʊ]</td>
<td>✗</td>
<td>✔️ ✔️ ✗</td>
</tr>
</tbody>
</table>

*Closed* | *Open*
Measuring Contrastiveness

All calculations were done using *Phonological CorpusTools* (Mackie et al., yesterday), based on the IPhOD corpus of English (Vaden et al. 2009), which in turn uses token frequencies from the SUBTLEX corpus (Brysbaert & New 2009).

<table>
<thead>
<tr>
<th>Vowel Pair</th>
<th>Shows Effect?</th>
<th>Functional Load</th>
<th>Type-Based UN-Predictability of Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Closed</td>
<td>Open</td>
</tr>
<tr>
<td>[i] / [ɪ]</td>
<td>✔️</td>
<td>98     / 8251</td>
<td>= 0.012</td>
</tr>
<tr>
<td>[e] / [ɛ]</td>
<td>✔️</td>
<td>86     / 8251</td>
<td>= 0.010</td>
</tr>
<tr>
<td>[o] / [ɔ]</td>
<td>✔️</td>
<td>41     / 8251</td>
<td>= 0.005</td>
</tr>
<tr>
<td>[u] / [ʊ]</td>
<td>✗</td>
<td>7      / 8251</td>
<td>= 0.001</td>
</tr>
</tbody>
</table>

All calculations were done using *Phonological CorpusTools* (Mackie et al., yesterday), based on the IPhOD corpus of English (Vaden et al. 2009), which in turn uses token frequencies from the SUBTLEX corpus (Brysbaert & New 2009).
Measuring Contrastiveness

**Functional load:** Pairs that show the effect have a higher functional load in closed syllables than open syllables, and have a higher functional load in closed syllables than the pair that doesn’t show the effect.

<table>
<thead>
<tr>
<th>Vowel Pair</th>
<th>Shows Effect?</th>
<th>Functional Load</th>
<th>Type-Based UN-Predictability of Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Closed</td>
<td>Open</td>
</tr>
<tr>
<td>[i] / [ɪ]</td>
<td>✔️</td>
<td>98 / 8251 = 0.012</td>
<td>0 / 8251 = 0</td>
</tr>
<tr>
<td>[e] / [ɛ]</td>
<td>✔️</td>
<td>86 / 8251 = 0.010</td>
<td>0 / 8251 = 0</td>
</tr>
<tr>
<td>[o] / [ɔ]</td>
<td>✔️</td>
<td>41 / 8251 = 0.005</td>
<td>7 / 8251 = 0.001</td>
</tr>
<tr>
<td>[u] / [ʊ]</td>
<td>✗</td>
<td>7 / 8251 = 0.001</td>
<td>0 / 8251 = 0</td>
</tr>
</tbody>
</table>
**Measuring Contrastiveness**

Type-based unpredictability of distribution: **Pairs that show the effect** have a greater unpredictability in closed syllables than open syllables, and have a higher unpredictability in closed syllables than **the pair that doesn’t show the effect**.

<table>
<thead>
<tr>
<th>Vowel Pair</th>
<th>Shows Effect?</th>
<th>Functional Load</th>
<th>Type-Based UN-Predictability of Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>![i] / ![ɪ]</td>
<td>✔</td>
<td>98 / 8251 = 0.012</td>
<td>0 / 8251 = 0</td>
</tr>
<tr>
<td>![e] / ![ɛ]</td>
<td>✔</td>
<td>86 / 8251 = 0.010</td>
<td>0 / 8251 = 0</td>
</tr>
<tr>
<td>![o] / ![ɔ]</td>
<td>✔</td>
<td>41 / 8251 = 0.005</td>
<td>7 / 8251 = 0.001</td>
</tr>
<tr>
<td>![u] / ![ʊ]</td>
<td>✗</td>
<td>7 / 8251 = 0.001</td>
<td>0 / 8251 = 0</td>
</tr>
</tbody>
</table>
The Role of Contrastiveness

• [u] / [ʊ] is distinguished from the other three pairs on both measures.

• Being relatively more lexically contrastive leads to greater relative hyperarticulation.
Difference vs. Threshold?

• For functional load, the *difference* between closed and open monosyllabic syllables is greater for pairs that show the effect than for [u] / [ʊ].
• This isn’t true for type-based unpredictability of distribution.
• Alternatively, there could be some threshold value beyond which relative hyperarticulation occurs (or below which hypoarticulation occurs?).
  — Somewhere between 7 and 41 minimal pairs...
  — ...or somewhere between 0.76 and 0.95 bits of uncertainty.
Conclusions

• So far, relatively clear evidence of articulatory differences between tense vowels in environments where they are more contrastive vs. those where they are less contrastive.

• The differences don't seem to be tied simply to syllable type, as [u] / [ʊ] do not pattern the same way as the other pairs.

• This holds both for cases where it's more clearly a neutralized contrast (e.g. [i]~[ɪ], [ɛ]~[ɛ]) and for cases where there are simply varying degrees of "full" contrast (e.g., [o]~[ɔ]).
Conclusions

• It does not hold for [u]~[ʊ]; simple binary distinction in phonological relations doesn't seem to explain the effect.
• Lexical contrastiveness—either functional load or type-based unpredictability of distribution—seems to matter.
• Regardless of whether this is based on difference or threshold values, these results suggest that speakers are making articulatory distinctions associated with lexical factors:
  – i.e., for sounds that are more important in terms of distinguishing lexical items, there is a greater articulatory difference.
Open Questions

• If there is a minimum threshold of contrastiveness, what is that threshold and why?
• Does the threshold depend on functional load or type-based unpredictability or both?
• Do other usage factors (token frequency, neighbourhood density) matter as well?
• Is this relative hyper- or hypo-articulation?
• How does a vowel's other contrasts come into play (e.g., the fact that [i] and [u] also contrast with each other)?
Open Questions

• Do these articulatory differences have acoustic effects?
  — if so, are such effects perceptible?
• Could the differences be tied simply to these particular vowels? What happens in a language where the same vowels have different relations in the same syllable positions (e.g., Canadian French)?
• Are the effects consistent across tasks?
Thank You!

• Especially to:
  – Eric Vatikiotis-Bateson
  – Bryan Gick
  – Molly Babel
  – The Vancouver Phonology Group
  – The UBC Inter-disciplinary Speech Research Lab
  – The UBC Speech in Context Lab

• Funding from UBC's Hampton Award
References


References

• Hall, Kathleen Currie. 2009. A probabilistic model of phonological relationships from contrast to allophony. Columbus, OH: The Ohio State University Doctoral dissertation.
References