Age vectors vs. axes of intraspeaker variation for North American and Scottish English vowel formants

Age vectors vs. axes of intraspeaker variation for North American and Scottish English vowel formants

Jeff Mielke, NCSU
Josef Fruehwald, University of Edinburgh
Erik Thomas, NCSU
Michael McAuliffe, McGill University
Morgan Sonderegger, McGill University
Robin Dodsworth, NCSU

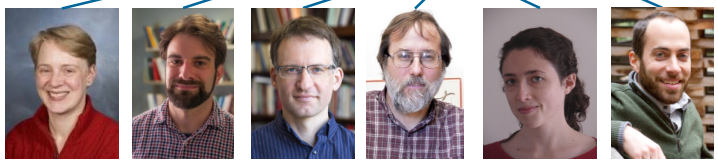
SPADE
SPeech Across Dialects of English

SPADE
SPeech Across Dialects of English

August 2017 – July 2020
http://spade.glasgow.ac.uk/

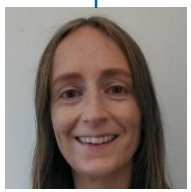
SPADE
SPeech Across Dialects of English

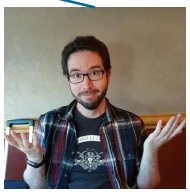
investigators



SPADE
SPeech Across Dialects of English

Rachel Macdonald





Michael McAuliffe

SPADE
SPeech Across Dialects of English

motivation

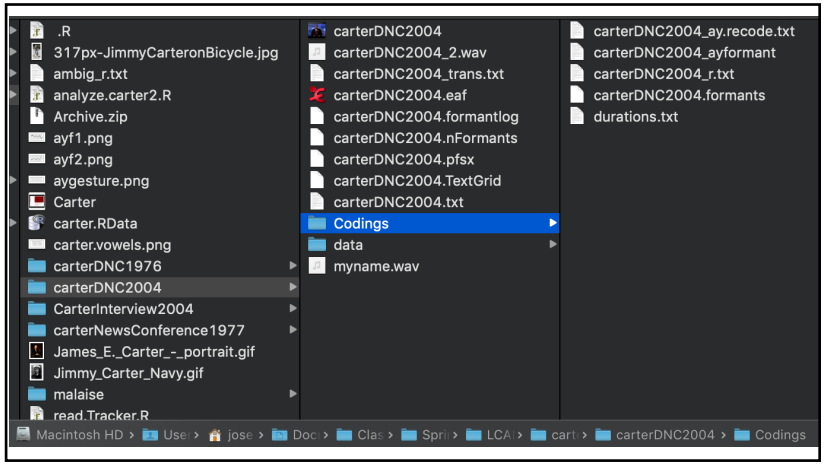
There is a large and ever growing collections of annotated speech data, improving automated analysis techniques, and a cultural shift towards greater data sharing.

SPADE
SPeech Across Dialects of English

motivation

But…

- These corpora are not all formatted in the same way.
- Many collections cannot ethically allow the raw audio to be shared.
- Even for the solo researcher, data structuring feels like reinventing the wheel every time.
- Lacking a common corpus analysis framework can make comparability across research groups challenging.



SPADE
SPeech Across Dialects of English

- 43+ datasets, 4 countries, 115 years
- heterogeneous corpus formats
- **to date 10 datasets are imported & measurements from 6 datasets have been generated ☺**

Vowels

It has been suggested that when a vowel is involved in a phonetic change in a speech community, this will be reflected in individuals' phonetic distributions, as they will spread out along the same axis as the change (Labov 1994).

Let's see if it pans out. We'll look at the "axis of variation" of speakers' vowel distributions.

Vowels

- The vowel formant data was extracted using a FAVE-like system.
  - Dialect specific initial prototypes based on hand measurements.
  - F1, F2 and F3 chosen from all formants at 5 candidate pole settings.
  - Speaker-specific re-estimation, iterating until results were identical between iterations.
  - Measured at 1/3 of duration.
- We only used words where there is shared lexical incidence across the varieties examined (Unisyn).

Vowels

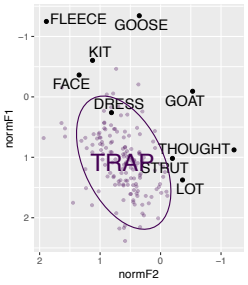
Corpus	ICECAN	Buckeye	SOTC	SCOTS
Dialect	Canada	Columbus	Glasgow	Scotland
N Vowels	11.6k	115.3k	105k	87.1k

Corpus	Raleigh	Santa Barbara	
Dialect	Raleigh	West	Northern Cities
N Vowels	162.8k	16.9k	6.1k

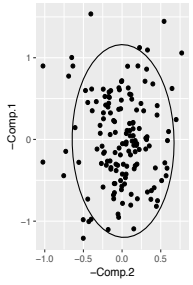
Vowels – Getting Axes of Variation

A representative speaker from Raleigh (Woman, DOB 1942)

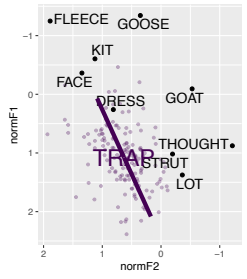
Original TRAP distribution



Principal components analysis rotates the data.

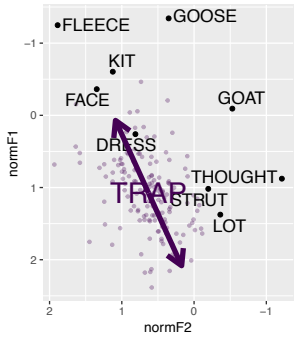


We can get the axis of variation for the original data.



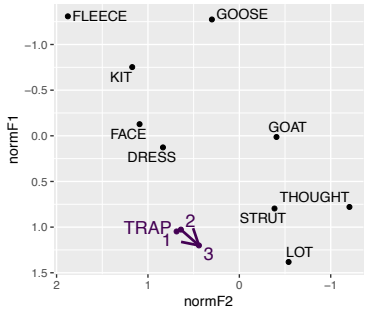
Vowels – Analyzing Axis of Variation

Is the angle across which this speaker's TRAP vowel is distributed the *same* angle in which TRAP is changing over time in Raleigh?

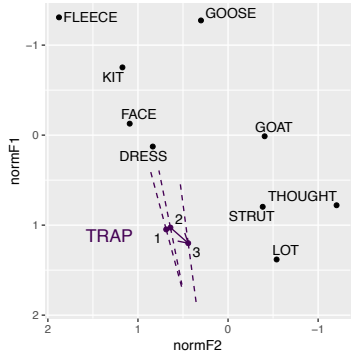


Vowels – Getting Axis of Variation

The cross generational change in Raleigh for TRAP is lowering and retracting, somewhat.

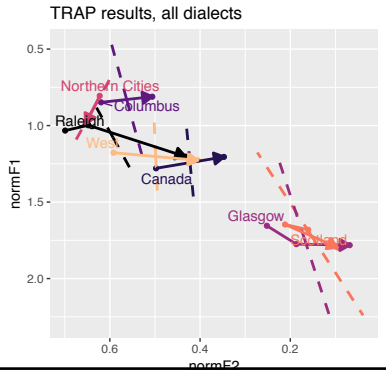


Vowels – Comparing Axis of Variation



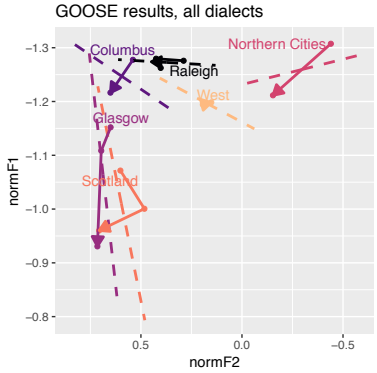
In Raleigh, the average axis of variation of speakers' distributions in each generation appears to be unrelated to the direction of cross generational change.

TRAP Results – All Dialects



Most varieties appear to have some cross generational TRAP retraction, but the average axis of variation of individuals' distributions is mostly vertical.

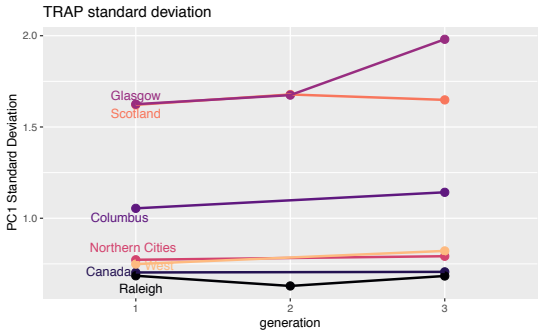
GOOSE Results, across dialects



On the other hand, most varieties have some form of GOOSE fronting, and the axis of variation more horizontal.

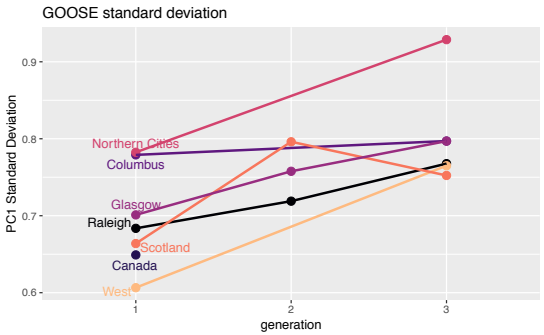
Exceptions are from the Scottish Corpora, which have both the frontest GOOSE, and are lowering.

TRAP – Spread along axis



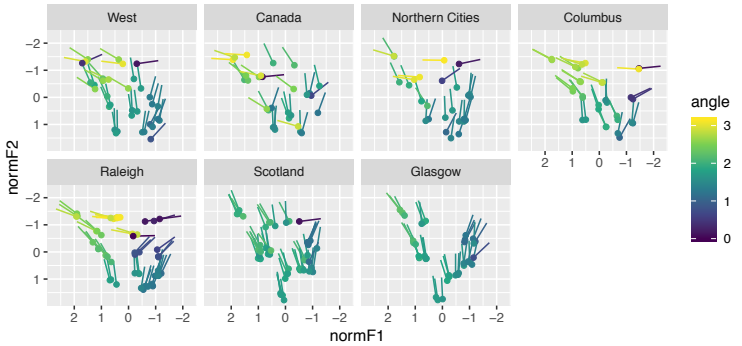
The degree to which TRAP is spread out along its primary axis is fairly stable across generations.

GOOSE – Spread along axis



GOOSE, on the other hand, seems to have an increasing spread along its primary axis over generations.

Axes – A broader view.



Why these axes?

For most vowels, the axis of variation is vertical. This could be understood in terms of variation in jaw opening.

The fact that GOOSE (and sometimes FLEECE) are more horizontally angled requires more investigation.

Conclusion

Weak evidence for individuals reflecting the trajectory of change within their own phonetic distributions. For the most part, their axis and spread seem to be a product of where in the vowel space they are located.

But, there is some evidence of dialect specificity with respect to both angle and spread.

Conclusion

A unified corpus analysis strategy, in combination with utilizing proper database methods, can help us investigate open conjectures about sound change.

Corpora Citations

Buckeye	Pitt, Mark A., Laura Dilley, Keith Johnson, Scott Kiesling, William Raymond, Elizabeth Hume, and Eric Fosler-Lussier. 2007. Buckeye Corpus of conversational speech (2nd release) [www.buckeyecorpus.osu.edu]. Columbus, OH.
ICECAN	Greenbaum, S., and G. Nelson. 1996. The international corpus of English (ICE) project. *World Englishes* 15:3–15.
Raleigh	Dodsworth, Robin, and Mary Kohn. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change* 24:221–245.
Santa Barbara	Du Bois, J. W., W. L. Chafe, C. Meyer, S. A. Thompson, and N. Martey. 2000. Santa Barbara Corpus of Spoken American English. *Linguistic Data Consortium* CD-ROM.
SOTC	Stuart-Smith, Jane. 2014. Fine phonetic variation and sound change: A real-time study of Glaswegian. Final Report: RPG-142 (Sounds of the City).
SCOTS	Anderson, J., D. Beavan, and C. Kay. 2007. SCOTS: Scottish corpus of texts and speech. In *Creating and digitizing language corpora*, 17–34. Palgrave Macmillan UK.


Thank you!
Please let us know if you would like us to work with your data
http://spade.glasgow.ac.uk/
arts-spade@glasgow.ac.uk
Trans-Atlantic Platform
Social Sciences and Humanities
DIGGING INTO DATA
CHALLENGE
Arts & Humanities Research Council
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
NSERC CRSNG
SSHRC CRSH
Social Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada
NSF