Order of Operations in Sociophonetic Data Analysis

Joseph A. Stanley

Brigham Young University

joey_stanley@byu.edu @joey_stan

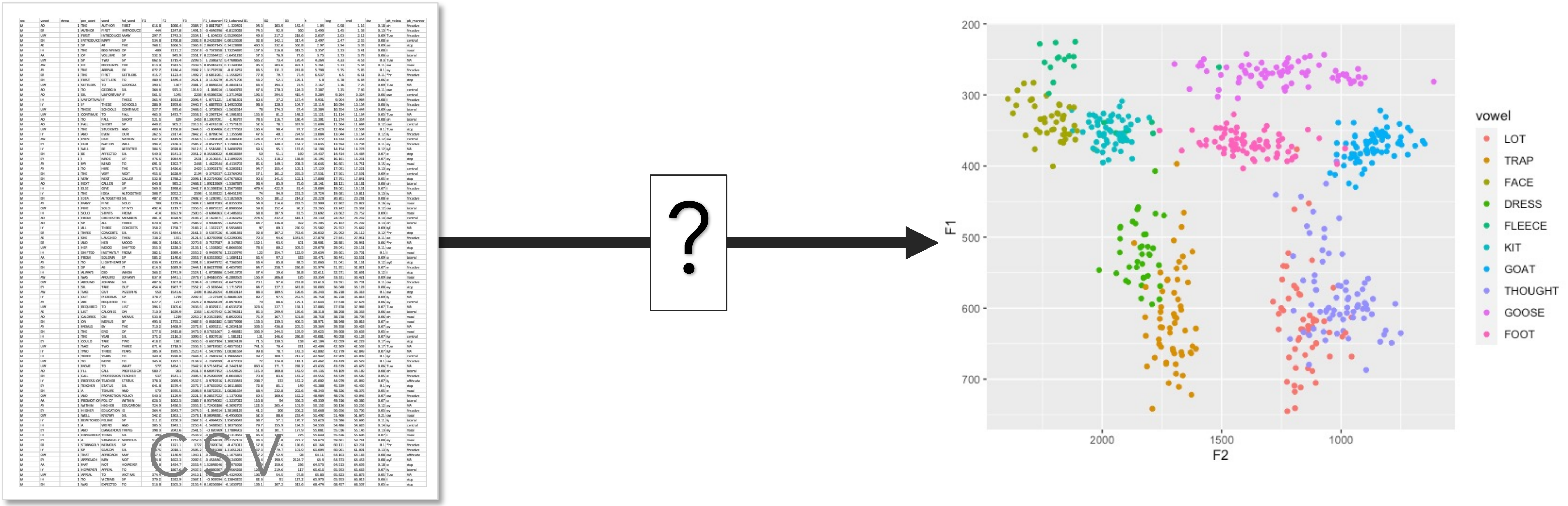
joeystanley.com

New Ways of Analyzing Variation 49

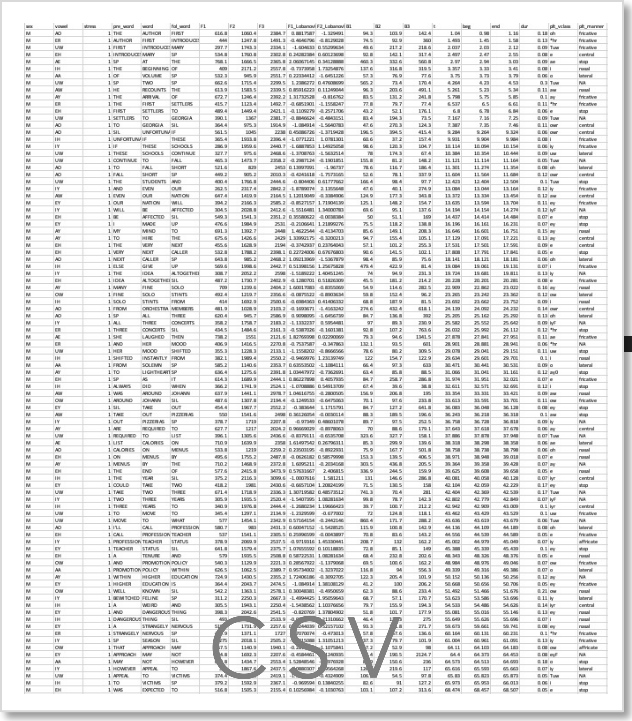
(Online)

October 2021

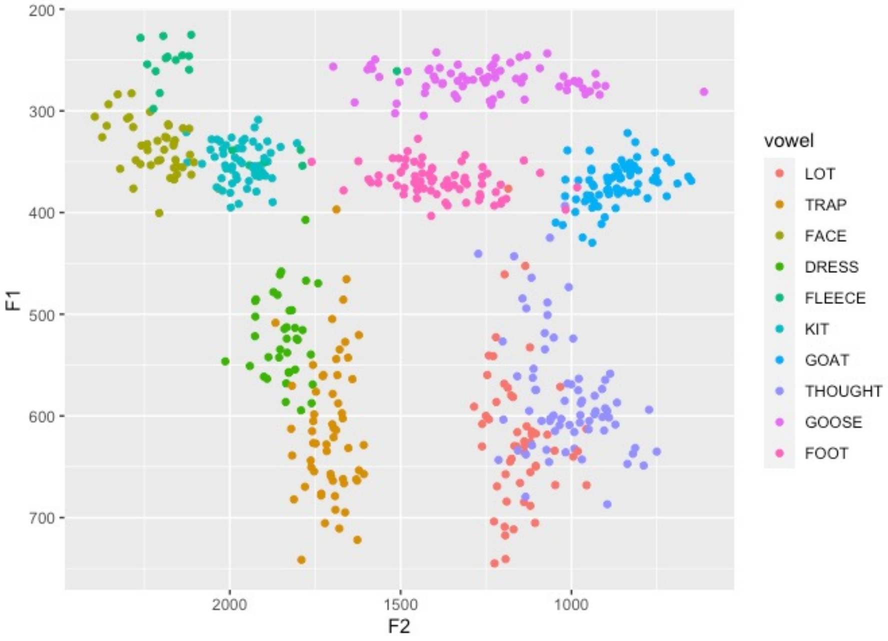
The Sociophonetic Pipeline



The Sociophonetic Pipeline



- remove outliers
- no stopwords
- no unstressed vowels
- normalization
- (no presonorants)
- no diphthongs
- midpoints only



Comparing Data Analysis Methods

Overlap

- Pillai is probably best (Nycz & Hall-Lew 2013)
- Well no, Bhttacharyya's affinity is better (Johnson 2015)
- Well, Pillai may be better at least for small, idealized data (Kelley & Tucker 2020)

Normalization

- We tested 12 ways and Lobanov is best (Adank et al. 2004)
- Adank et al miscalculated and Lobanov is actually not great (Barreda 2021a *LVC*)

Statistical Analysis

- Ditch Goldvarb and use mixed-effects models (Johnson 2009, 2014)
- Goldvarb still has its uses (Tagliamonte & D'Arcy 2017)

Formant estimation

- Different Praat settings can produce different results (Kendall & Vaughn 2020)
- Find the best Praat settings (Barreda 2021b *Ling Van.*)

Order of Operations

Lots of debate on which individual method is best. (Good!)

- We all know that we need to run the same steps (Good!)
- We all know we need to report details of the pipeline (Good!)

To my knowledge, no discussion of the order of operations

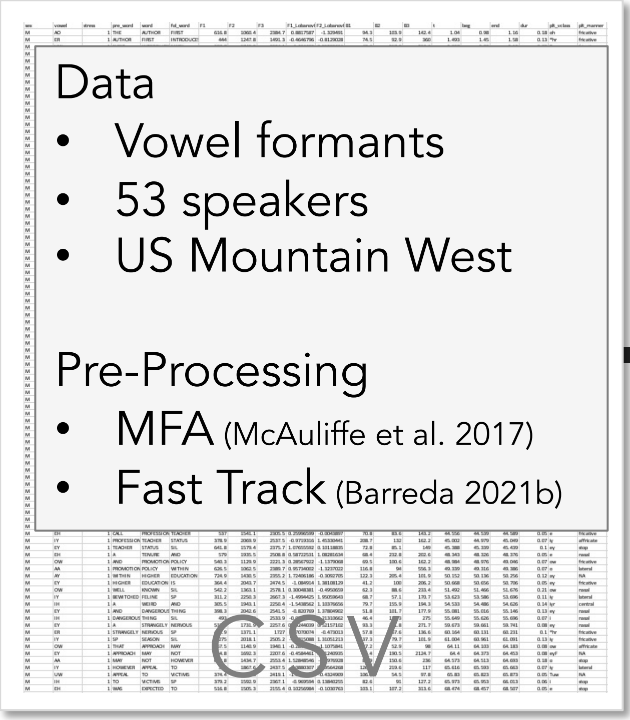
- Some configurations make no difference
- Other configurations make a sizeable difference

Hypothesis:

- We can get substantially different interpretations of the data depending on the order of operations.
- Order of operations matters!

Methods

Methods

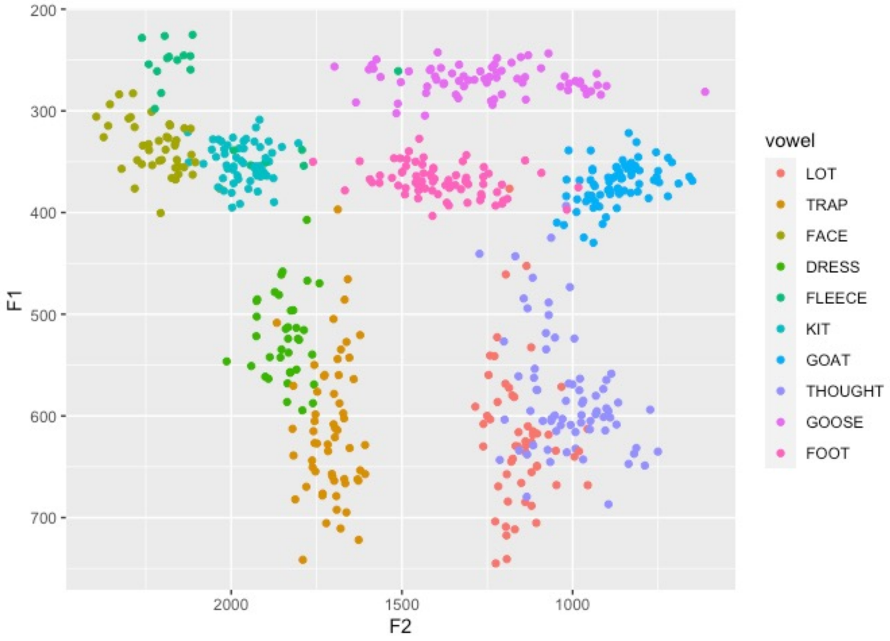


Data
- Vowel formants
- 53 speakers
- US Mountain West

Pre-Processing
- MFA (McAuliffe et al. 2017)
- Fast Track (Barreda 2021b)

Permutation 1:

1. remove outliers
2. no stopwords
3. no unstressed vowels
4. normalize
5. (no presonorants)
6. no diphthongs
7. midpoints only



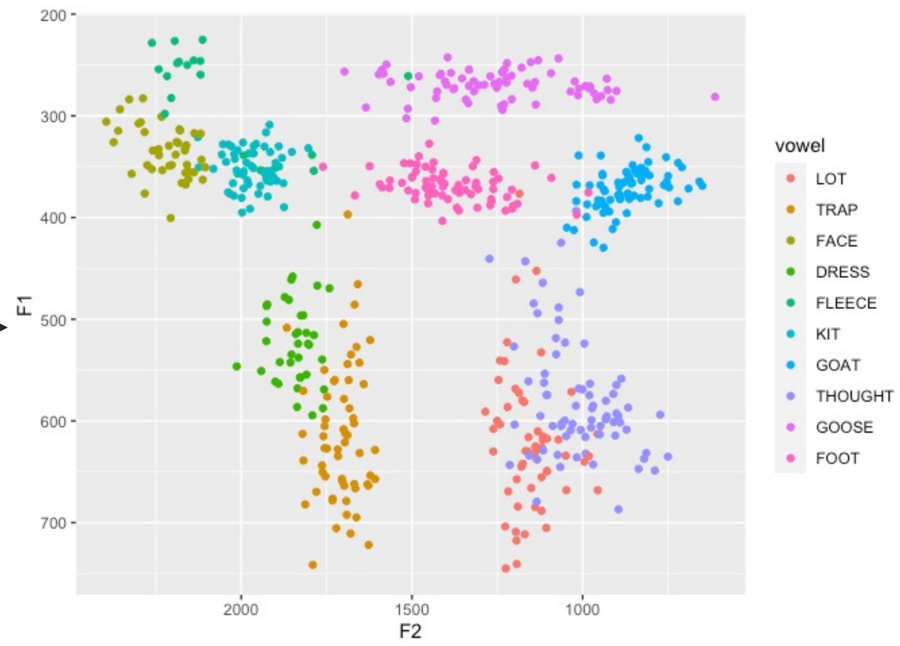
Methods

Data
- Vowel formants
- 53 speakers
- US Mountain West

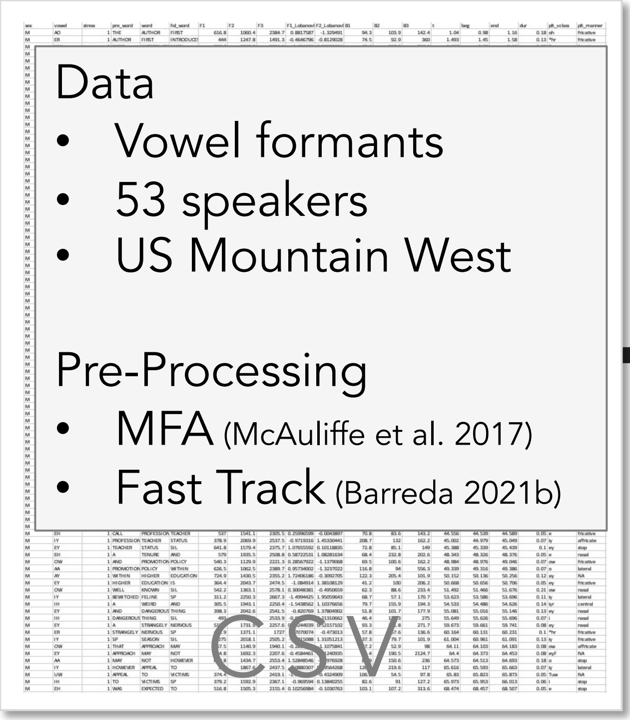
Pre-Processing
- MFA (McAuliffe et al. 2017)
- Fast Track (Barreda 2021b)

CSV

Permutation 2:
1. remove outliers
2. no stopwords
3. no unstressed vowels
4. normalize
5. (no presonorants)
6. midpoints only
7. no diphthongs



Methods

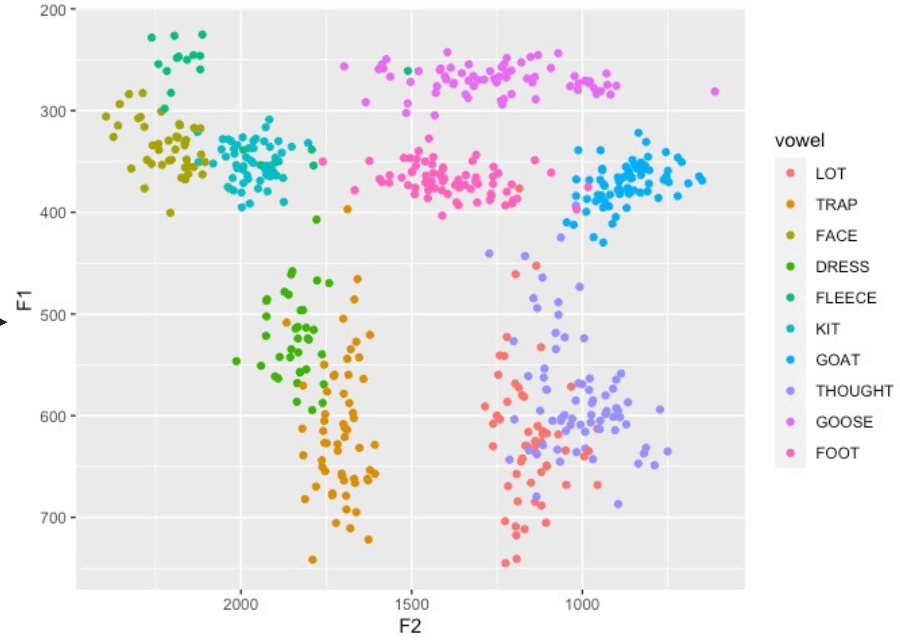


Data
- Vowel formants
- 53 speakers
- US Mountain West

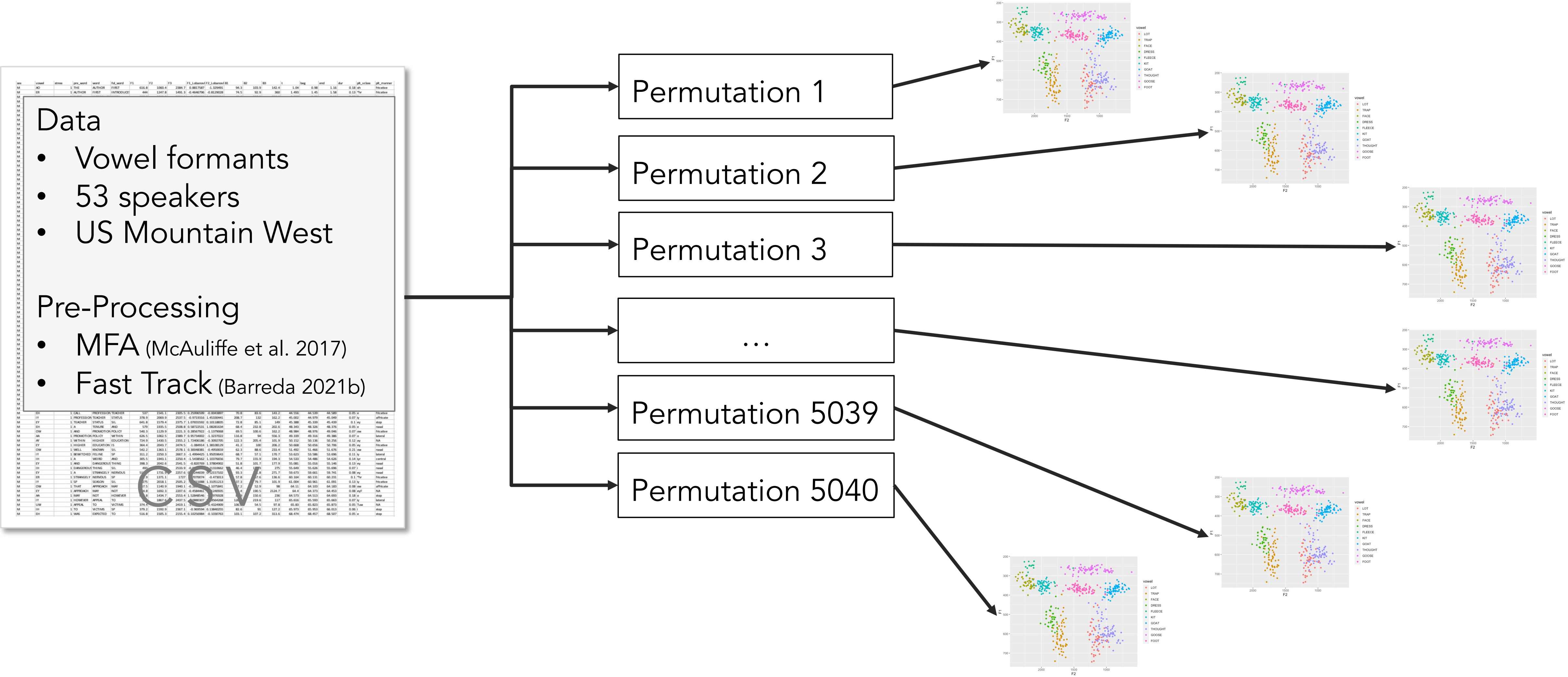
Pre-Processing
- MFA (McAuliffe et al. 2017)
- Fast Track (Barreda 2021b)

Permutation 3:

1. remove outliers
2. no stopwords
3. no unstressed vowels
4. normalize
5. midpoints only
6. (no presonorants)
7. no diphthongs



Methods



Data
• Vowel formants
• 53 speakers
• US Mountain West
Pre-Processing
• MFA (McAuliffe et al. 2017)
• Fast Track (Barreda 2021b)
CSV
Permutation 1
Permutation 2
Permutation 3
…
Permutation 5039
Permutation 5040

Three Experiments

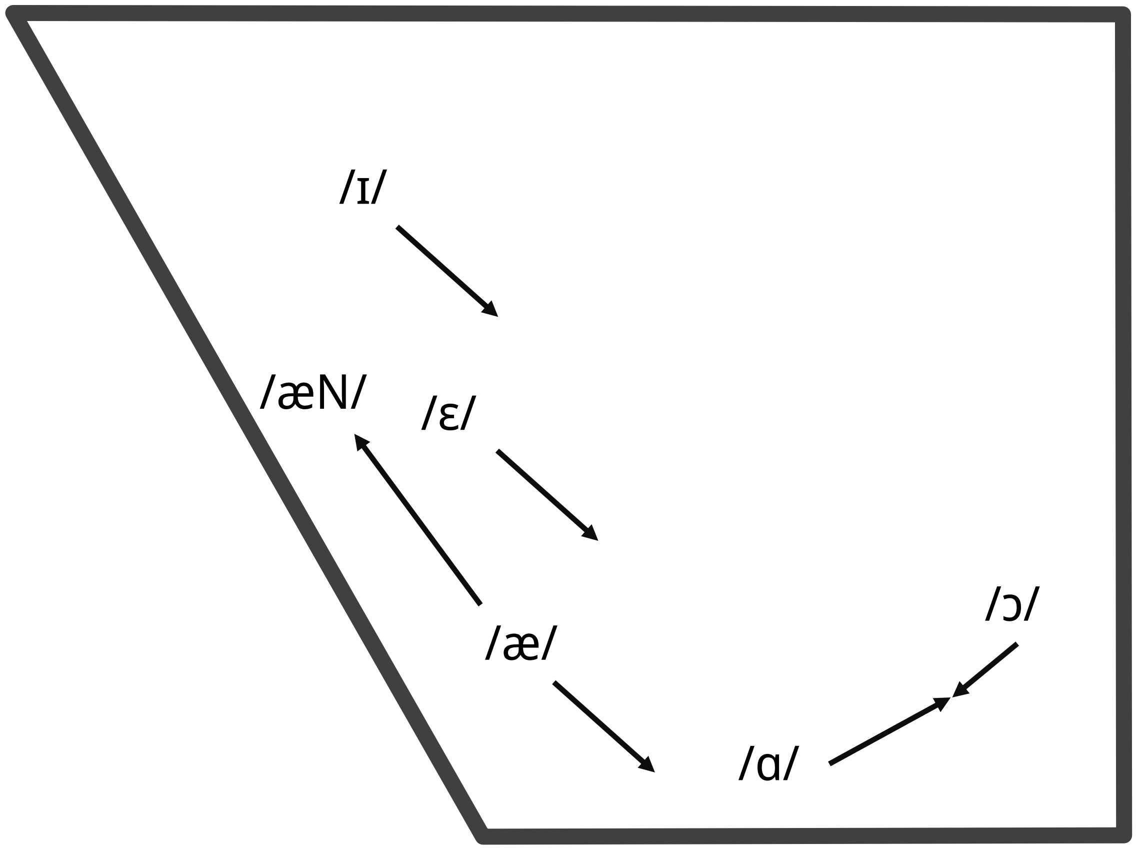
Low-Back-Merger Shift (Becker 2019)

- Includes the low back merger
- Front lax vowels lower and centralize

Merger and shifts are difficult to quantify

How do order of operations affect three common metrics?

- Experiment 1: Pillai scores
- Experiment 2: ANAE "benchmarks"
- Experiment 3: LBMS Index



Experiment 1: Pillai Scores

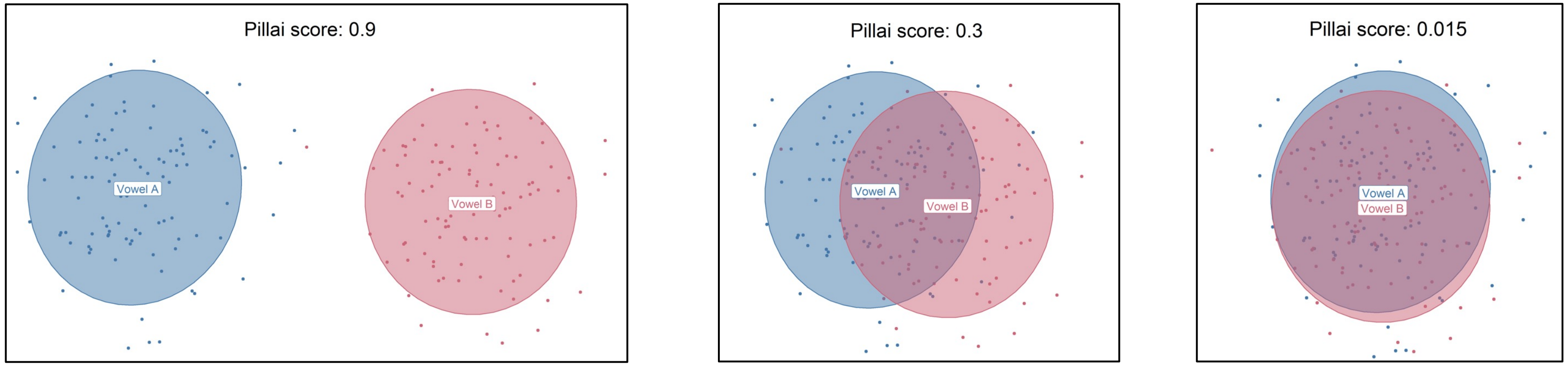
Measuring Vowel Overlap

Mergers have always been difficult to measure

- Simple solution: Does someone realize two historically different vowels in the same space?
- However, F1-F2 isn't everything (voice quality, length, trajectories, speaker intuition)

Most common solution: Pillai scores (Hay et al. 2006)

- Output of MANOVA test: 0 = complete overlap, 1 = complete separation



Pillai Scores

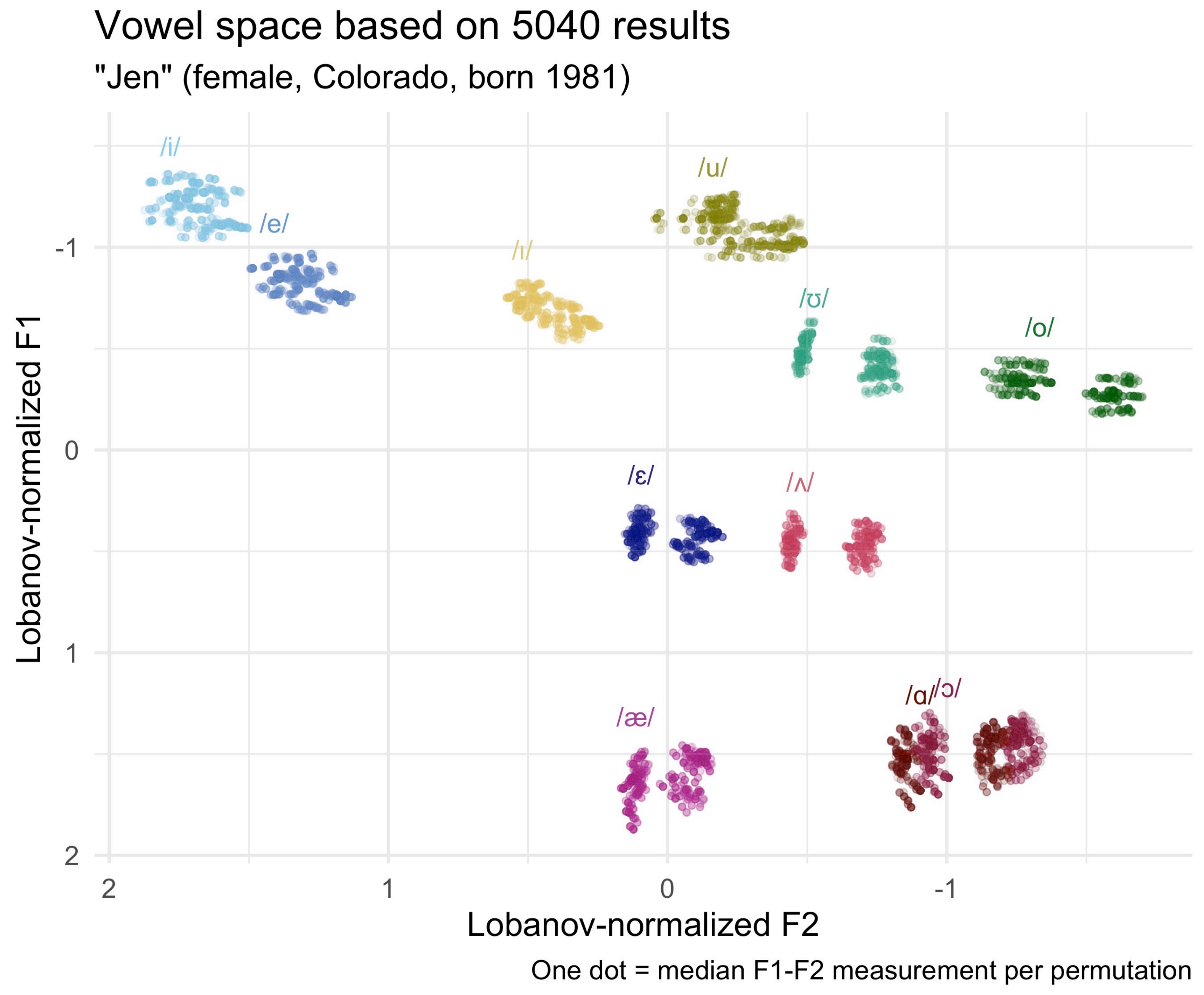
Two vowel pairs

- the low back merger (/ɑ/ vs. /ɔ/)
- the ”prenasal split” (raising of /æN/ compared to pre-obstruent /æ/)

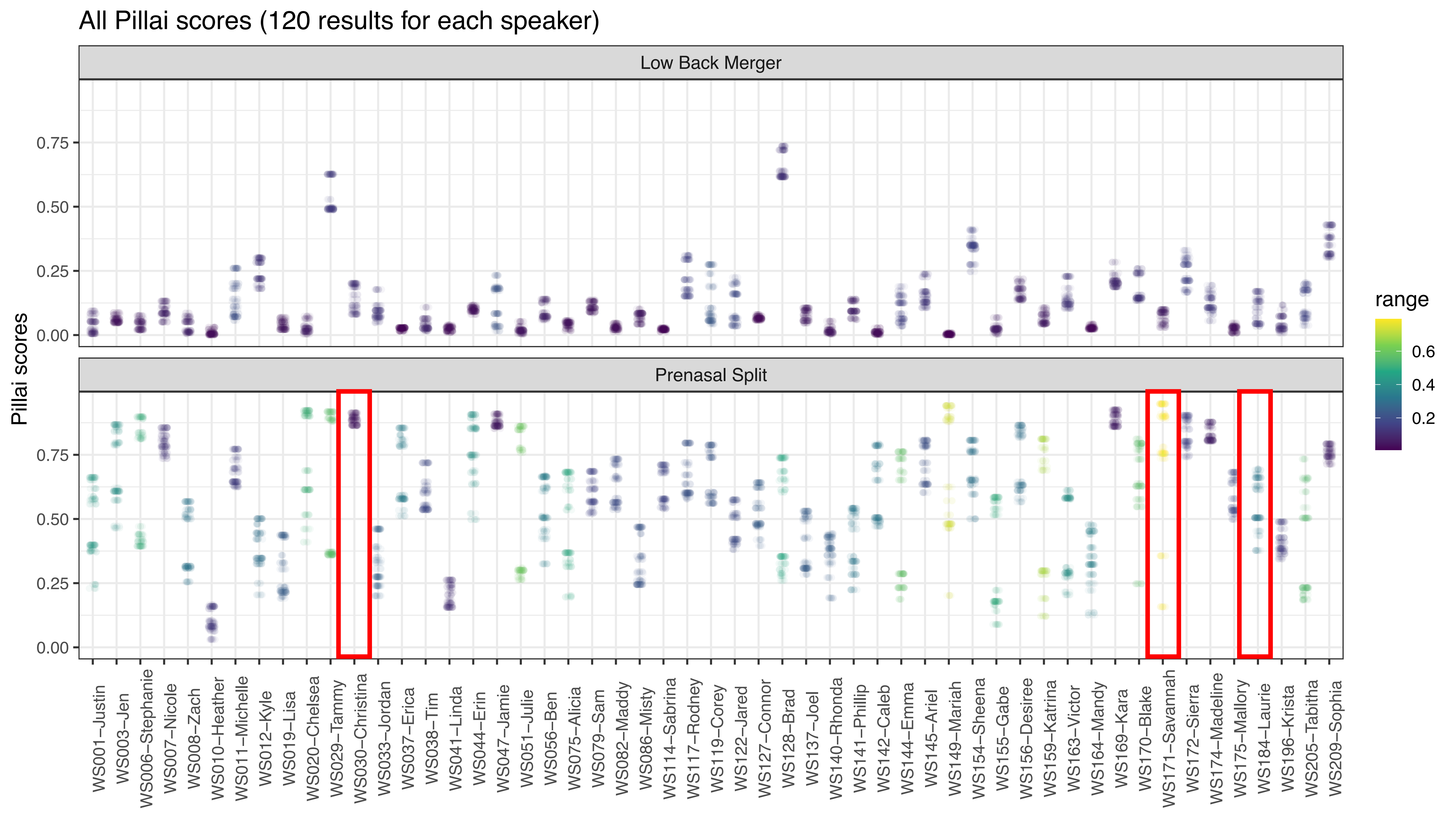
Steps in the pipeline

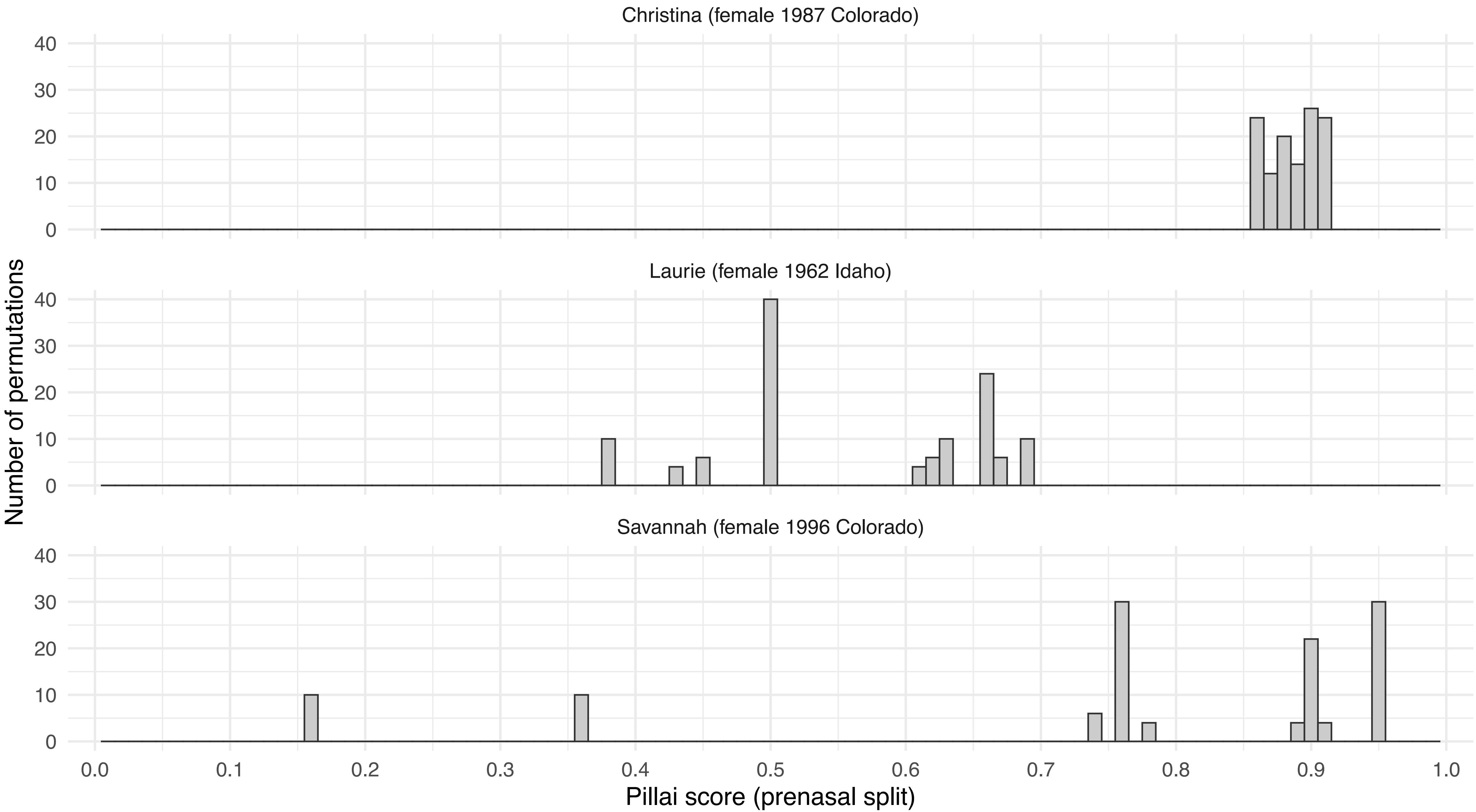
- remove stopwords
- remove outliers
- isolating midpoints from trajectories
- remove unstressed tokens
- For low-back merger: remove presonorant tokens.
- For prenasal split: defining vowel classes (i.e. treat /æN/ differently than /æ/)

Five steps = 120 permutations



Vowel space based on 5040 results
"Jen" (female, Colorado, born 1981)
/i/
/e/
/ɪ/
/u/
/ʊ/
/o/
/ɛ/
/ʌ/
/æ/
/ɑ/
/ɔ/
-1
0
1
2
2
1
0
-1
Lobanov-normalized F1
Lobanov-normalized F2
One dot = median F1-F2 measurement per permutation


All Pillai scores (120 results for each speaker)
Low Back Merger
Prenasal Split
Pillai scores
0.00
0.25
0.50
0.75
range
0.2
0.4
0.6
WS001–Justin
WS003–Jen
WS006–Stephanie
WS007–Nicole
WS008–Zach
WS010–Heather
WS011–Michelle
WS012–Kyle
WS019–Lisa
WS020–Chelsea
WS029–Tammy
WS030–Christina
WS033–Jordan
WS037–Erica
WS038–Tim
WS041–Linda
WS044–Erin
WS047–Jamie
WS051–Julie
WS056–Ben
WS075–Alicia
WS079–Sam
WS082–Maddy
WS086–Misty
WS114–Sabrina
WS117–Rodney
WS119–Corey
WS122–Jared
WS127–Connor
WS128–Brad
WS137–Joel
WS140–Rhonda
WS141–Phillip
WS142–Caleb
WS144–Emma
WS145–Ariel
WS149–Mariah
WS154–Sheena
WS155–Gabe
WS156–Desiree
WS159–Katrina
WS163–Victor
WS164–Mandy
WS169–Kara
WS170–Blake
WS171–Savannah
WS172–Sierra
WS174–Madeline
WS175–Mallory
WS184–Laurie
WS196–Krista
WS205–Tabitha
WS209–Sophia



Christina (female 1987 Colorado)
Laurie (female 1962 Idaho)
Savannah (female 1996 Colorado)
Number of permutations
Pillai score (prenasal split)
0
10
20
30
40
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Experiment 2: ANAE Benchmarks

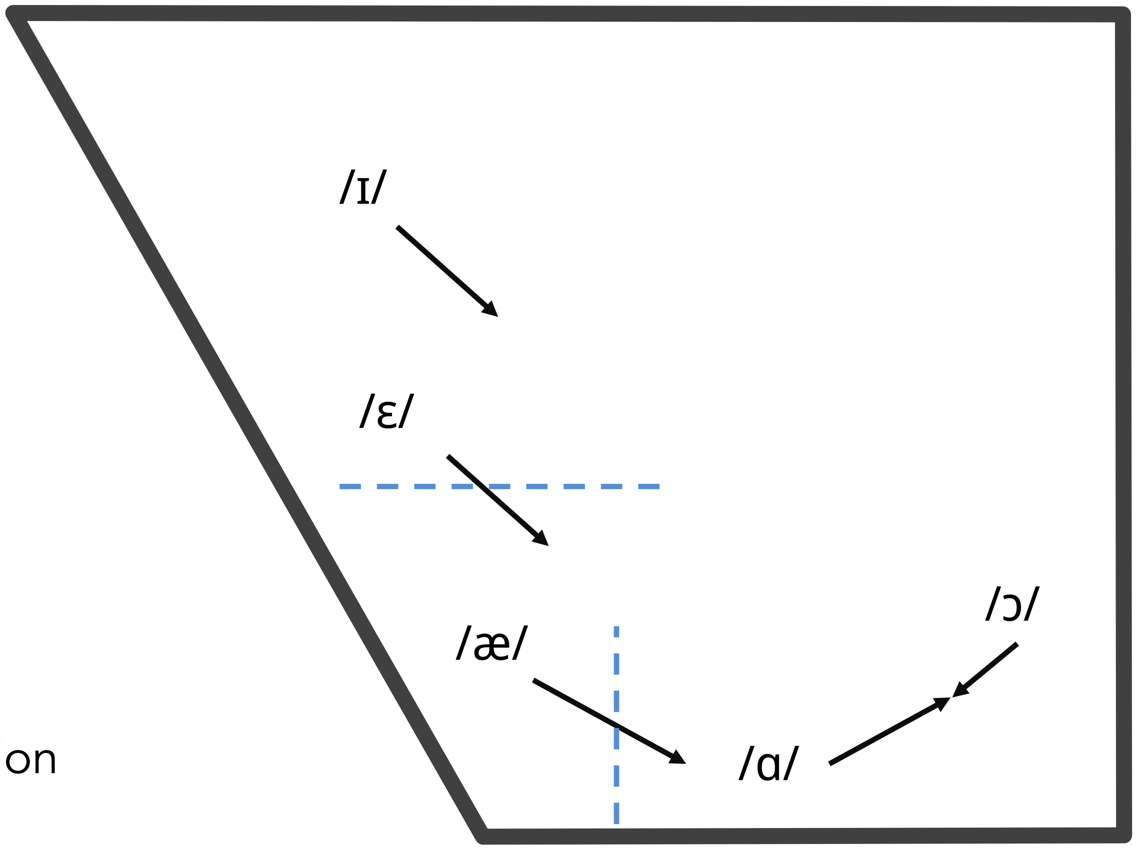
How to measure shifting vowels?

It's difficult to say objectively whether a vowel is shifted

- Relationship to other vowels (which may also be shifting)
- Benchmarks don't work because of normalization and dialect differences

Most typical method: compare to supposed "benchmarks"

- *Atlas of North American English* (Labov, Ash, & Boberg 2006)
- Using these benchmarks implies ANAE normalization

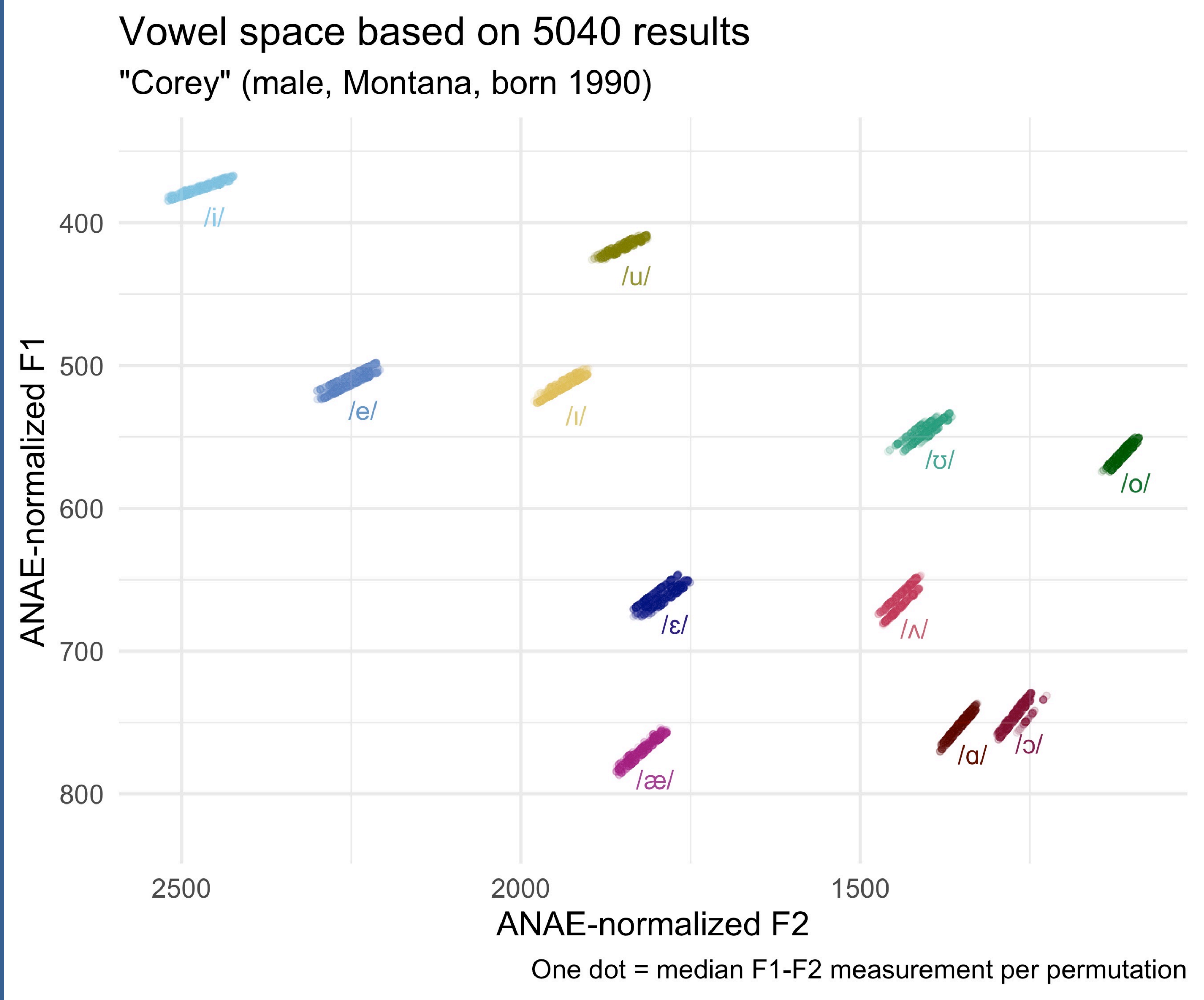


Compare to ANAE Benchmarks

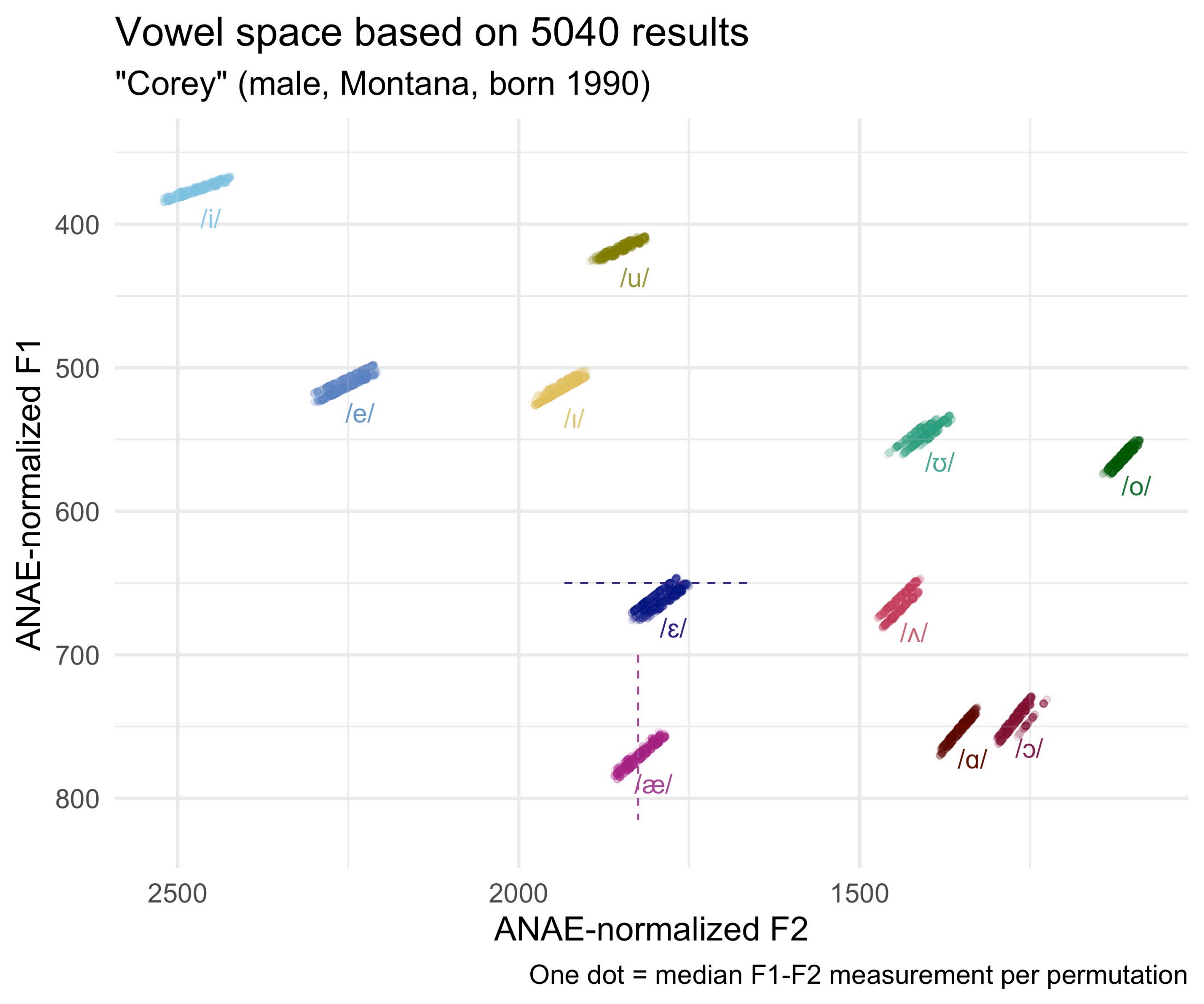
More processing involved

- remove stopwords
- remove outliers
- remove diphthongs
- remove presonorants
- isolating midpoints
- a specific normalization procedure
- remove unstressed vowels

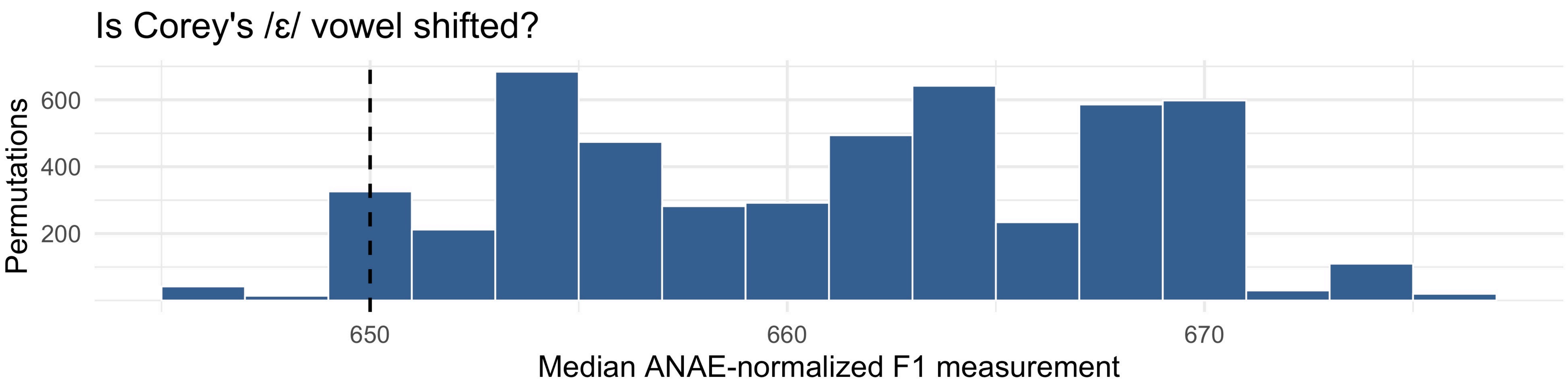
Seven steps = 5040 combinations

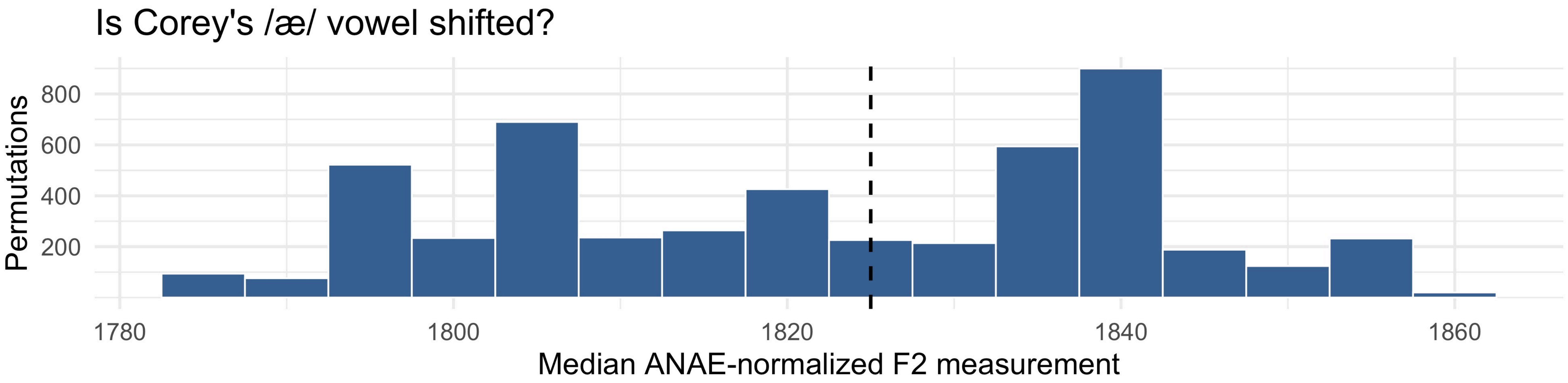


Vowel space based on 5040 results
"Corey" (male, Montana, born 1990)
ANAE-normalized F1
ANAE-normalized F2
400
500
600
700
800
2500
2000
1500
/i/
/u/
/e/
/ɪ/
/ʊ/
/o/
/ɛ/
/ʌ/
/ɑ/
/ɔ/
/æ/
One dot = median F1-F2 measurement per permutation



Vowel space based on 5040 results
"Corey" (male, Montana, born 1990)
ANAE-normalized F1
ANAE-normalized F2
400
500
600
700
800
2500
2000
1500
/i/
/u/
/e/
/ɪ/
/ʊ/
/o/
/ɛ/
/ʌ/
/æ/
/ɑ/
/ɔ/
One dot = median F1-F2 measurement per permutation


Is Corey's /ɛ/ vowel shifted?
Permutations
600
400
200
650
660
670
Median ANAE-normalized F1 measurement


Is Corey's /æ/ vowel shifted?
Permutations
800
600
400
200
1780
1800
1820
1840
1860
Median ANAE-normalized F2 measurement

ANAE Benchmark Results

If shifting is measure by those benchmarks, the results varied widely

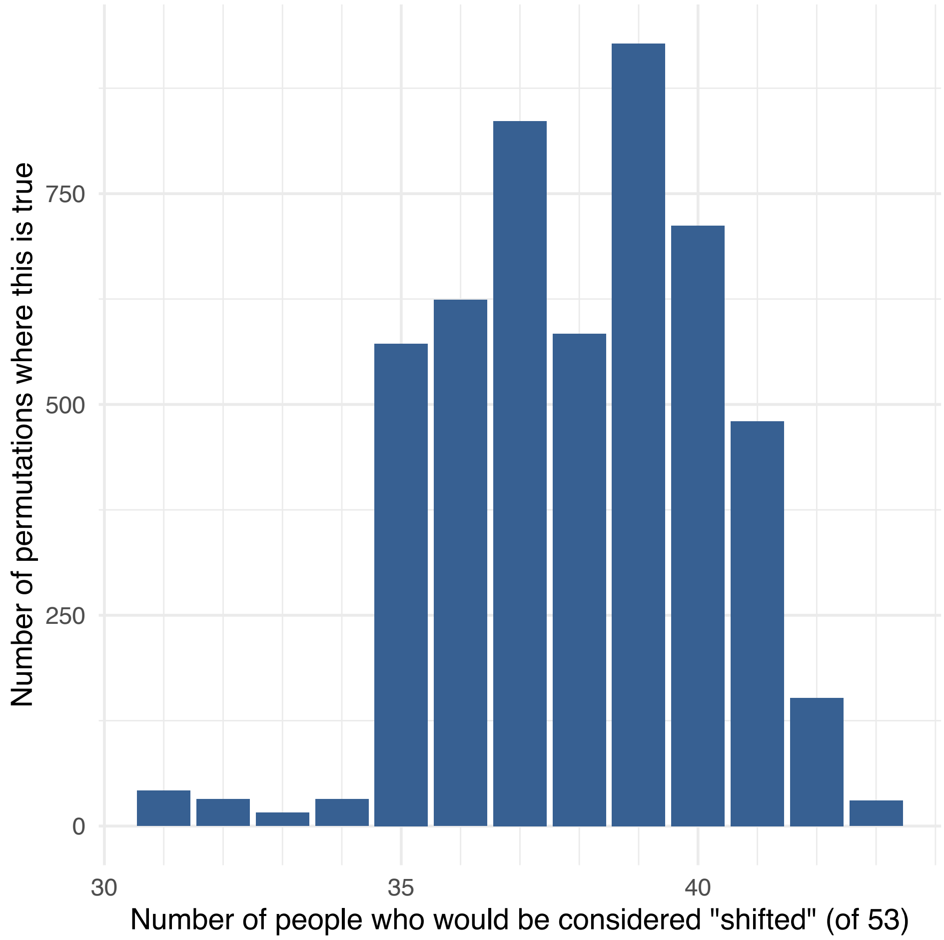
Individuals varied

- For 20 people their overall status of shifted vs non-shifted was the same in all 5040 combinations
- But for 33 people, whether they were shifted or not depended on the permutation.

Overall results varied

- 31 people were considered shifted in some permutations
- 43 people were in others

So how many people in my sample are shifted??



Experiment 3: LBMS-Index

Another measure for shifting vowels

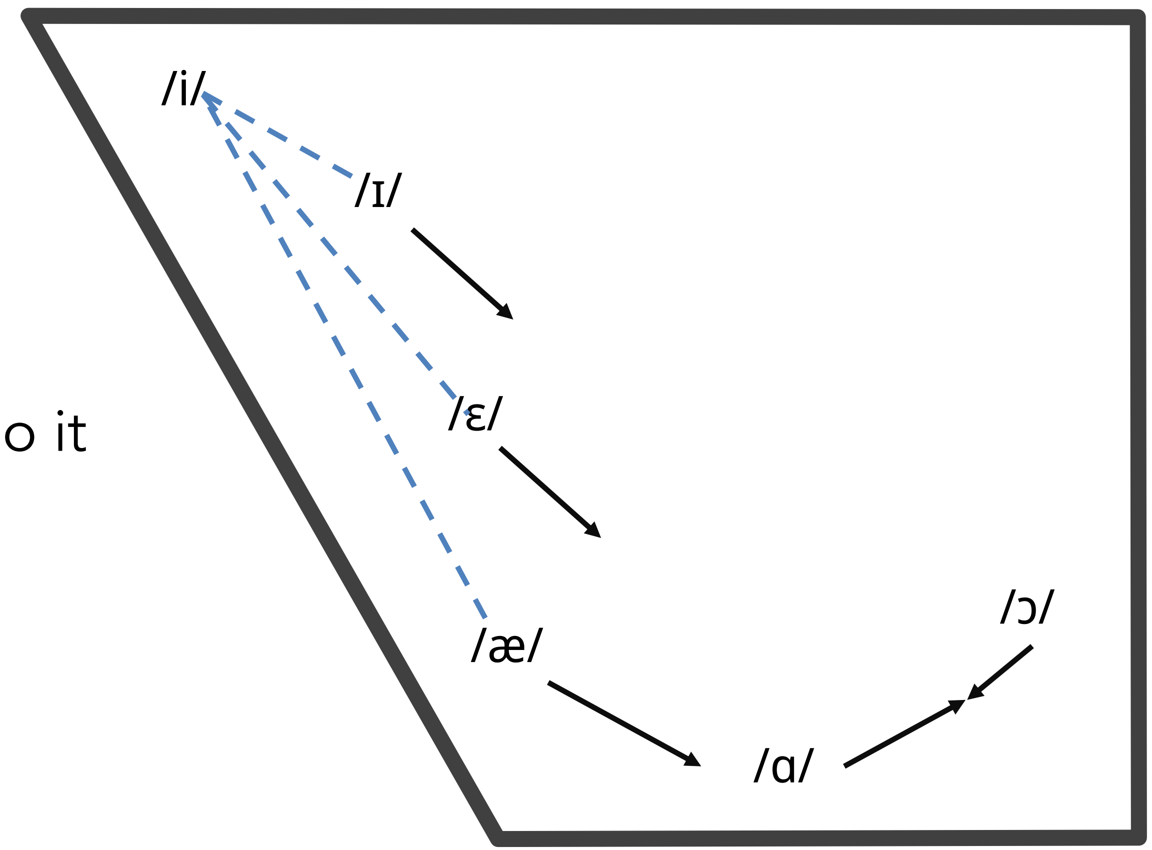
The "LBMS Index" (Becker 2019)

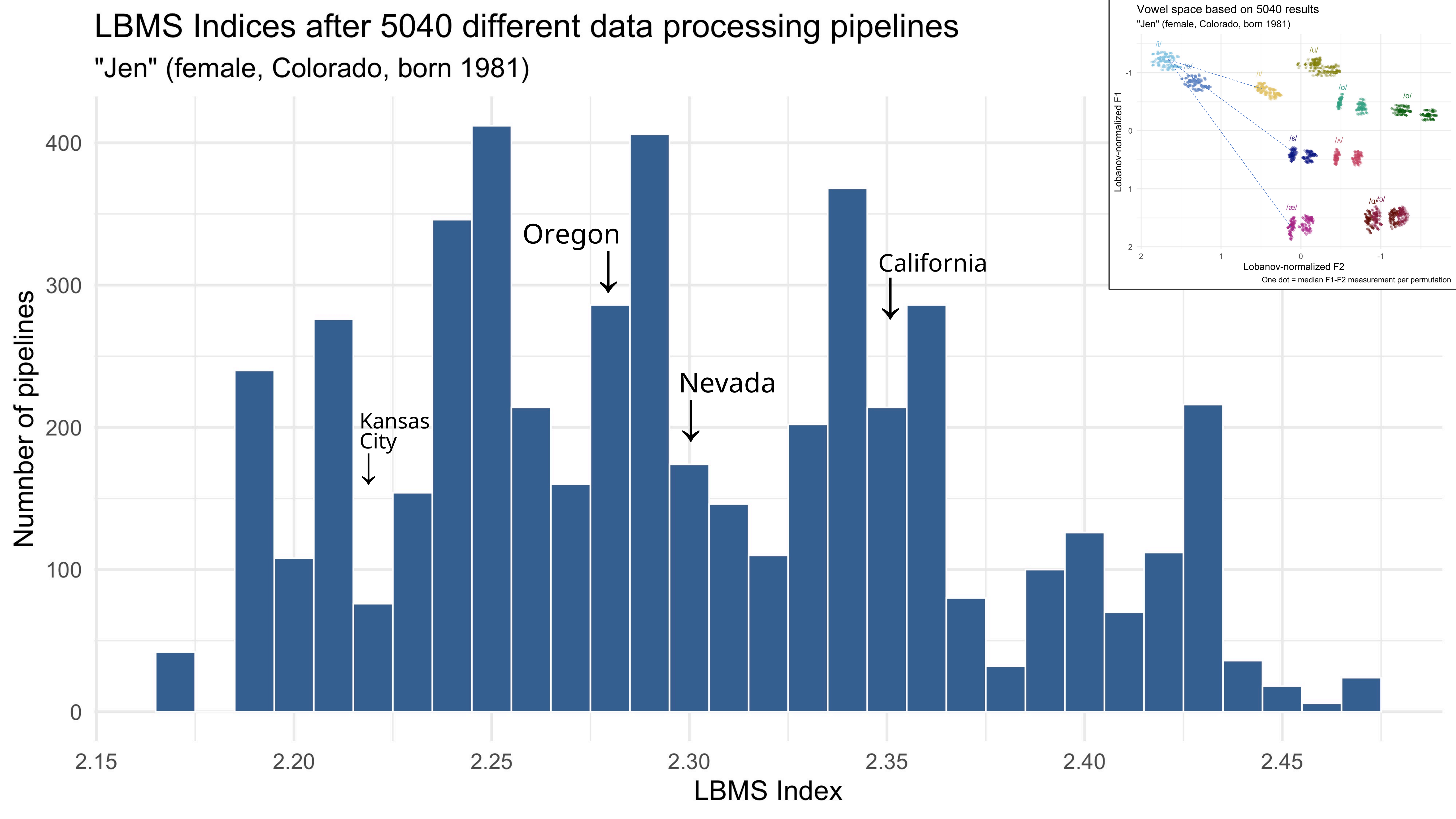
- Distance between /ɪ/ and /i/
- Distance between /ɛ/ and /i/
- Distance between /æ/ and /i/
- Average these three

Becker (2019) provides guidelines on how to do it

- Use Lobanov normalization
- Remove presonorants
- The other steps in the pipeline are assumed
- No order is specified.

Same steps as Experiment 2: 5040 combos




LBMS Indices after 5040 different data processing pipelines
"Jen" (female, Colorado, born 1981)
Numnber of pipelines
LBMS Index
0
100
200
300
400
2.15
2.20
2.25
2.30
2.35
2.40
2.45
Kansas City
Oregon
Nevada
California
Vowel space based on 5040 results
"Jen" (female, Colorado, born 1981)
Lobanov-normalized F1
Lobanov-normalized F2
One dot = median F1-F2 measurement per permutation

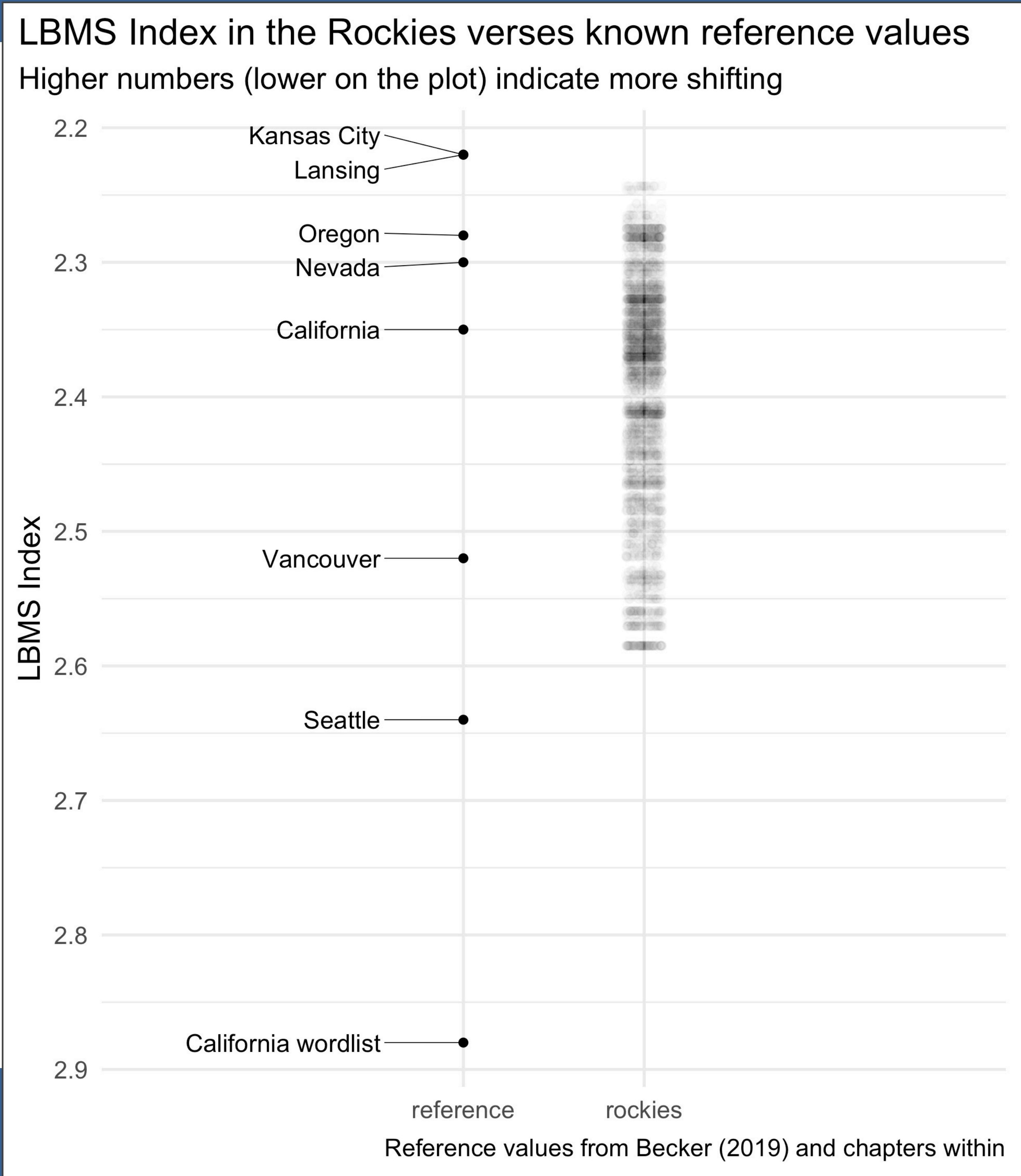
LBMS Results

Average the results by permutation

Wide range of results

- Least shifted: 2.24
- Most shifted: 2.58

A wide enough range to be considered sociolinguistically significant!



Discussion

Overview

Pillai scores

- Up to 16 outcomes
- Mostly small differences in low back, though some in /æN/-raising were large enough to matter

"Benchmarks" and LBMS Index

- Hundreds of outcomes for each person
- Differences were wide enough to reclassify people as "shifted" or not

In all three measures, the order of operations mattered

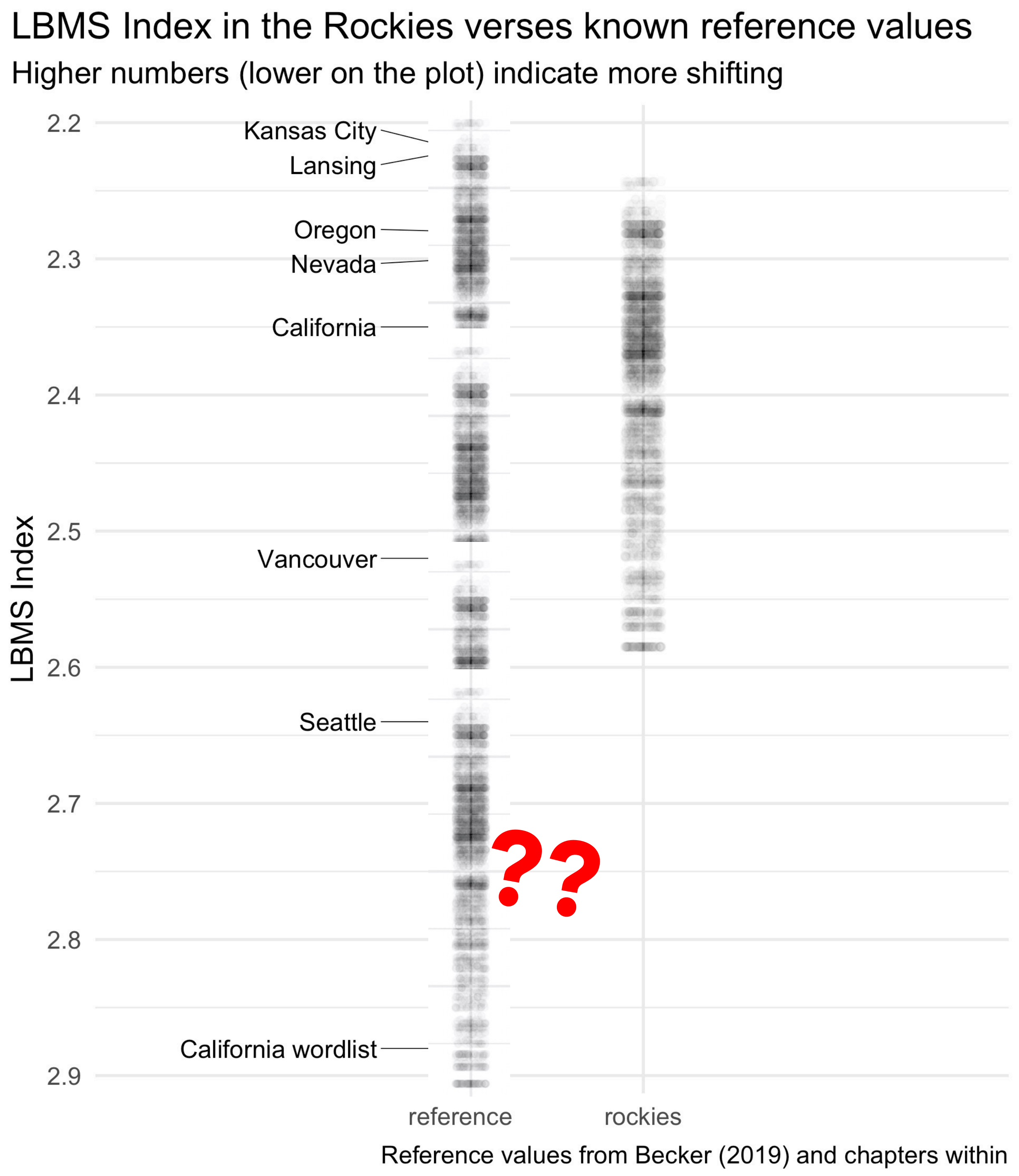
- The differences were the same magnitude as reported sociolinguistic differences
- Where normalization happens is the biggest factor

Who cares?

A naive analysis of the Rockies would be a random draw from the possible range of values.

Because order is not reported, all previous studies' results are random draws from their range of possible values!

Comparisons across studies are potentially meaningless.



Recommendations

Define a "proper" order of operations

- My recommendation:
  1. **Filter noise first** (stopwords before outliers)
  2. **Normalize**
  3. **Subset the data if necessary** (midpoints only, presonorants, monophthongs)
- Report this as part of methods
- Take numbers from studies that use different/unknown order with a grain of salt

Perhaps find different metrics

- Perhaps this instability is a sign that the metric is bad

Be smart about how we process our data

- More papers talking about methods

References

Adank, Patti, Roel Smits, and Roeland Van Hout. "2004. A Comparison of Vowel Normalization Procedures for Language Variation Research." *The Journal of the Acoustical Society of America* 116(5). 2004: 3099–3107.

Barreda, Santiago. 2021a. "Perceptual Validation of Vowel Normalization Methods for Variationist Research." *Language Variation and Change*, 1–27. https://doi.org/10.1017/S0954394521000016.

Barreda, Santiago. 2021b. "Fast Track: Fast, (Nearly) Automatic Formant-Tracking Using Praat." *Linguistics Vanguard* 7(1).

Becker, Kara, ed. 2019. *The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and Short Front Vowel Shifts across North America*. Publication of the American Dialect Society 104. Durham, NC: Duke University Press.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006. "Factors Influencing Speech Perception in the Context of a Merger-in-Progress." *Journal of Phonetics*, Modelling Sociophonetic Variation 34(4): 458–84. https://doi.org/10.1016/j.wocn.2005.10.001.

Johnson, Daniel Ezra. 2009. "Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis." *Language and Linguistics Compass* 3(1): 359–83.

Johnson, Daniel Ezra. 2014. "Progression in Regression: Why Natural Language Data Calls for Mixed-Effects Models." Self-published manuscript.

Johnson, Daniel Ezra. 2015. "Quantifying Vowel Overlap with Bhattacharyya's Affinity." Presented at the New Ways of Analyzing Variation (NWAV44), Toronto.

Kelley, Matthew C., and Benjamin V. Tucker. 2020. "A Comparison of Four Vowel Overlap Measures." *The Journal of the Acoustical Society of America* 147(1): 137–45. https://doi.org/10.1121/10.0000494.

Kendall, Tyler, and Charlotte Vaughn. 2015. "Measurement Variability in Vowel Formant Estimation: A Simulation Experiment." In *ICPhS*.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Walter de Gruyter.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger. 2017. "Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi." *Proceedings of the 18th Conference of the International Speech Communication Association*.

Nycz, Jennifer, and Lauren Hall-Lew. 2013. "Best Practices in Measuring Vowel Merger." *Proceedings of Meetings on Acoustics* 20(1): 060008. https://doi.org/10.1121/1.4894063.

Tagliamonte, Sali A., and Alexandra D'Arcy. 2017. "Individuals, Communities and the Sociolinguistic Canon." Presented at the New Ways of Analyzing Variation (NWAV) 46, Madison, WI.

Order of Operations in Sociophonetic Data Analysis

Joseph A. Stanley

Brigham Young University
joey_stanley@byu.edu @joey_stan

joeystanley.com

New Ways of Analyzing Variation 49
(Online)
October 2021