# 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.
  - e.g., Goldsmith (1995), Ladd (2006), Scobbie & Stuart-Smith (2008), Hall (2013)
- 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

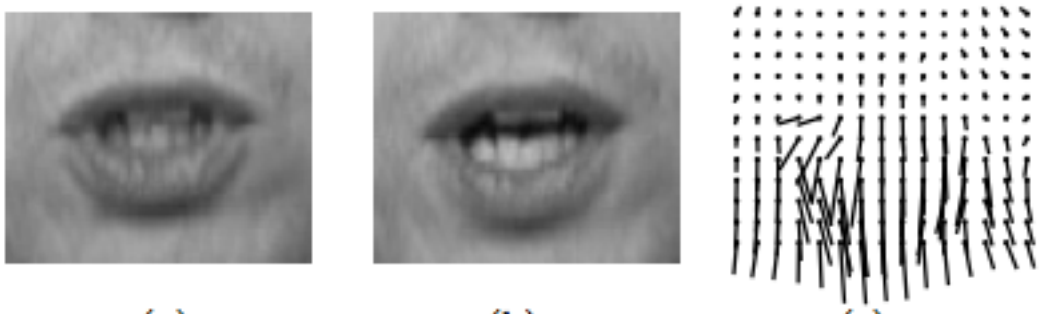
- For both measures, higher numbers indicate greater contrastiveness.
- Both functional load and unpredictability of distribution can be calculated from corpora (e.g., using *Phonological CorpusTools* (Mackie et al., yesterday).

# 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 indivdual pixels from frame to frame.
  - *See Horn & Schunck (1981), Barbosa et al. (2008), Fais et al. (2010; tomorrow!).*



(a) (b) (c)

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:

| | [i] / [ɪ] | [e] / [ɛ] | [u] / [ʊ] | [o] / [ɔ] |
|---|---|---|---|---|
| **Closed Syllables** | Contrastive (*bead* / *bid*) | Contrastive (*bayed* / *bed*) | Contrastive (*who'd* / *hood*) | Contrastive (*node* / *gnawed*) |
| **Open Syllables** | Non-contrastive (*bee* / *[bɪ]) | Non-contrastive (*bay* / *[bɛ]) | Non-contrastive (*who* / *[hʊ]) | Contrastive (*know* / *gnaw*) |

# Target Sounds

- Predictions:
  - Contrastive sounds show greater differences in articulatory measures than non-contrastive sounds
- Phonological relations:

| | [i] / [ɪ] | [e] / [ɛ] | [u] / [ʊ] | [o] / [ɔ] |
|---|---|---|---|---|
| **Closed Syllables** | Contrastive (*bead* / *bid*) | Contrastive (*bayed* / *bed*) | Contrastive (*who'd* / *hood*) | Contrastive (*node* / *gnawed*) |
| **Open Syllables** | Non-contrastive (*bee* / *[bɪ]) | Non-contrastive (*bay* / *[bɛ]) | Non-contrastive (*who* / *[hʊ]) | Contrastive (*know* / *gnaw*) |

# Stimuli

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

| | [i] / [ɪ] | [e] / [ɛ] | [u] / [ʊ] | [o] / [ɔ] |
|---|---|---|---|---|
| **Closed Syllables** | 10 [i] / 10 [ɪ] | 10 [e] / 10 [ɛ] | 10 [u] / 9 [ʊ] | 10 [o] / 10 [ɔ] |
| **Open Syllables** | 10 [i] | 10 [e] | 8 [u] | 10 [o] / 5 [ɔ] |

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

# Stimuli

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

| | [i] / [ɪ] | [e] / [ɛ] | [u] / [ʊ] | [o] / [ɔ] |
|---|---|---|---|---|
| **Closed Syllables** | 10 [i] / 10 [ɪ] | 10 [e] / 10 [ɛ] | 10 [u] / 9 [ʊ] | 10 [o] / 10 [ɔ] |
| **Open Syllables** | 10 [i] | 10 [e] | 8 [u] | 10 [o] / 5 [ɔ] |

- 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"

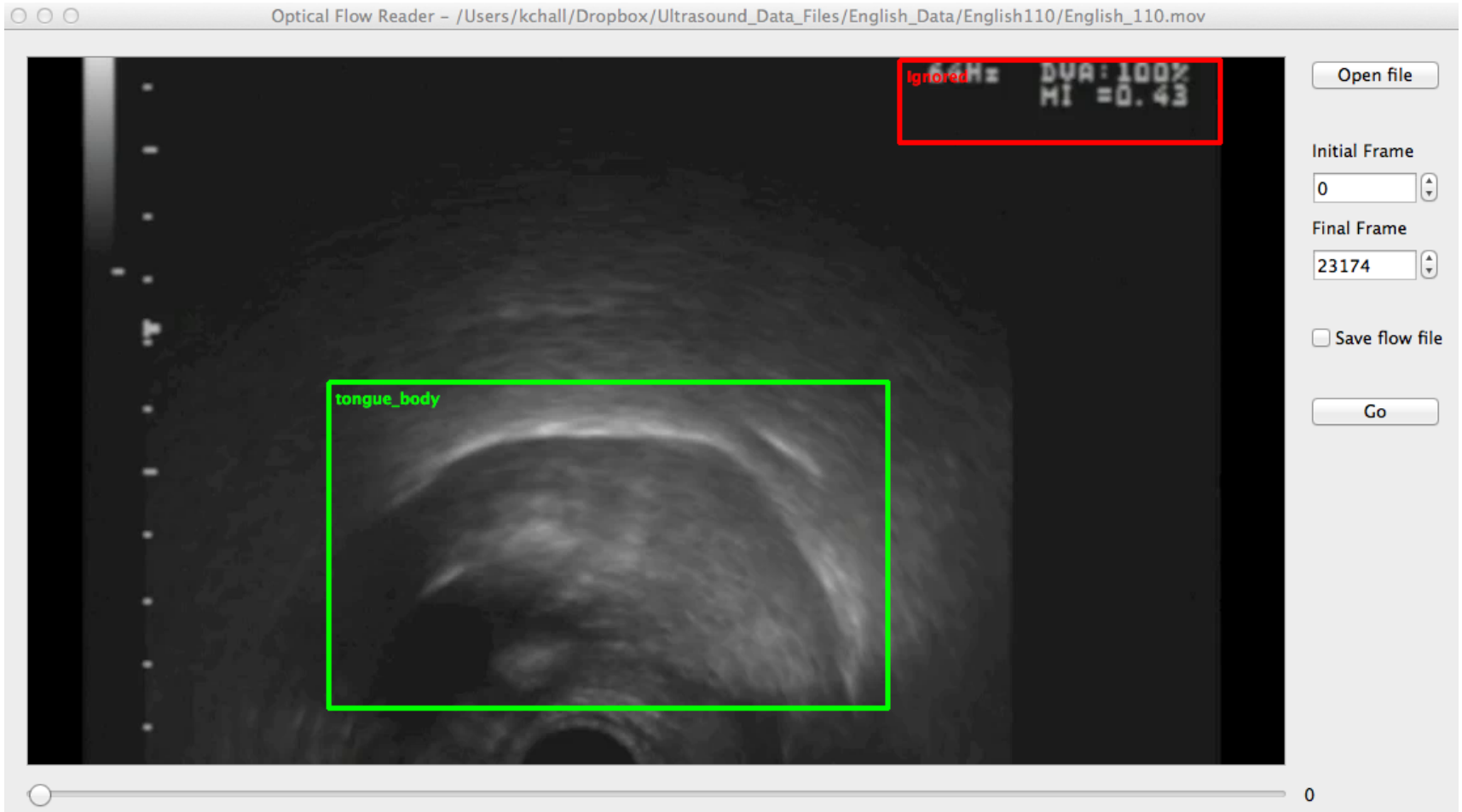
# Data Extraction

- 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.

# FlowAnalyzer




NB: In the current study, the full video was used, with no selection of individual tongue areas.

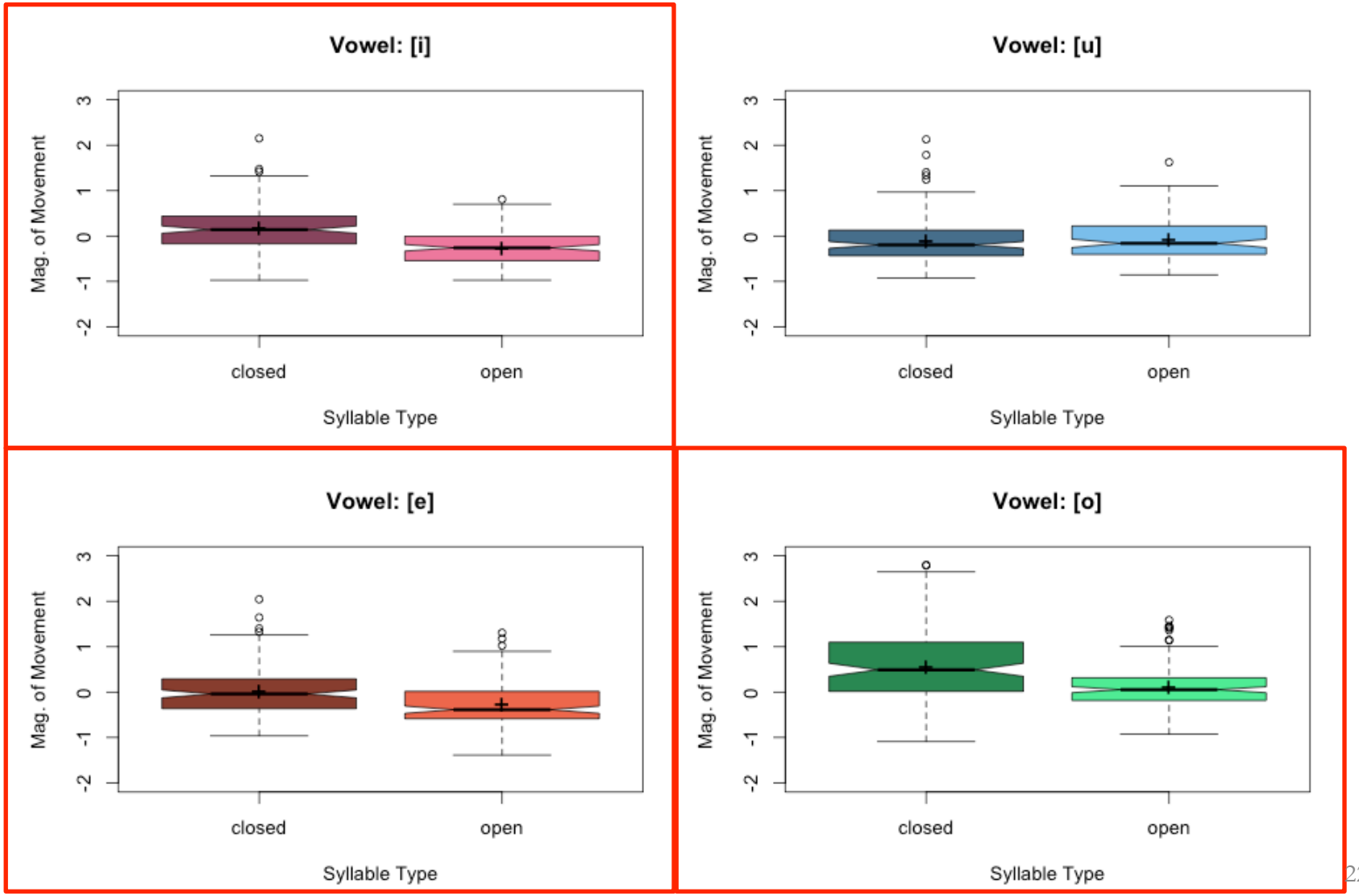
# Normalization

- 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.

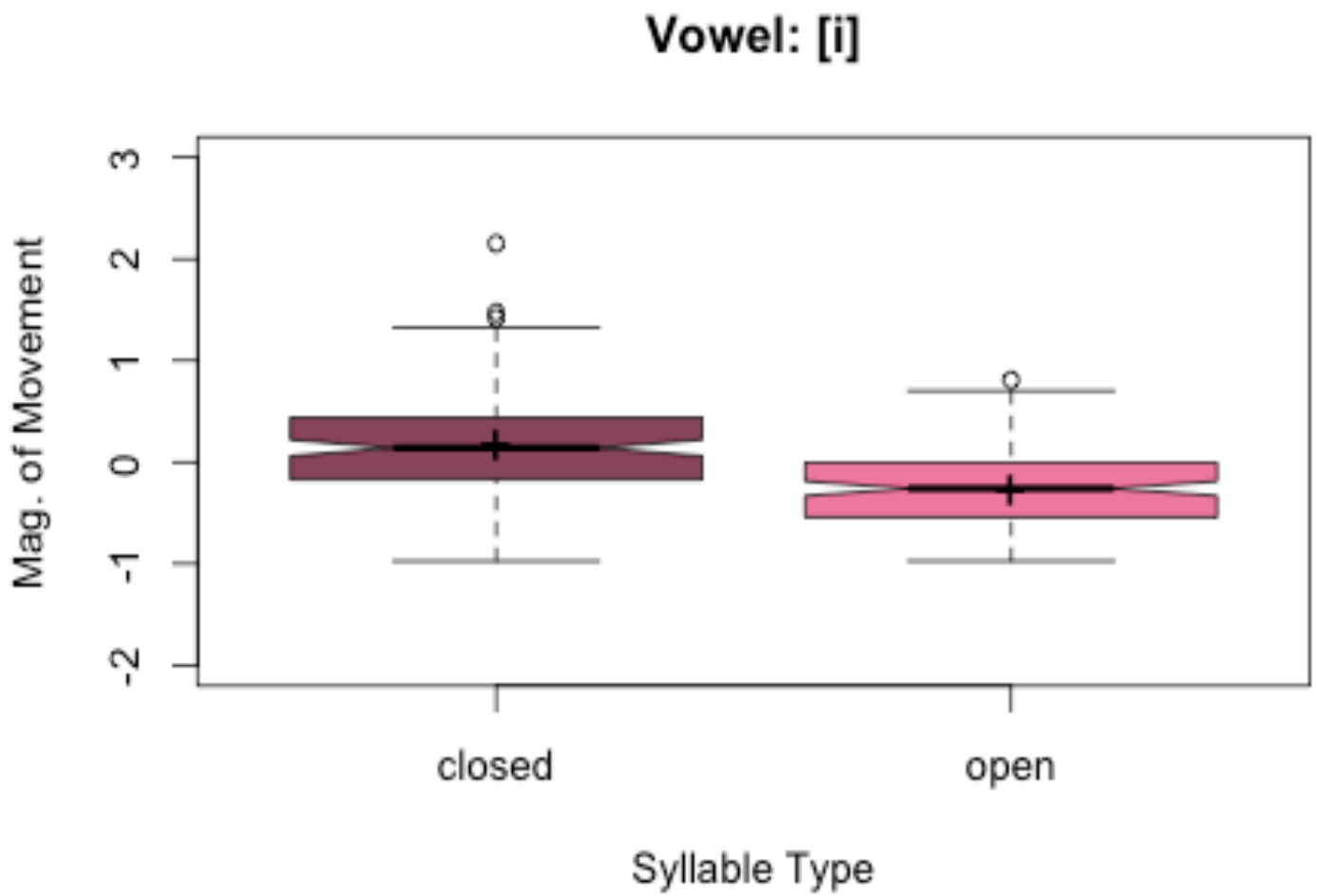
# Word-Embedding

- 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)

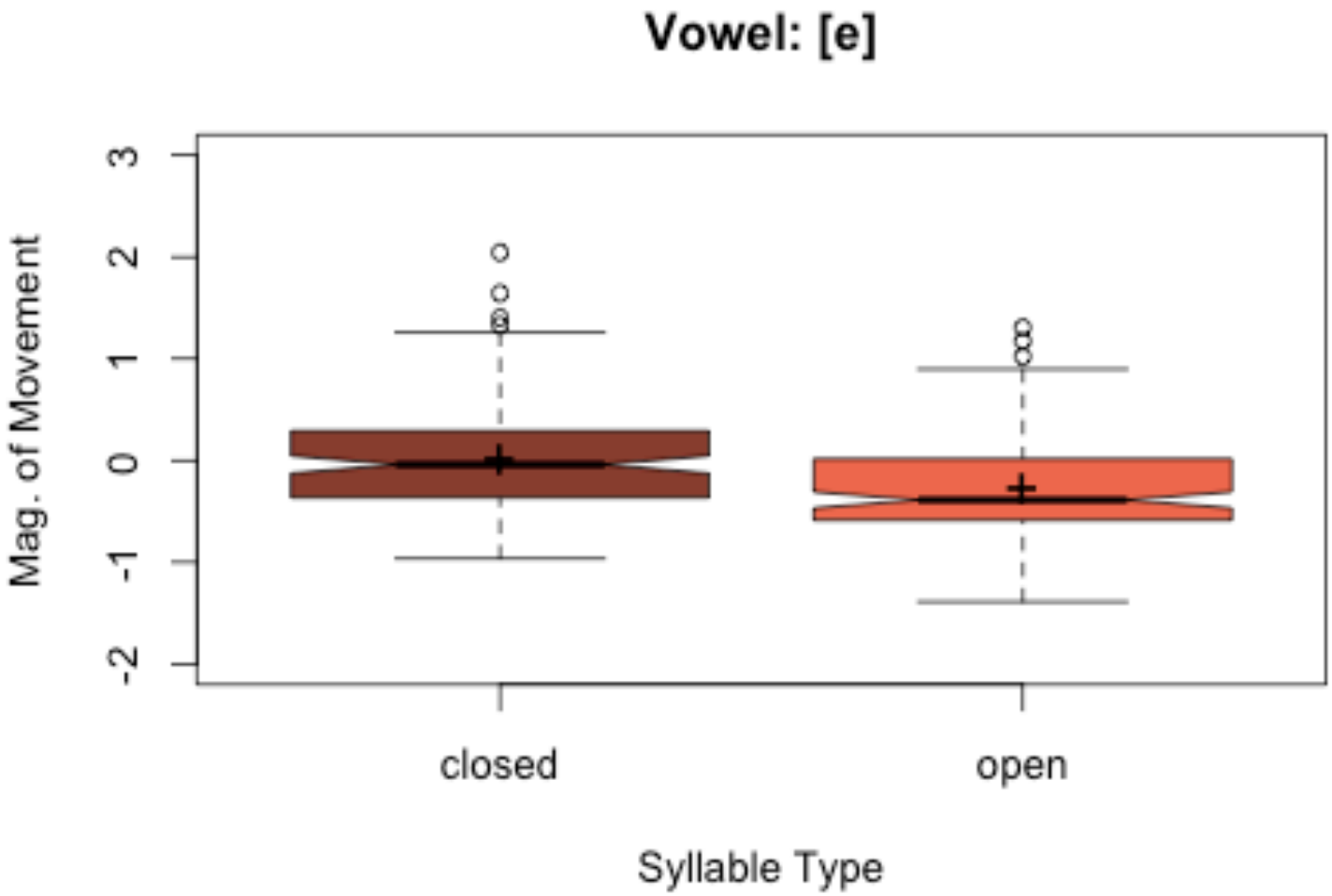


# Results (14 Speakers, Wordlist)



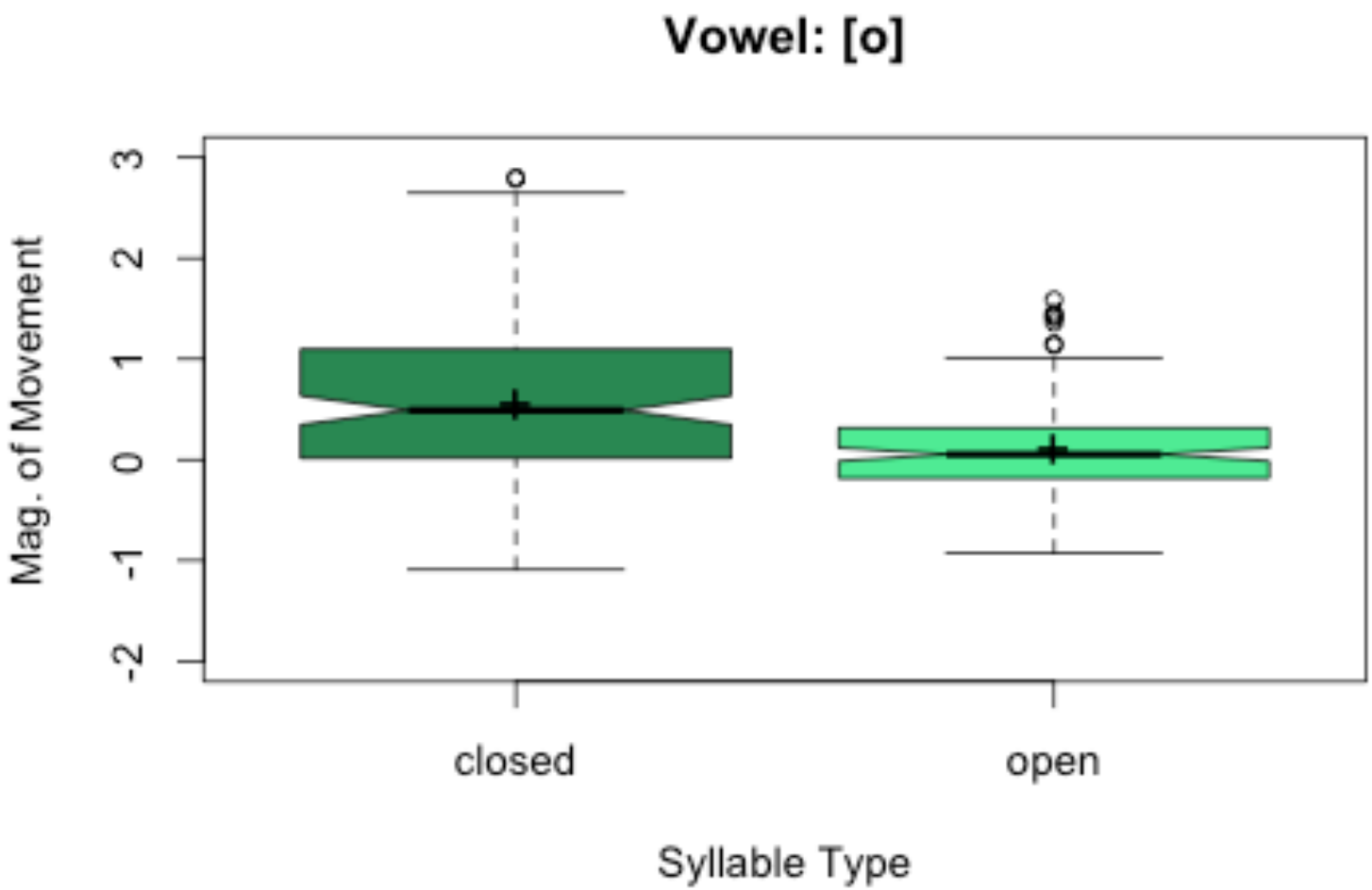

t-test on magnitudes:
$t(250.281) = 7.92$
$p < 0.001$

# Results (14 Speakers, Wordlist)



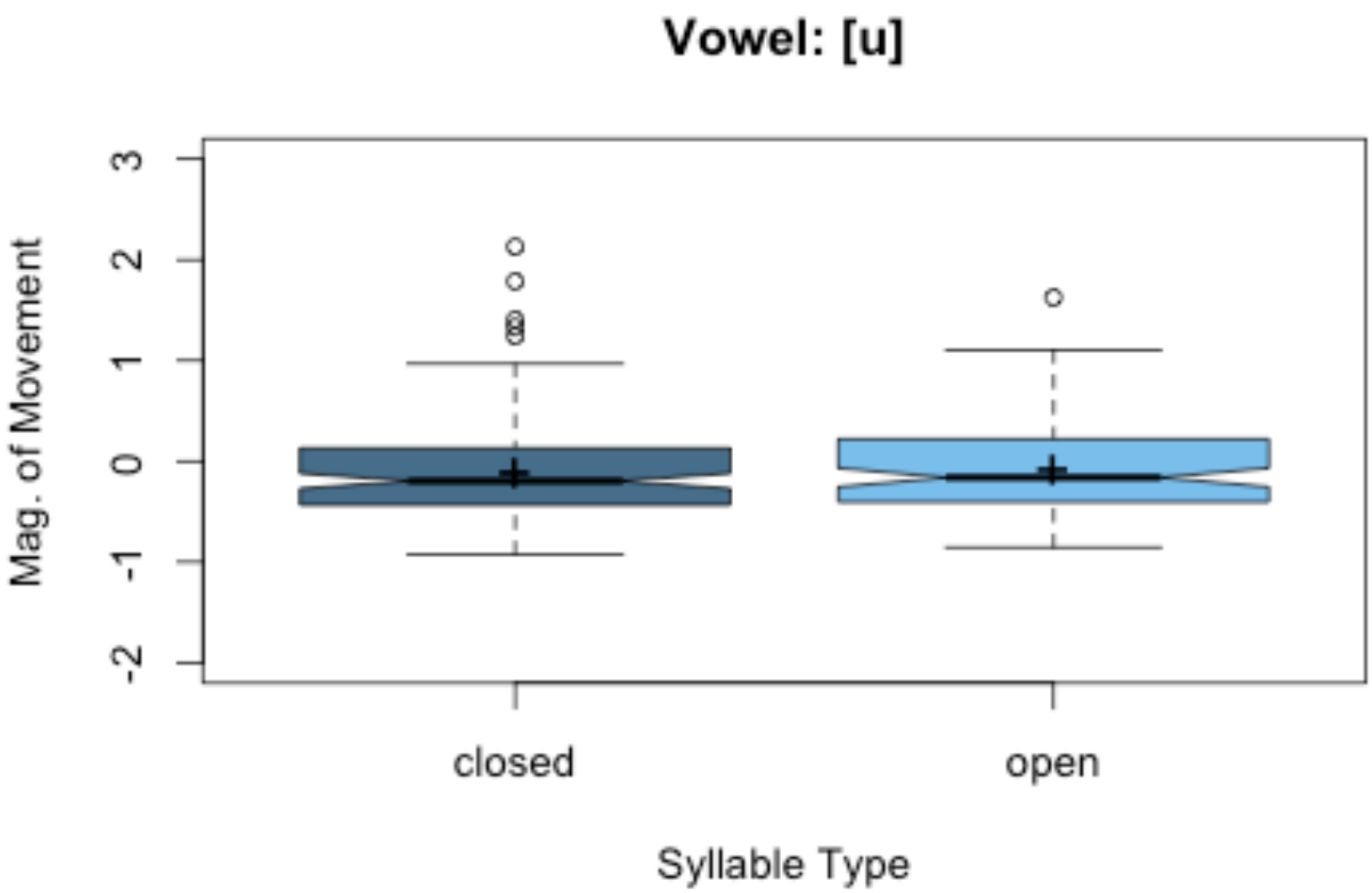

t-test on magnitudes:
$t(275.82) = 4.56$
$p < 0.001$

# Results (14 Speakers, Wordlist)




t-test on magnitudes:
$t(231.69) = 6.04$
$p < 0.001$

# Results (14 Speakers, Wordlist)




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

# Measuring Contrastiveness

| Vowel Pair | Shows Effect? | Traditional Contrast | |
|---|---|---|---|
| | | *Closed* | *Open* |
| [i] / [ɪ] | ✔ | ✔ | ✗ |
| [e] / [ɛ] | ✔ | ✔ | ✗ |
| [o] / [ɔ] | ✔ | ✔ | ✔ |
| [u] / [ʊ] | ✗ | ✔ | ✗ |

# 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).

| Vowel Pair | Shows Effect? | Functional Load | | Type-Based UN-Predictability of Distribution | |
|---|---|---|---|---|---|
| | | *Closed* | *Open* | *Closed* | *Open* |
| [i] / [ɪ] | ✔ | 98 / 8251 = 0.012 | 0 / 8251 = 0 | 0.95 | 0.00 |
| [e] / [ɛ] | ✔ | 86 / 8251 = 0.010 | 0 / 8251 = 0 | 0.996 | 0.00 |
| [o] / [ɔ] | ✔ | 41 / 8251 = 0.005 | 7 / 8251 = 0.001 | 0.99 | 0.76 |
| [u] / [ʊ] | ✗ | 7 / 8251 = 0.001 | 0 / 8251 = 0 | 0.70 | 0.00 |

# 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.

| Vowel Pair | Shows Effect? | Functional Load | | Type-Based UN-Predictability of Distribution | |
|---|---|---|---|---|---|
| | | *Closed* | *Open* | *Closed* | *Open* |
| [i] / [ɪ] | ✔ | 98 / 8251 = 0.012 | 0 / 8251 = 0 | 0.95 | 0.00 |
| [e] / [ɛ] | ✔ | 86 / 8251 = 0.010 | 0 / 8251 = 0 | 0.996 | 0.00 |
| [o] / [ɔ] | ✔ | 41 / 8251 = 0.005 | 7 / 8251 = 0.001 | 0.99 | 0.76 |
| [u] / [ʊ] | ✗ | 7 / 8251 = 0.001 | 0 / 8251 = 0 | 0.70 | 0.00 |

# 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.

| Vowel Pair | Shows Effect? | Functional Load | | Type-Based UN-Predictability of Distribution | |
|---|---|---|---|---|---|
| | | *Closed* | *Open* | *Closed* | *Open* |
| [i] / [ɪ] | ✔ | 98 / 8251 = 0.012 | 0 / 8251 = 0 | 0.95 | 0.00 |
| [e] / [ɛ] | ✔ | 86 / 8251 = 0.010 | 0 / 8251 = 0 | 0.996 | 0.00 |
| [o] / [ɔ] | ✔ | 41 / 8251 = 0.005 | 7 / 8251 = 0.001 | 0.99 | 0.76 |
| [u] / [ʊ] | ✗ | 7 / 8251 = 0.001 | 0 / 8251 = 0 | 0.70 | 0.00 |

# The Role of Contrastiveness

- [u] / [ʊ] is distinguished from the other three pairs on both measures.
- <u>Being relatively more lexically contrastive leads to greater relative hyperarticulation.</u>

# 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], [e]~[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


- Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. Language and Cognitive Processes 24.527-54.
- Barbosa, Adriano V., Hani C. Yehia & Eric Vatikiotis-Bateson. 2008. Linguistically valid movement behavior measured non-invasively. Auditory and visual speech processing - AVSP08, ed. by R. Göcke, P. Lucey & S. Lucey, 173-77. Moreton Island, Australia: Caual Productions.
- Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume & Keith Johnson. 2008. The influence of allophony vs. contrast on perception: the case of Spanish and English. *Contrast in phonology: Perception and acquisition*, ed. by P. Avery, B.E. Dresher & K. Rice, 145-71. Berlin: Mouton.
- Brysbaert, Marc, & Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4): 977-990.
- Cristia, Alejandrina & Amanda Seidl. 2013. The hyperarticulation hypothesis of infant-directed speech. Journal of Child Language 41.1-22.
- Fleet, David J. & Yair Weiss. 2006. Optical Flow Estimation. Handbook of Mathematical Models in Computer Vision, ed. by N. Paradgios, Y. Chen & O.D. Faugeras, 239-57. Secaucus, NJ: Springer.
- Gick, Bryan, Douglas Pulleyblank, Fiona Campbell & Ngessimo Mutaka. 2006. Low vowels and ATR harmony in Kinande. Phonology 23.1-20.

# References


- Goldsmith, John. 1995. Phonological theory. The Handbook of Phonological Theory, ed. by J.A. Goldsmith, 1-23. Cambridge, MA: Blackwell.
- Hall, Kathleen Currie. 2009. A probabilistic model of phonological relationships from contrast to allophony. Columbus, OH: The Ohio State University Doctoral dissertation.
- Hall, Kathleen Currie. 2012. Phonological relationships: A probabilistic model. McGill Working Papers in Linguistics 22.
- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. The Linguistic Review 30.215-75.
- Hockett, Charles F. 1966. The quantification of functional load: A linguistic problem. U.S. Air Force Memorandum RM-5168-PR.
- Horn, Berthold K. P. & Brian G. Schunck. 1981. Determining optical flow. Artificial Intelligence 17.185-203.
- Huang, Tsan. 2001. The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. Studies on the interplay of speech perception and phonology, ed. by E. Hume & K. Johnson, 23-42. Columbus, OH: Ohio State University Working Papers in Linguistics.
- Johnson, Keith & Molly Babel. 2010. On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. Journal of Phonetics 38.127-36.

# References


- Kazanina, Nina, Colin Phillips & William J. Idsardi. 2006. The influence of meaning on the perception of speech sounds. Proceedings of the National Academy of Sciences of the United States of America 103.11381-86.
- Ladd, D. Robert. 2006. "Distinctive phones" in surface representation. Laboratory Phonology 8, ed. by L.M. Goldstein, D.H. Whalen & C.T. Best, 3-26. Berlin: Mouton de Gruyter.
- Martinet, André. 1955. Économie des changements phonétiques Bern: Francke.
- Pierrehumbert, Janet B., Beckman, Mary E., & Ladd, D. R. (2001). Conceptual foundations of phonology as laboratory science. In N. Burton-Roberts, P. Carr & G. Docherty (Eds.), *Phonological Knowledge: Its Nature and Status* (pp. 273-304). Cambridge: Cambridge UP.
- Scobbie, James M. & Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the indeterminacy of the segmental inventory: examples from Scottish English. Contrast in phonology: Perception and acquisition, ed. by P. Avery, B.E. Dresher & K. Rice. Berlin: Mouton.
- Surendran, Dinoj & Partha Niyogi. 2003. Measuring the functional load of phonological contrasts. In *Tech. Rep. No. TR-2003-12*. . Chicago.
- Vaden, K.I., Halpin, H.R., Hickok, G.S. (2009). Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. Available from http://www.iphod.com.
- Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. Cognition 128.179-86.
- Whalen, D. H., Catherine T. Best & Julia R. Irwin. 1997. Lexical effects in the perception and production of American English /p/ allophones. Journal of Phonetics 25.501-28.