The SPeech Across Dialects of English (SPADE) Project

Opportunities for variationist research

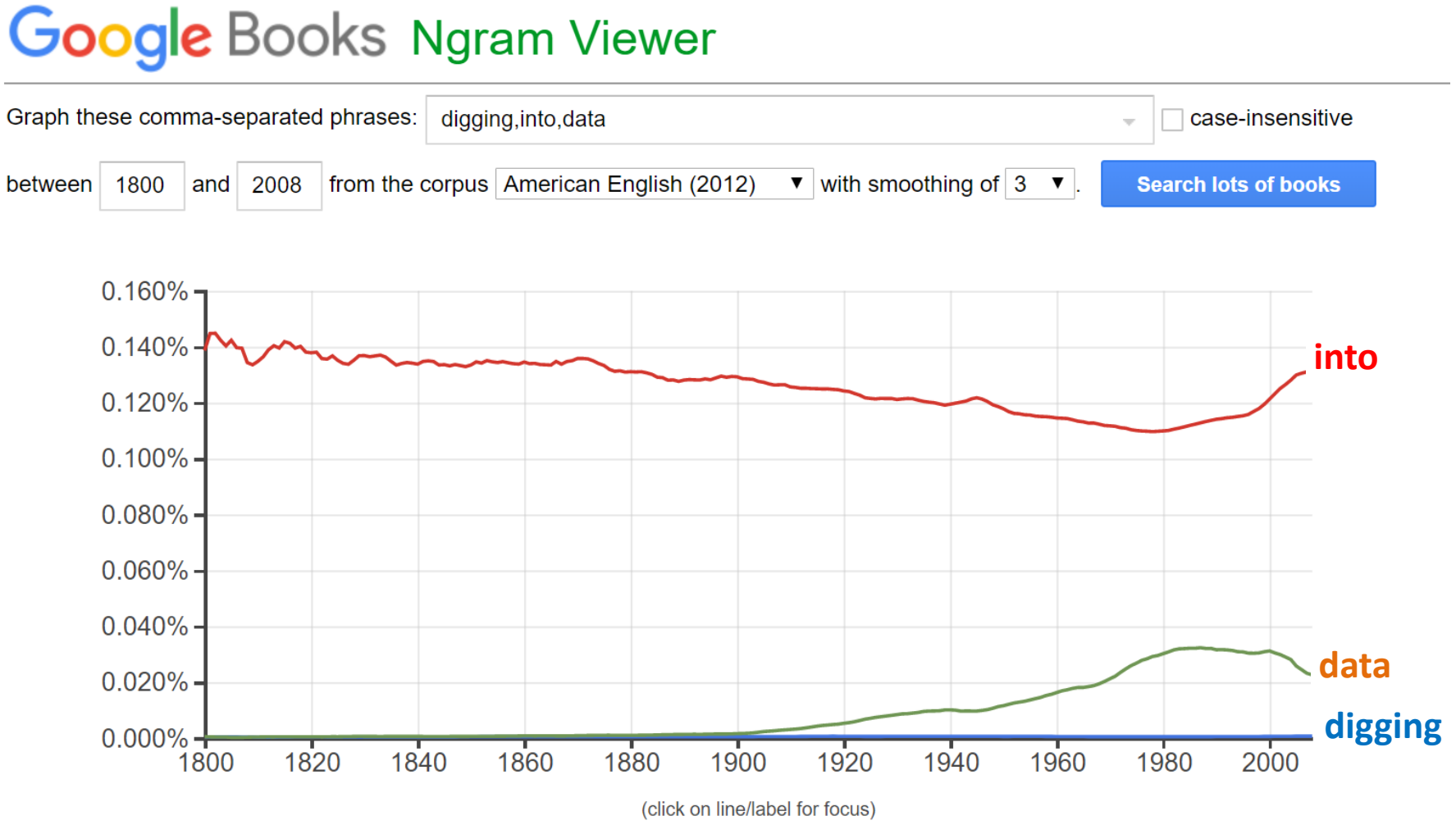
James Tanner & Jane Stuart-Smith

Unfashionable Language workshop, 13th March 2025

SPADE

SPeech Across Dialects of English

Text over time and space



`culturomics': Michel et al. (2011) *Science*

Huge amounts of annotated speech exist…

barriers

- $$ ££
- Software
- Ethics

Scientific questions

Goal: overcome barriers and scale up scientific study of speech

SPADE

SPeech Across Dialects of English

Jane Stuart-Smith (UK PI)

Morgan Sonderegger (Canada PI)

Jeff Mielke (US PI)

2017-2020…

http://spade.glasgow.ac.uk/

Trans-Atlantic Platform
Social Sciences and Humanities

DIGGING INTO DATA CHALLENGE

THE UNIVERSITY of EDINBURGH

University of Glasgow

McGill

NC STATE UNIVERSITY

UNIVERSITY OF OREGON

SPADE

SPeech Across Dialects of English

Project goals

Software large-scale speech analysis

Data from ~40 datasets (socio)linguistic surveys

Research

'English' sounds over time and space

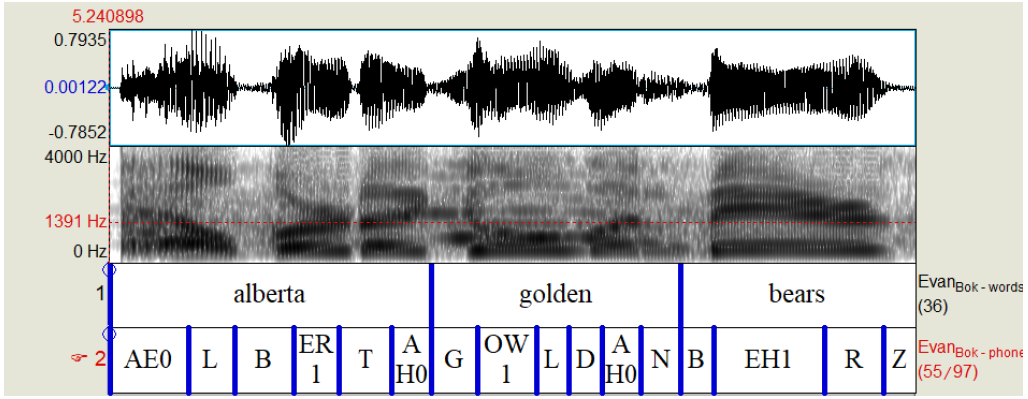
Sonderegger et al. (2021), *Open Handbook Ling. Data Management*

Data: The SPADE Consortium

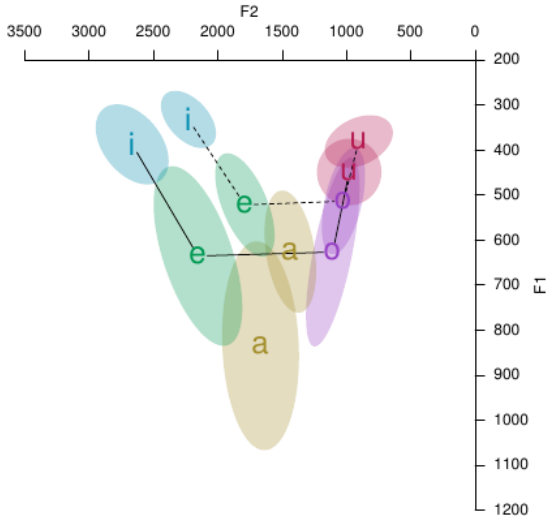
- 66 datasets: public/private, 4 countries, 115 years
- 10,086 speakers, ~2400 hours
- measures available at: https://osf.io/4jfrm/

- Typical speech research study:

Raw data



Analysis



scripts, manual entry, software

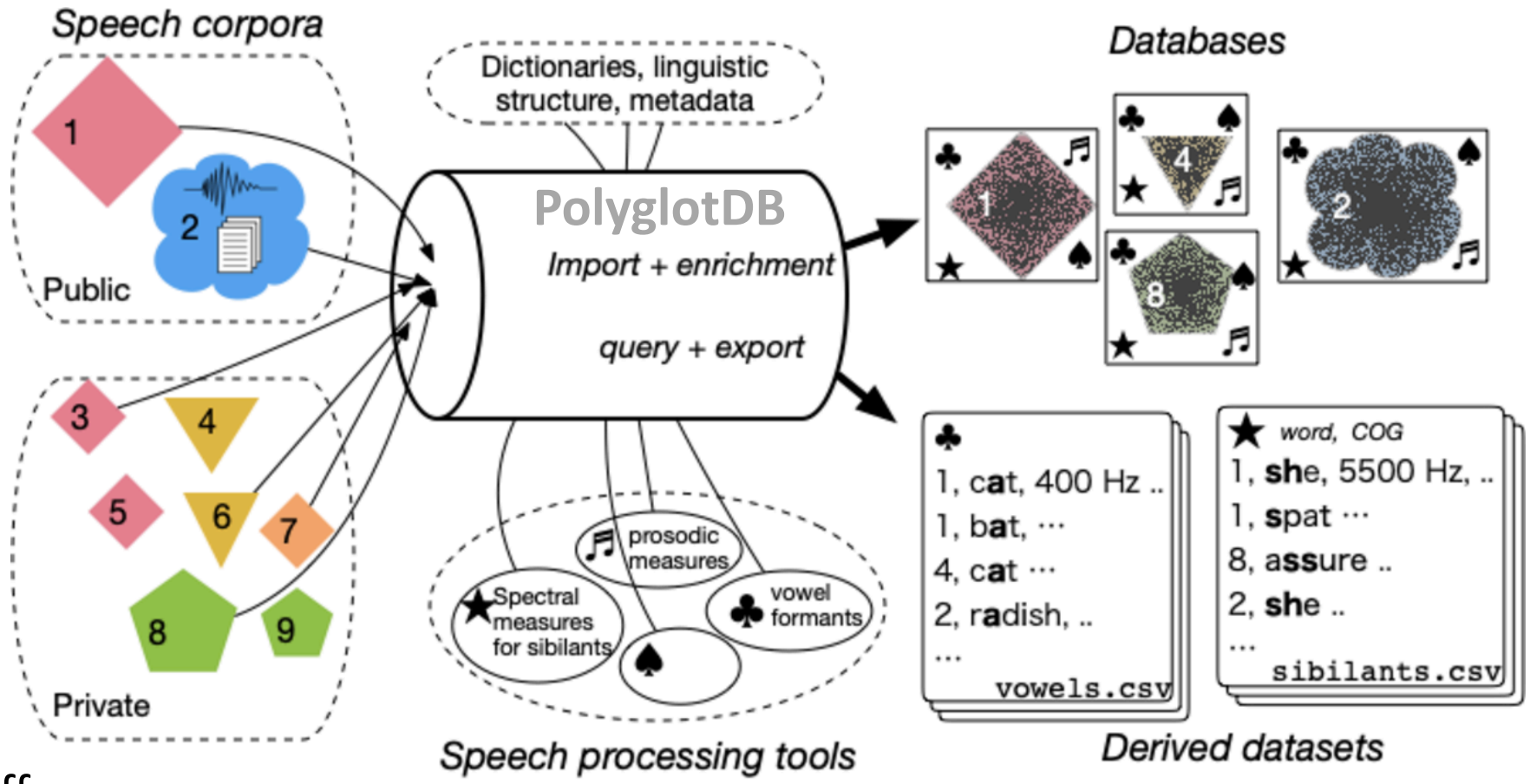
bottleneck

Data file (CSV)

- Gender
- formants
- Vowel duration
- Age
- adjacent segments

- To scale up: need software for integrated speech corpus analysis
- Integrating speech datasets
- Analyzing and querying across them

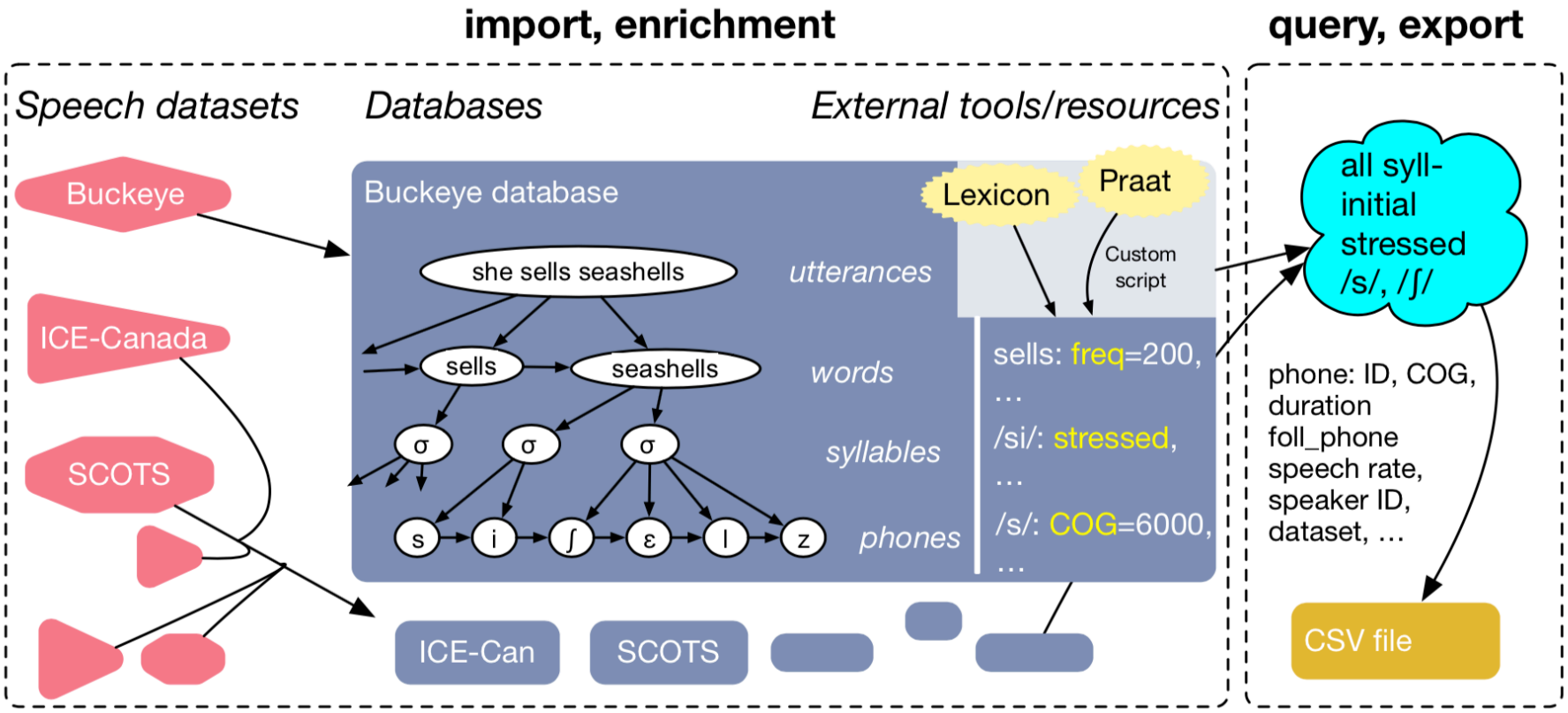
Integrated Speech Corpus ANalysis



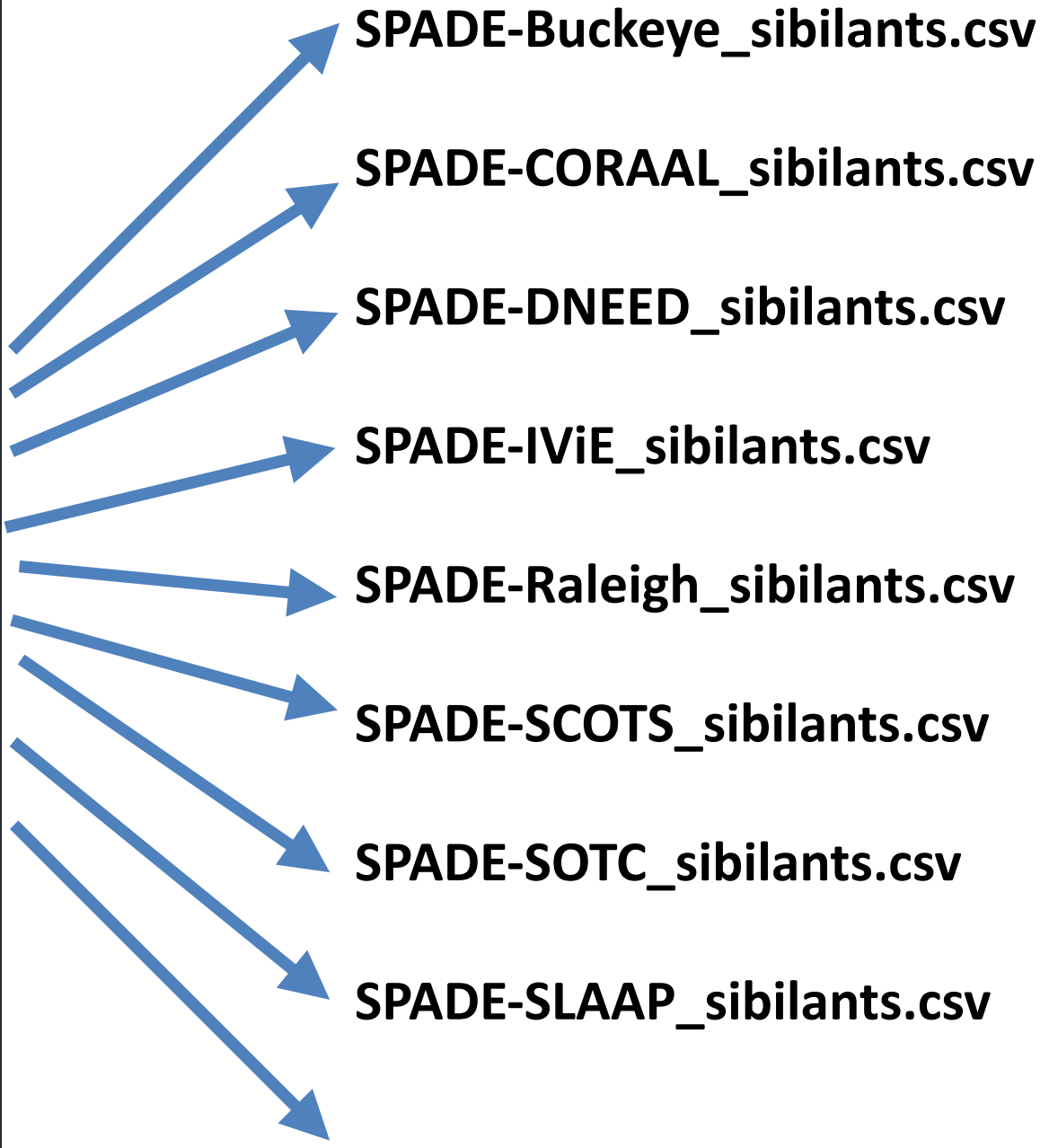
Michael McAuliffe
Software development

McAuliffe et al. (2019) *Proc. ICPhS*

Ex: Finding and measuring sibilants



```python
##################################
## SPADE sibilant analysis script ##
##################################

## Processes and extracts linguistic and acoustic properties of sibilants from
## corpora collected as part of the SPeech Across Dialects of English (SPADE) project.

## Input:
## - corpus name (e.g., Buckeye SOTC)
## - corpus metadata (placed in an associated YAML file)
##   specifies paths to corpus audio, transcripts, and metadata
## Output:
## - CSV of sibilant measurements

import sys
import os
import argparse

base_dir = os.path.dirname(os.path.abspath(__file__))
script_dir = os.path.join(base_dir, 'Common')

sys.path.insert(0, script_dir)

import common

from polyglotdb.utils import ensure_local_database_running
from polyglotdb.config import CorpusConfig

if __name__ == '__main__':
    # process command-line arguments
    parser = argparse.ArgumentParser()
    parser.add_argument('corpus_name', help='Name of the corpus')
    parser.add_argument('-r', '--reset', help="Reset the corpus", action='store_true')
    parser.add_argument('-d', '--docker', help="This script is being called from Docker", action='store_true')

    args = parser.parse_args()
    corpus_name = args.corpus_name
    reset = args.reset
    docker = args.docker
    directories = [x for x in os.listdir(base_dir) if os.path.isdir(x) and x != 'Common']

    # check that the corpus has an associated YAML configuration file
    if args.corpus_name not in directories:
        print(
            'The corpus {0} does not have a directory (available: {1}).  Please make it with a {0}.yaml file inside.'.format(
                args.corpus_name, ', '.join(directories)))
        sys.exit(1)
    corpus_conf = common.load_config(corpus_name)

    ## Process configuration file
    included_speakers = corpus_conf.get('speakers', [])
    ignored_speakers = corpus_conf.get('ignore_speakers', [])
    print('Processing...')
    if reset:
        common.reset(corpus_name)
    ip = common.server_ip
    if docker:
        ip = common.docker_ip
    with ensure_local_database_running(corpus_name, port=common.server_port, ip=ip, token=common.load_token()) as params:
        print(params)
        config = CorpusConfig(corpus_name, **params)
        config.formant_source = 'praat'

        # Process corpus and enrich with information
        # about the lexical properties of words in the
```



Current ISCAN/polyglot functionality

- Aligner support
  - MFA, LaBB-CAT, FAVE, MAUS
- Acoustic analysis
  - Vowel formants (static + trajectories), duration, F0, etc
  - Custom Praat scripts (e.g. CoG, spectral peak, etc)
- Contextual information
  - Surrounding segments, stress, speech rate, etc
- Metadata
  - Speaker & demographic information
  - Word information (lexical set, frequency, etc)

SPADE
SPeech Across Dialects of English

Case Studies (so far)

- Sibilants Stuart-Smith et al. *ICPhS 2019, LabPhon 2020*; Sonderegger et al. *ICPhS 2023*
- Vowels
  - static Mielke et al. *ICPhS 2019*; Smith et al. *LabPhon 2024, Interspeech 2024*
  - dynamic Tanner et al. *SIGMORPHON 2022*
  - duration Tanner et al. *Frontiers in AI 2020*
  - duration (SVLR) Stuart-Smith & Macdonald *New Cambridge History of Eng Lang, in press*
- Speech rate Tanner et al. *LabPhon 2024, Interspeech 2024*
- /r/ & /l/
- Stops

SPADE
SPeech Across Dialects of English

Case Studies (so far)

- Sibilants Stuart-Smith et al. *ICPhS 2019, LabPhon 2020*; Sonderegger et al. *ICPhS 2023*
- Vowels
  - static Mielke et al. *ICPhS 2019*; Smith et al. *LabPhon 2024, Interspeech 2024*
  - dynamic Tanner et al. *SIGMORPHON 2022*
  - duration Tanner et al. *Frontiers in AI 2020*
  - duration (SVLR) Stuart-Smith & Macdonald *New Cambridge History of Eng Lang, in press*
- Speech rate Tanner et al. *LabPhon 2024, Interspeech 2024*
- /r/ & /l/
- Stops

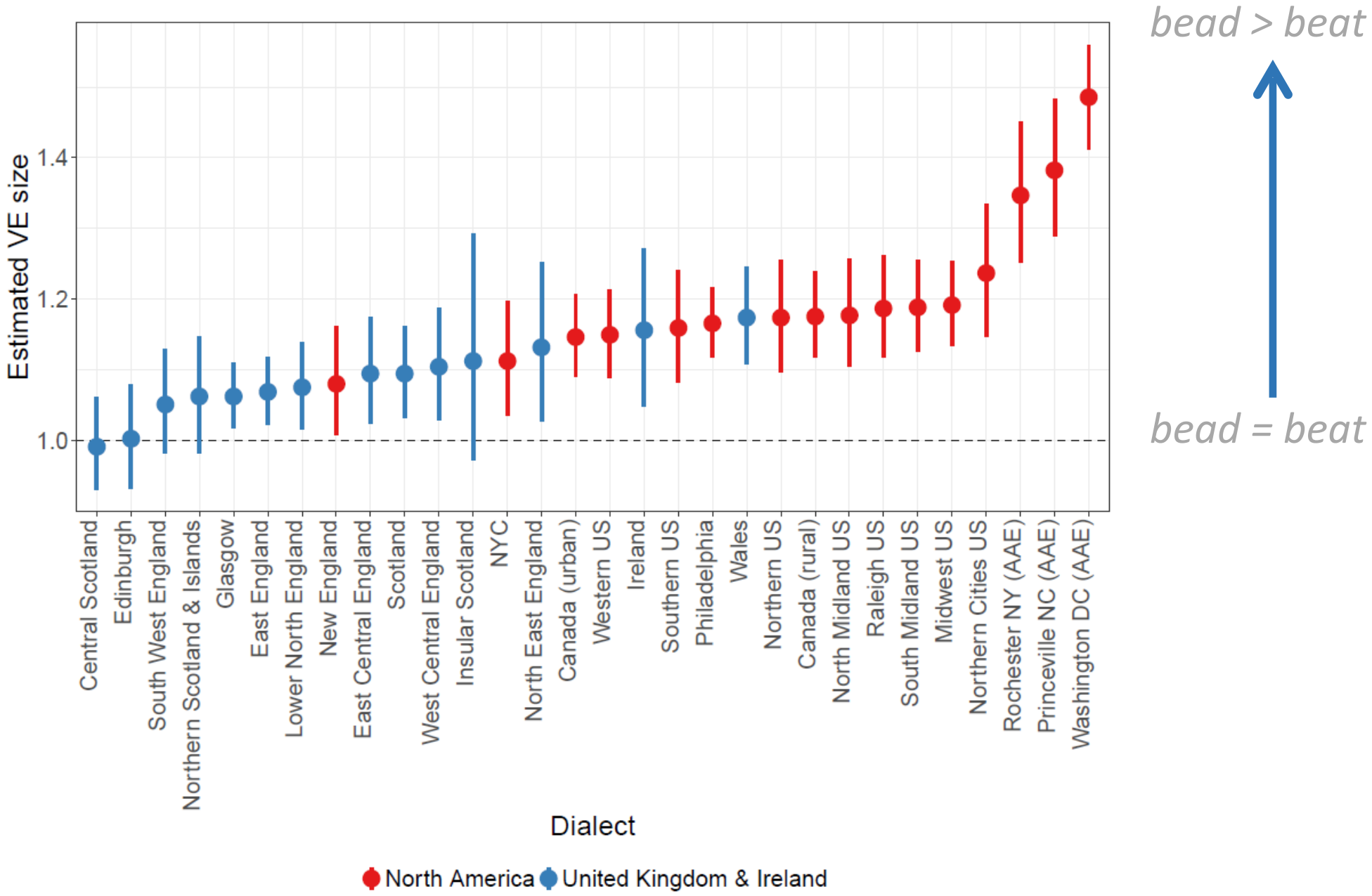
English Voicing Effect (VE)

- Vowels longer before **voiced** consonants
  - bead > beat
- Varies in *size* across languages
  - English = ~1.6x; French = ~1.1x; Russian = 1.2x
- Studied only in **lab speech** (often single words!)
- Dialect-specific vowel patterning
  - SVLR
  - AAE

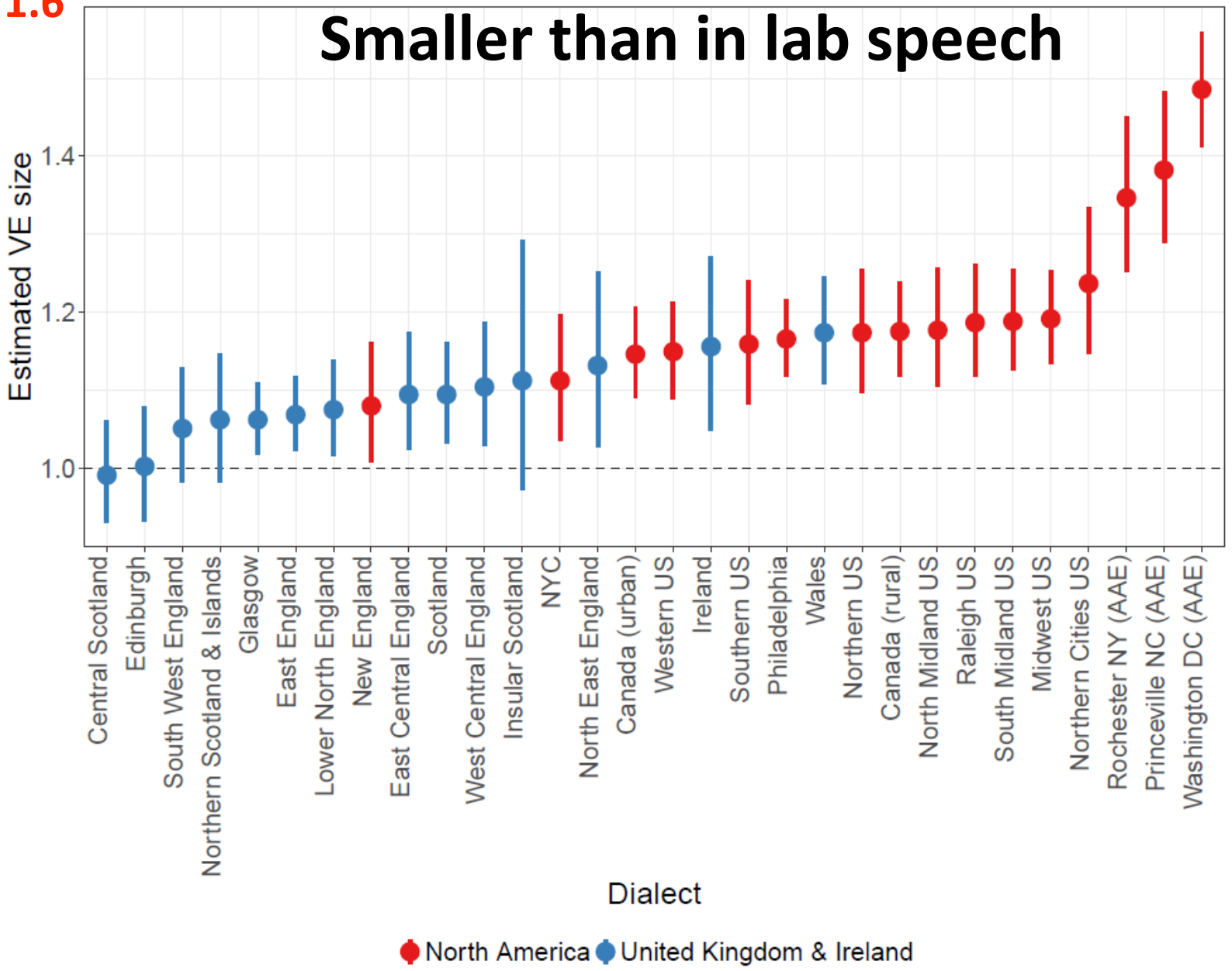
Aitken (1981), Holt (2016)

English Voicing Effect (VE)

- *How robust is the Voicing Effect*
  - *in spontaneous speech?*
  - *across English dialects?*
- Examine VE *across* SPADE corpora
- Utterance-final monosyllabic CVC (e.g. *beat*, *beat)*
  - ~230k tokens
  - ~2k speakers
  - 15 SPADE corpora -> 30 US & UK English dialects


bead > beat
bead = beat
Estimated VE size
1.4
1.2
1.0
Central Scotland
Edinburgh
South West England
Northern Scotland & Islands
Glasgow
East England
Lower North England
New England
East Central England
Scotland
West Central England
Insular Scotland
NYC
North East England
Canada (urban)
Western US
Ireland
Southern US
Philadelphia
Wales
Northern US
Canada (rural)
North Midland US
Raleigh US
South Midland US
Midwest US
Northern Cities US
Rochester NY (AAE)
Princeville NC (AAE)
Washington DC (AAE)
Dialect
North America
United Kingdom & Ireland

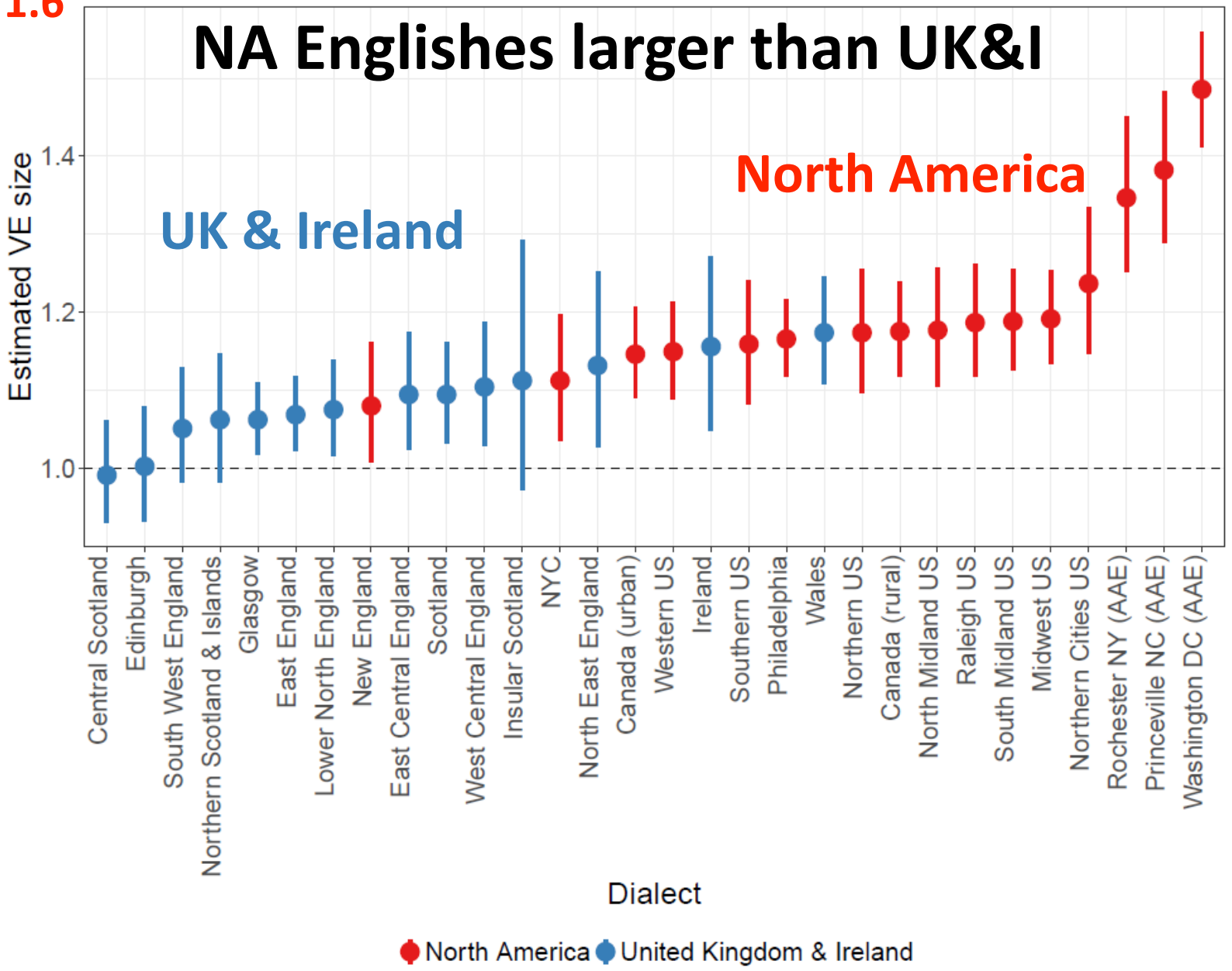
Lab speech ~1.6



bead > beat

bead = beat

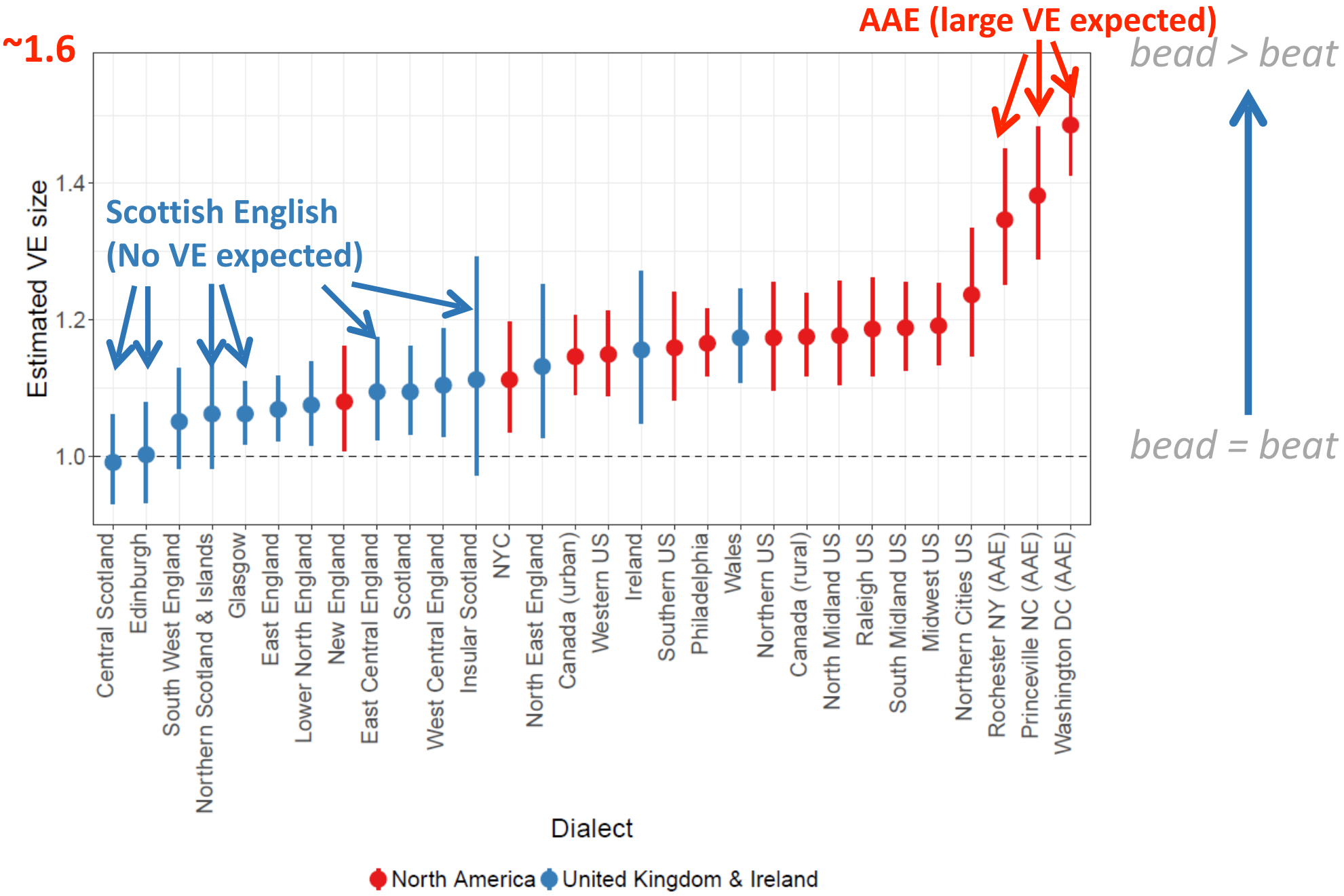
Lab speech ~1.6



bead > beat

bead = beat

Lab speech ~1.6



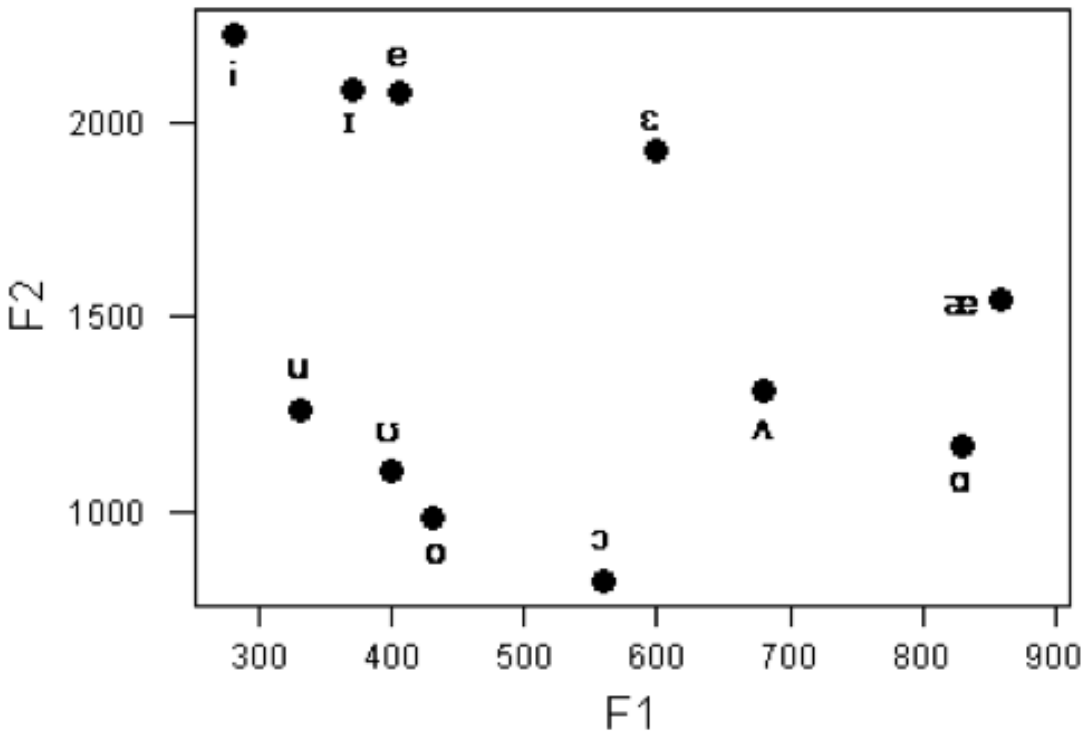
SPADE
SPeech Across Dialects of English

Case Studies (so far)

- Sibilants Stuart-Smith et al. *ICPhS 2019, LabPhon 2020*; Sonderegger et al. *ICPhS 2023*
- Vowels
  - static Mielke et al. *ICPhS 2019*; Smith et al. *LabPhon 2024, Interspeech 2024*
  - dynamic Tanner et al. *SIGMORPHON 2022*
  - duration Tanner et al. *Frontiers in AI 2020*
  - duration (SVLR) Stuart-Smith & Macdonald *New Cambridge History of Eng Lang, in press*
- Speech rate Tanner et al. *LabPhon 2024, Interspeech 2024*
- /r/ & /l/
- Stops

Key acoustic dimensions of vowels

- Duration (e.g., tense vs lax)
- F1 x F2



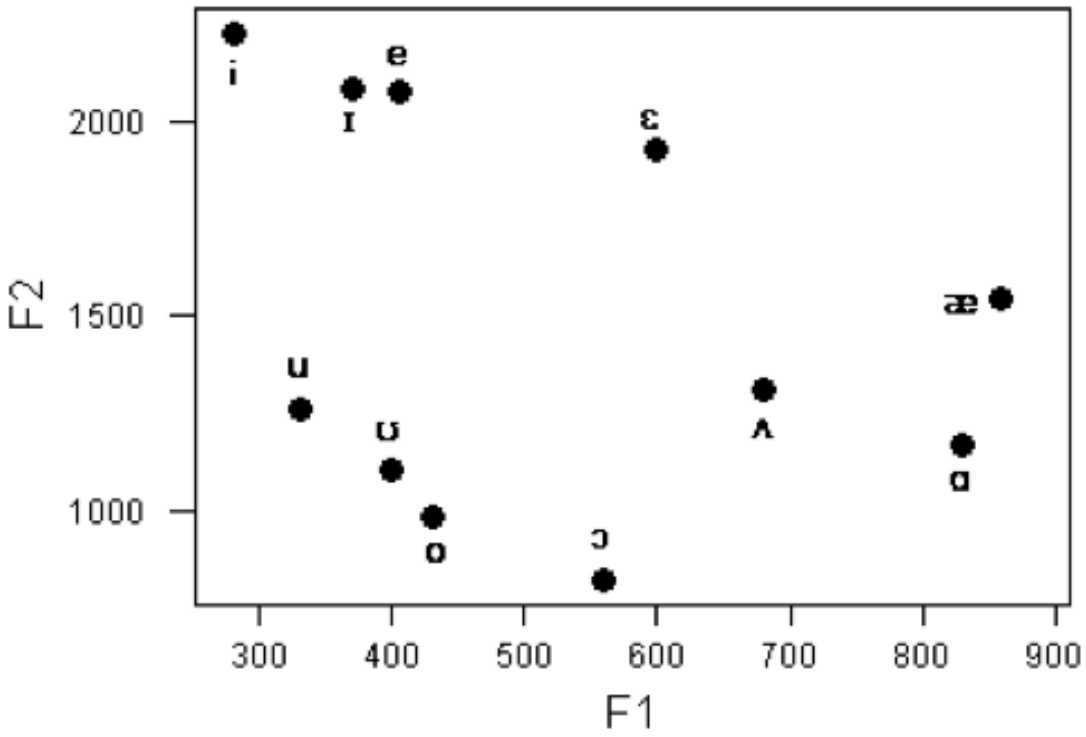
Peterson & Barney (1952), Fant (1960), Ladefoged (1963)

Key acoustic dimensions of vowels

- Duration (e.g., tense vs lax)
- F1 x F2

‘the complex acoustical patterns … are not adequately represented by a single section, but require a more complex portrayal.’

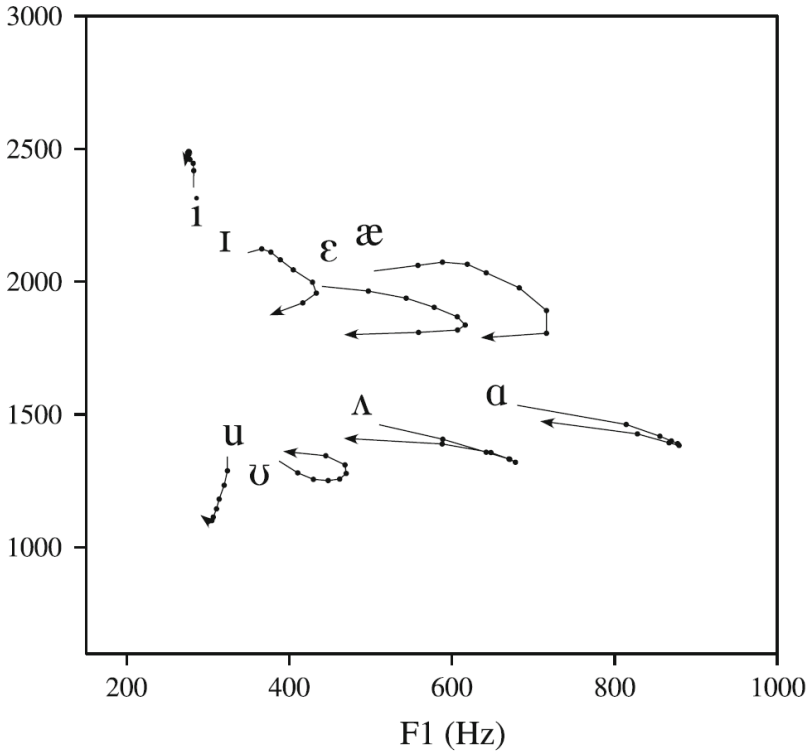
Peterson & Barney (1952: 184)



Peterson & Barney (1952), Fant (1960), Ladefoged (1963)

Key acoustic dimensions of vowels

- Duration (e.g., tense vs lax)
- F1 x F2
- Change in F1,F2 *over vowel timecourse*

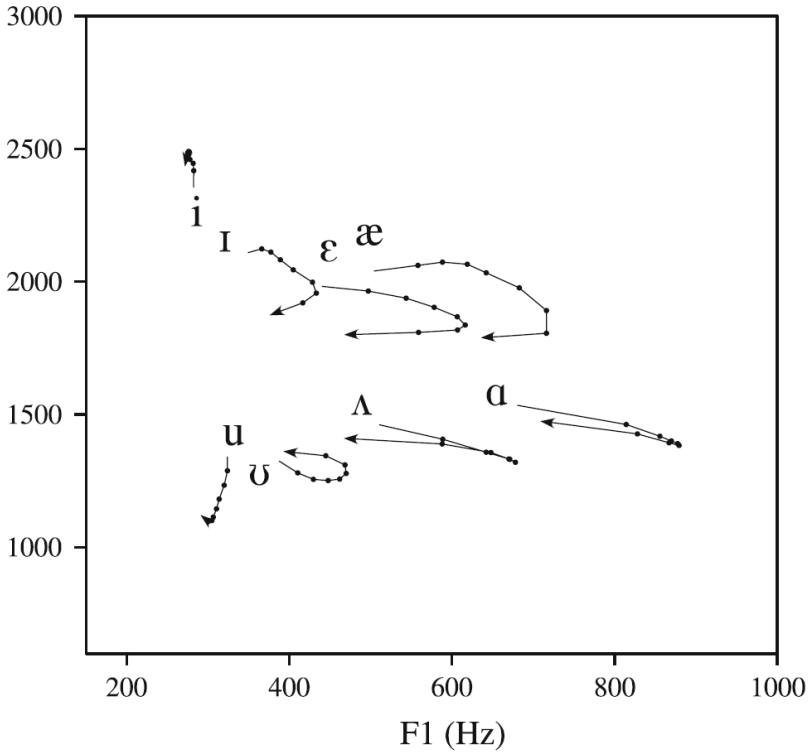


Peterson & Barney (1952), Fant (1960), House (1961). Nearey & Assmann (1986), Watson & Harrington (1999), Morrison & Assmann (2013)

Key acoustic dimensions of vowels

- Duration (e.g., tense vs lax)
- F1 x F2
- Change in F1,F2 *over vowel timecourse*

What about across dialects?

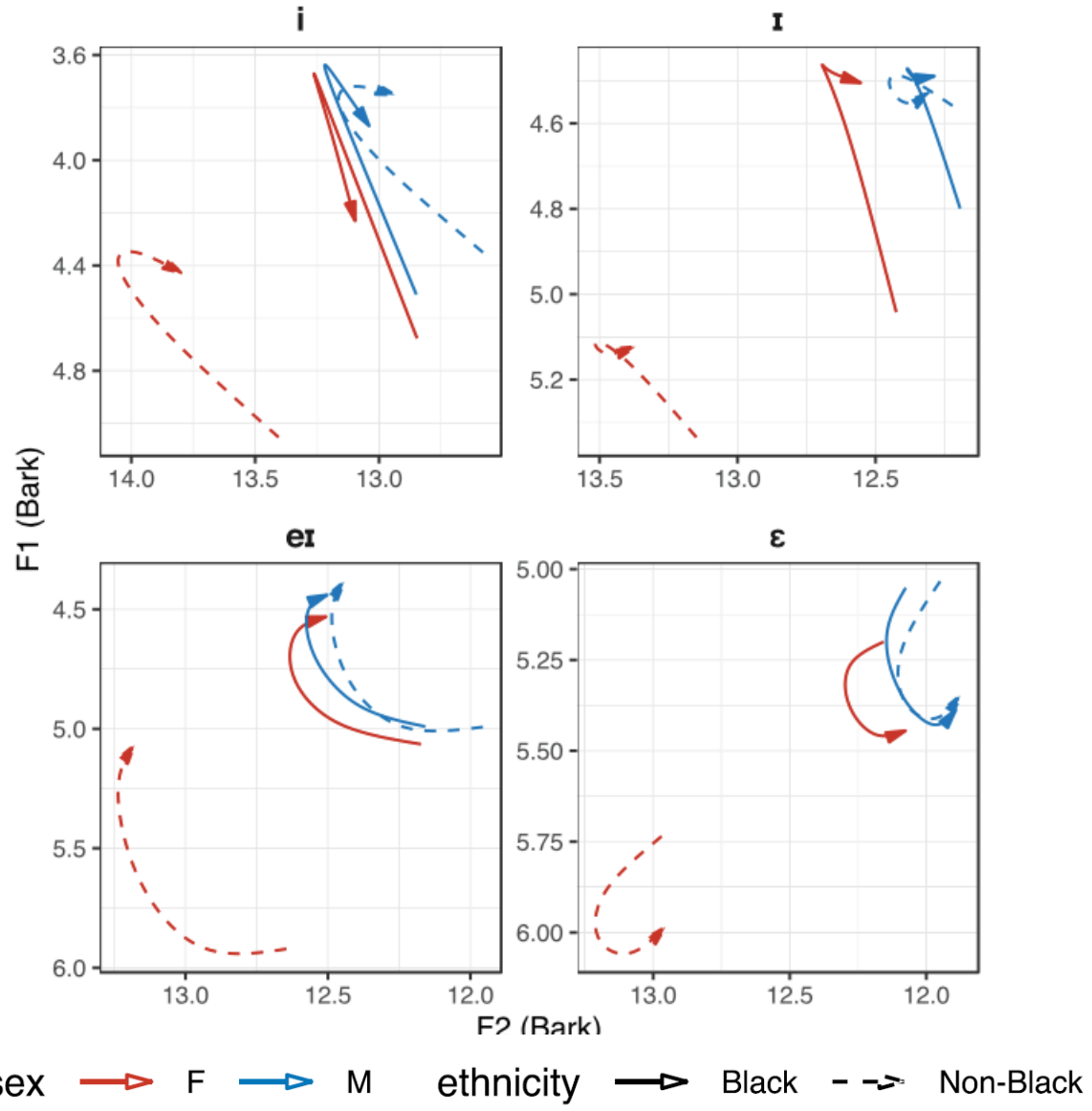


Peterson & Barney (1952), Fant (1960), House (1961). Nearey & Assmann (1986), Watson & Harrington (1999), Morrison & Assmann (2013)

Vowels in English dialects

- differences within + across dialects in *dynamic quality*

- formants
  - up to 4 dialects
  - multipoint/curve parameterization
  - differing sets of vowels selected
- duration
  - up to 15 dialects



Watson & Harrington (1999), Thomas (2001), Jacewicz et al (2007), Tauberer & Evanini (2009), Jacewicz & Fox (2013), Williams & Escudero (2014), Risdal & Kohn (2014), Farrington et al. (2018), Cox & Palethorpe (2019), Williams et al (2019) Renwick & Stanley (2020)

Research Questions

What is the role of *dynamic* information for English vowels?

1. How are dynamic properties of vowels structured across many dialects?
2. How do dialects vary in the dynamic properties of their vowels?

Vowels for this study

FLEECE GOOSE FACE GOAT PRICE MOUTH CHOICE

Monophthongal Diphthongal

- ‘true’/’nominal’ monophthongs … FLEECE GOOSE
- ‘phonetic’ diphthongs FACE GOAT
- ‘true’/’nominal’ diphthongs PRICE MOUTH CHOICE

Labov et al. (2006)

Measurements

- F1 & F2 extracted at 5% increments (21 points)
- Central 60% of vowel
- Z-normalized

Mielke et al. (2019)

English dialects

England East
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Standard Southern British English

Wales South
Ireland North
Ireland South
Scottish Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Scottish Standard English

Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South

Hughes et al (2012); Skretkowicz and Rennie (2005); Labov et al (2006)

English dialects - ethnicity

Black English
British Asian

England East
England West Central
England Merseyside
England East Central
England Lower North
England Northeast

Standard Southern British English

Wales South

Ireland North
Ireland South

Scottish Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern

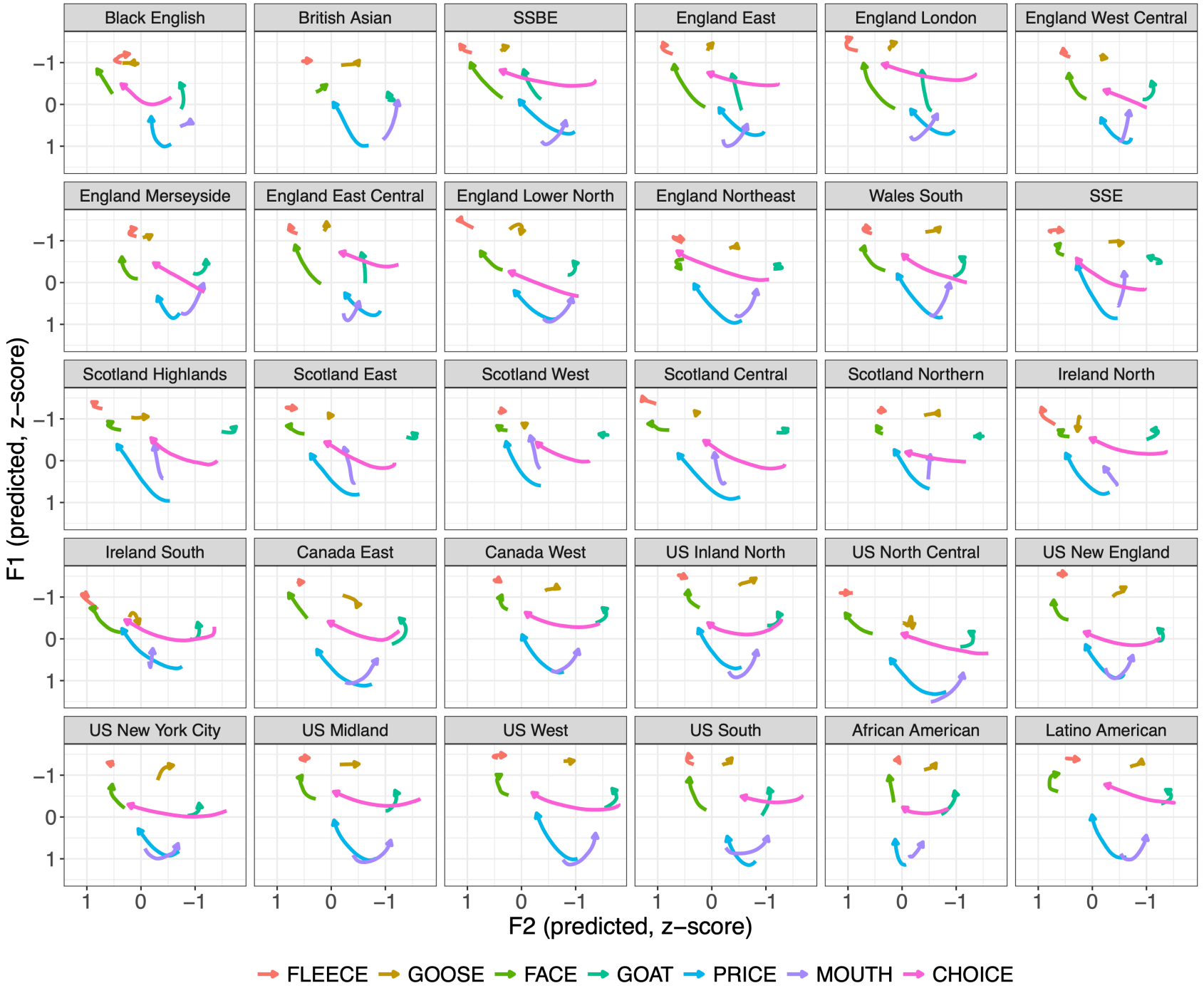
Scottish Standard English

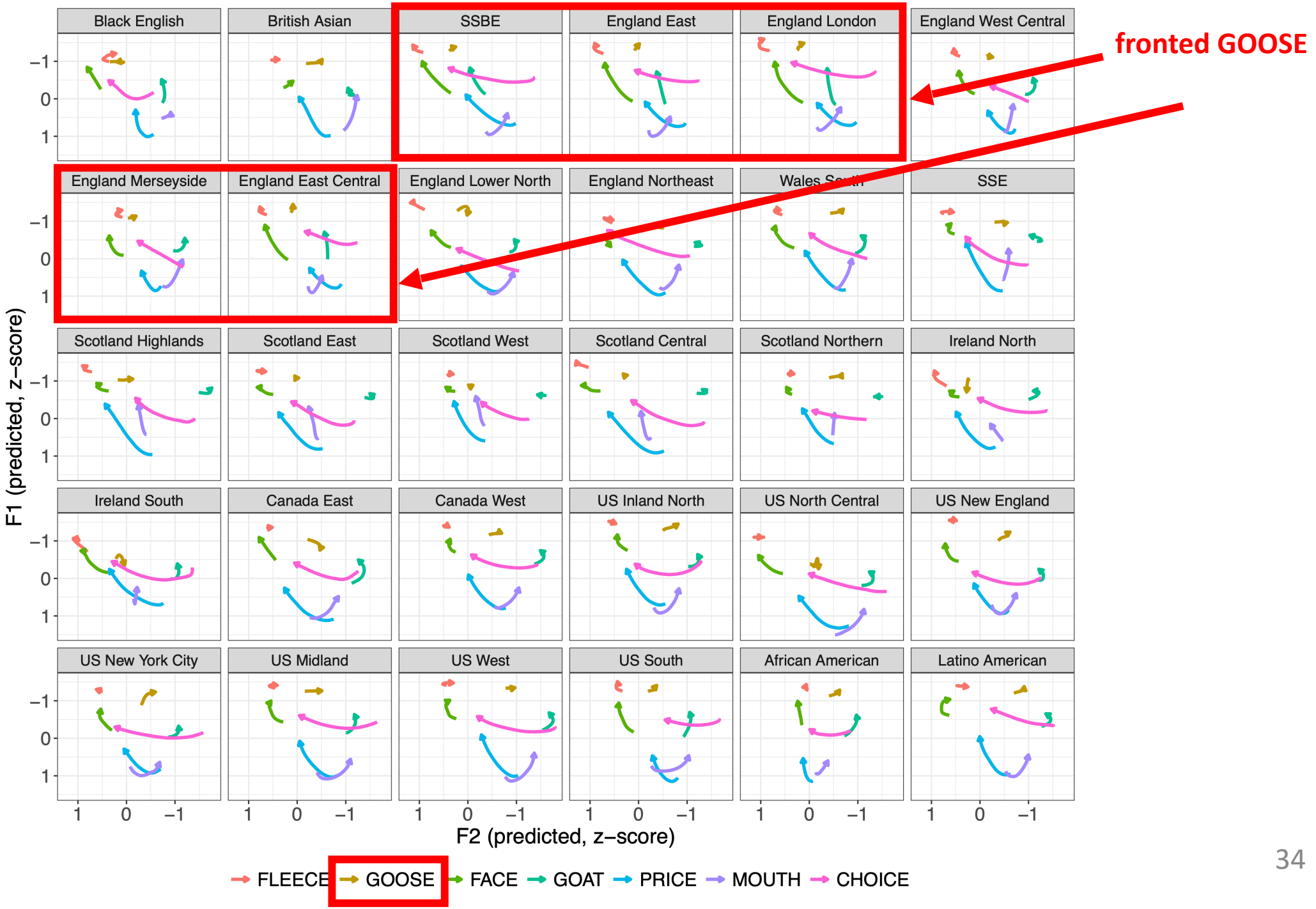
African American
Latino American

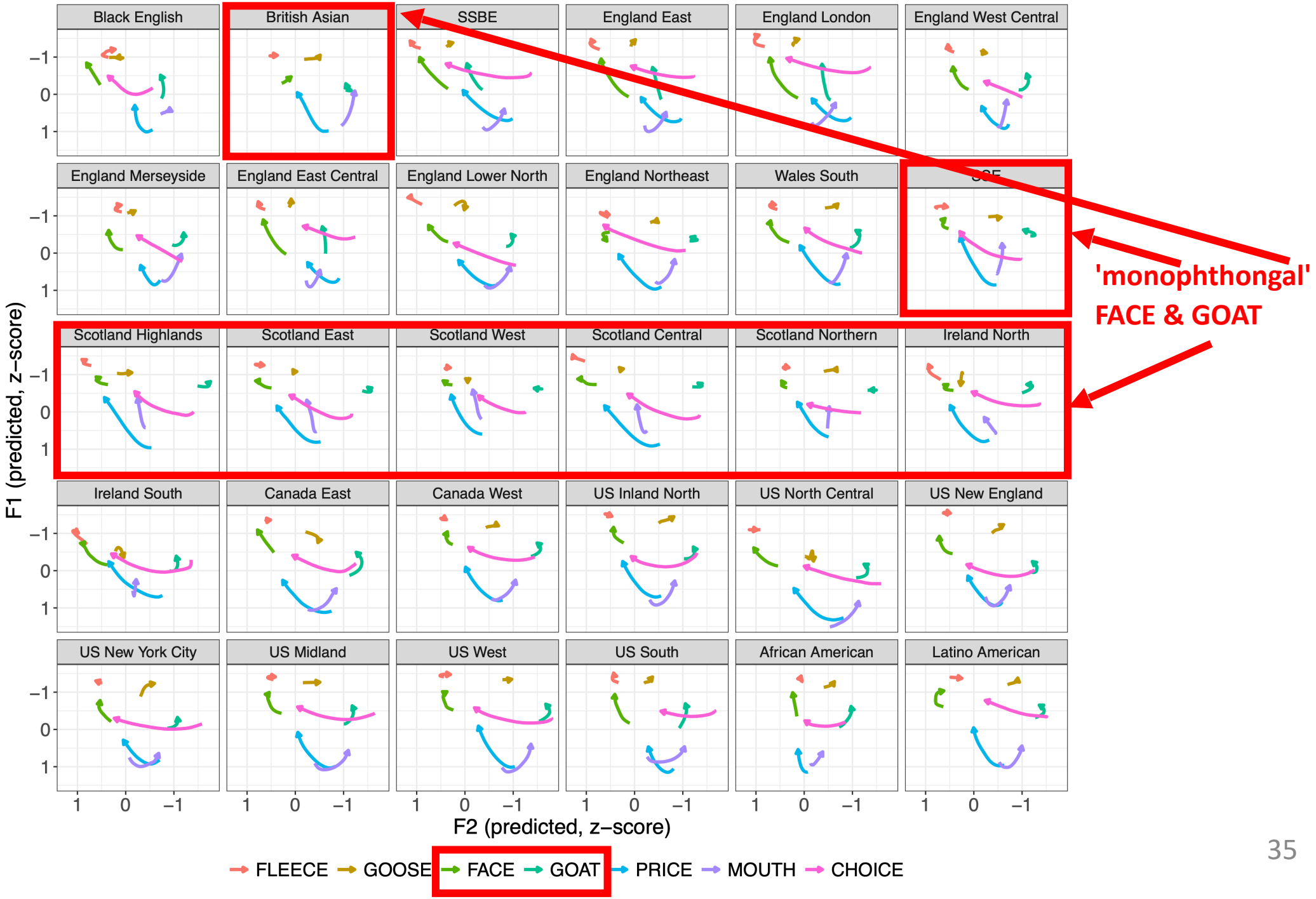
Canada East
Canada West

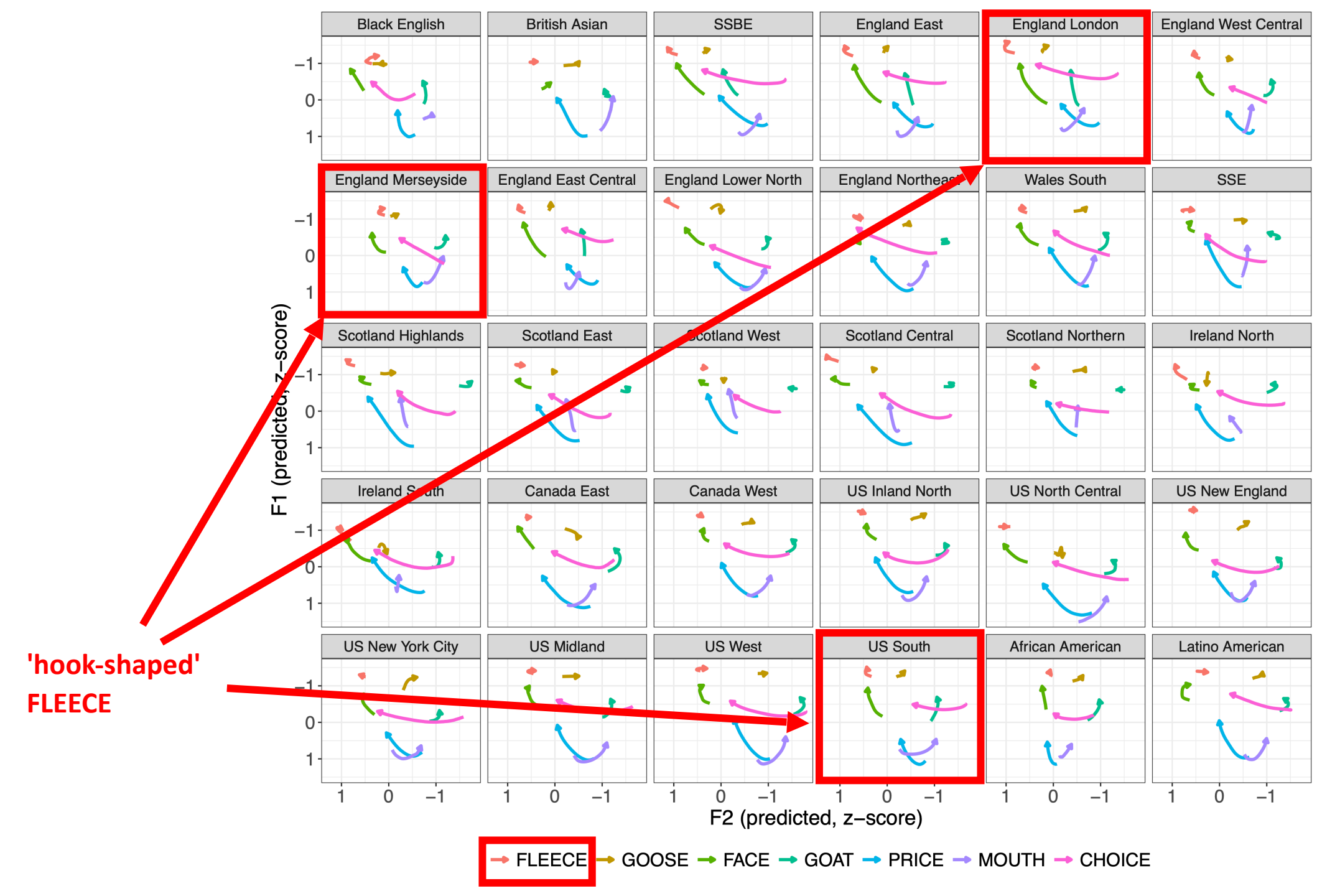
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South

Hughes et al (2012); Skretkowicz and Rennie (2005); Labov et al (2006)

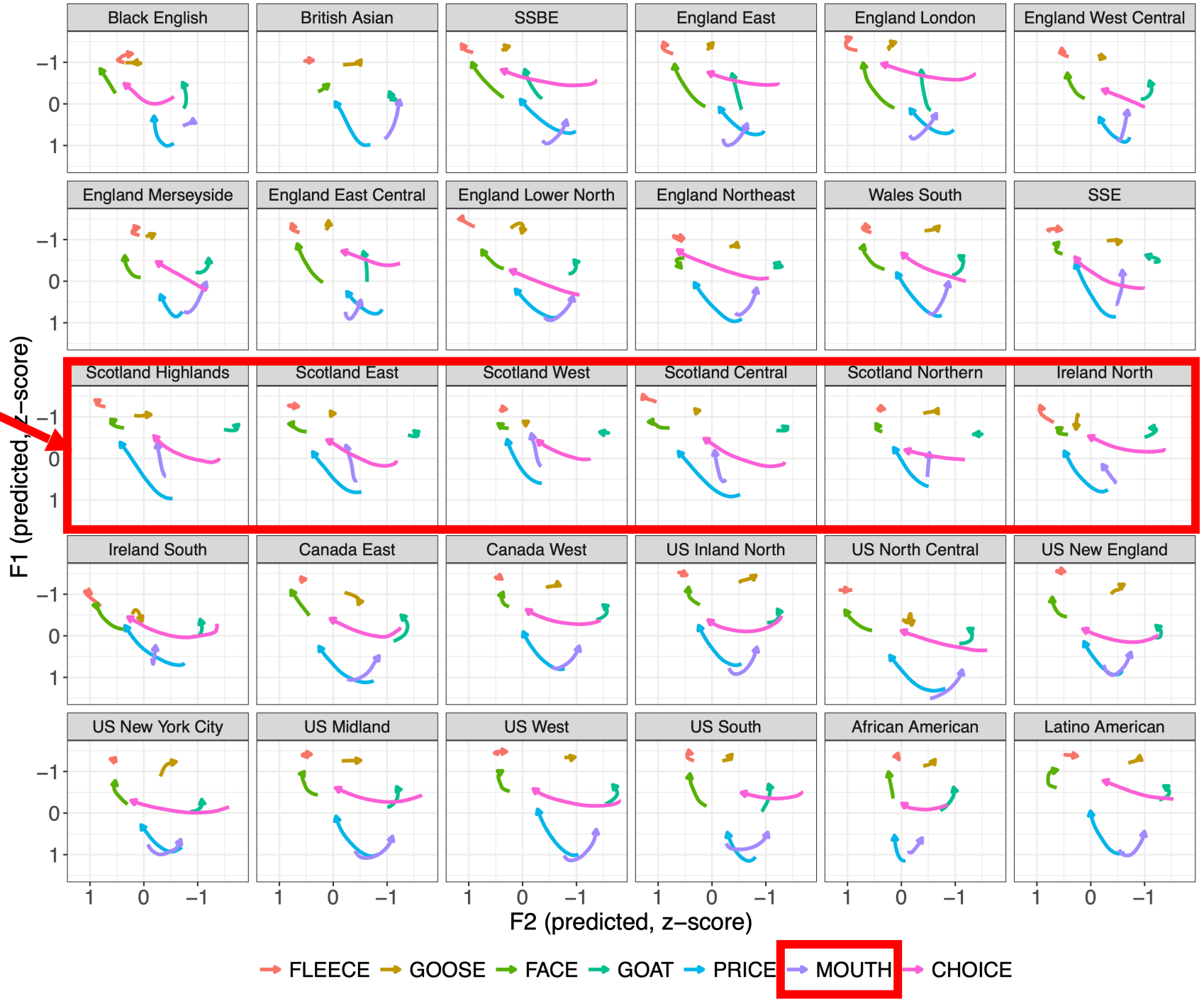

Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South
African American
Latino American
F1 (predicted, z−score)
F2 (predicted, z−score)
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE


fronted GOOSE
Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South
African American
Latino American
F1 (predicted, z−score)
F2 (predicted, z−score)
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE


Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South
African American
Latino American
F1 (predicted, z−score)
F2 (predicted, z−score)
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE
'monophthongal'
FACE & GOAT


Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South
African American
Latino American
F1 (predicted, z−score)
F2 (predicted, z−score)
'hook-shaped'
FLEECE
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE

front-raising MOUTH



SPeech Across Dialects of English

What to do with this app -->

Select distance (in degrees) for clustering in the map

0 | 1 | 100

0 0.01 0.02 0.05 0.25 0.5 1 3 5 10 20 25 50 100

Select birth year range to show in the map

1838 | 2012

1838 1856 1874 1892 1910 1928 1946 1964 1982 2000 2012

Minimum number of tokens to display variable

1 | 100

0 10 20 30 40 50 60 70 80 90 100

Choose the variable

FACE

Choose an acoustic measure

normF1

Choose the scatterplot x-axis

birthyear

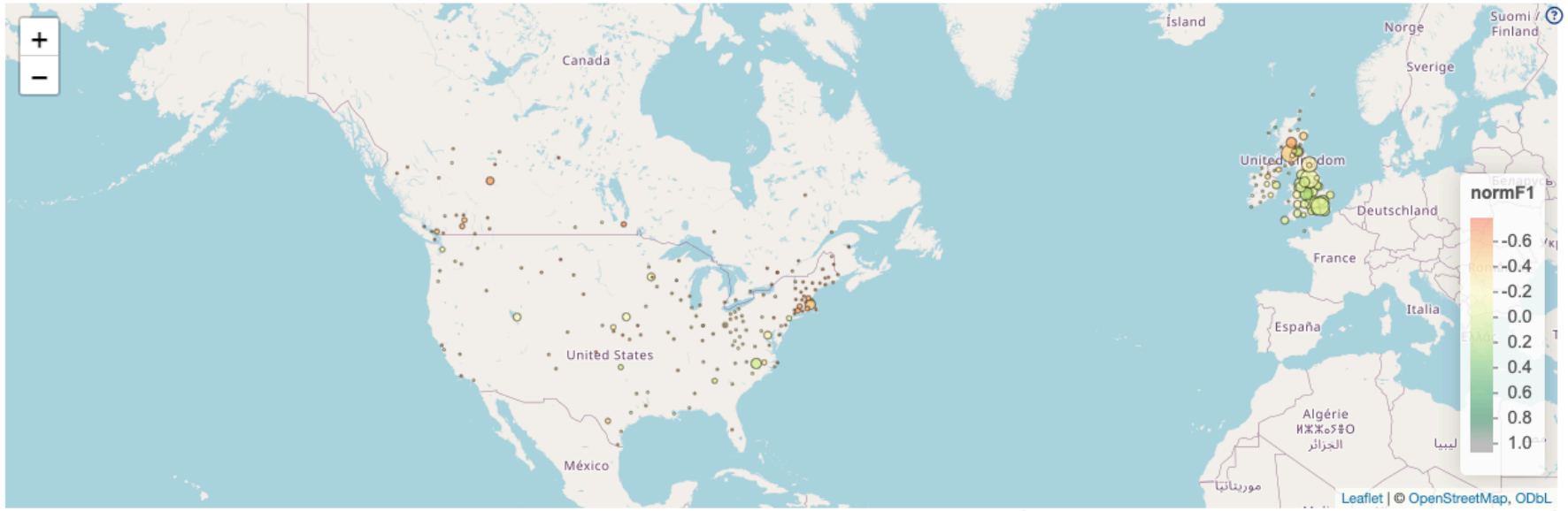
Choose dialects to include in the scatterplot

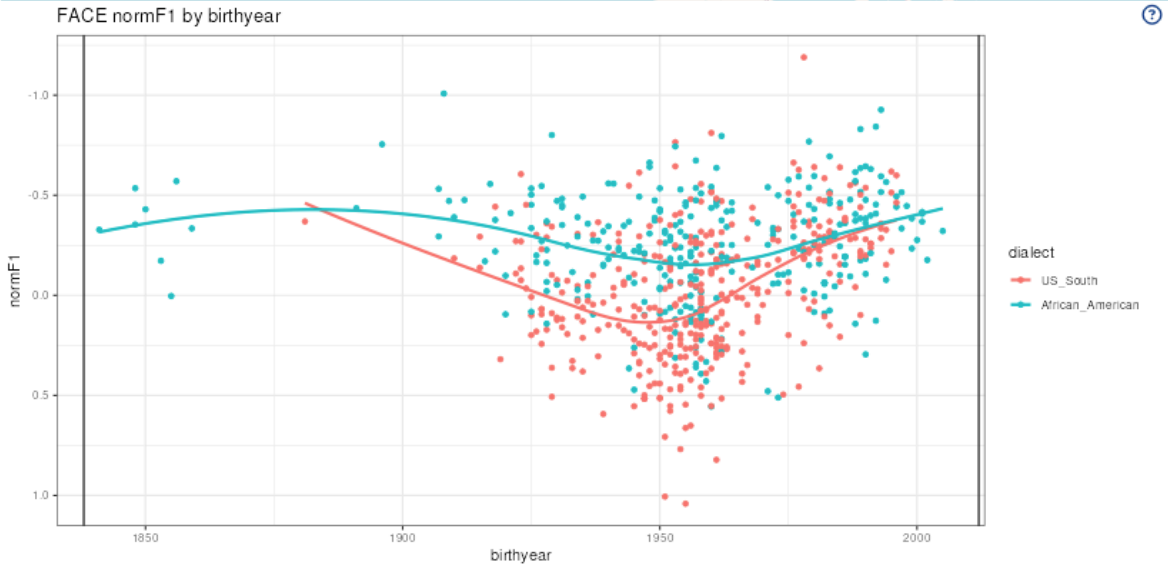
US_South African_American

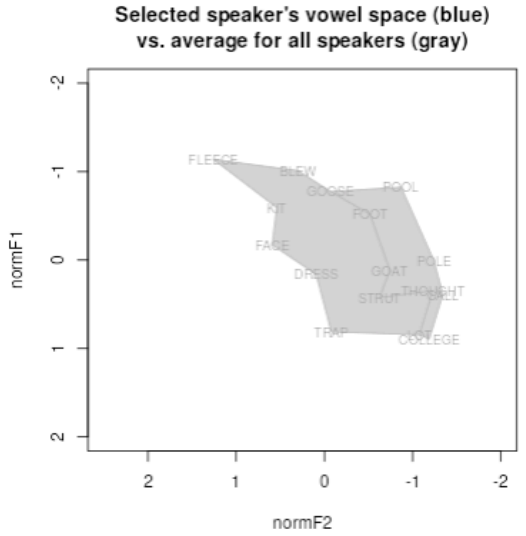
- [ ] Show regions on map
- [ ] Split by gender
- [x] Make scatterplot with ggplot
- [ ] Make vowelplot with ellipses

Exclude Data with Uncertainty

- [ ] birthyear
- [ ] place







Please click a speaker

0:00 / 0:00

https://shiny.chass.ncsu.edu/spade/latest/

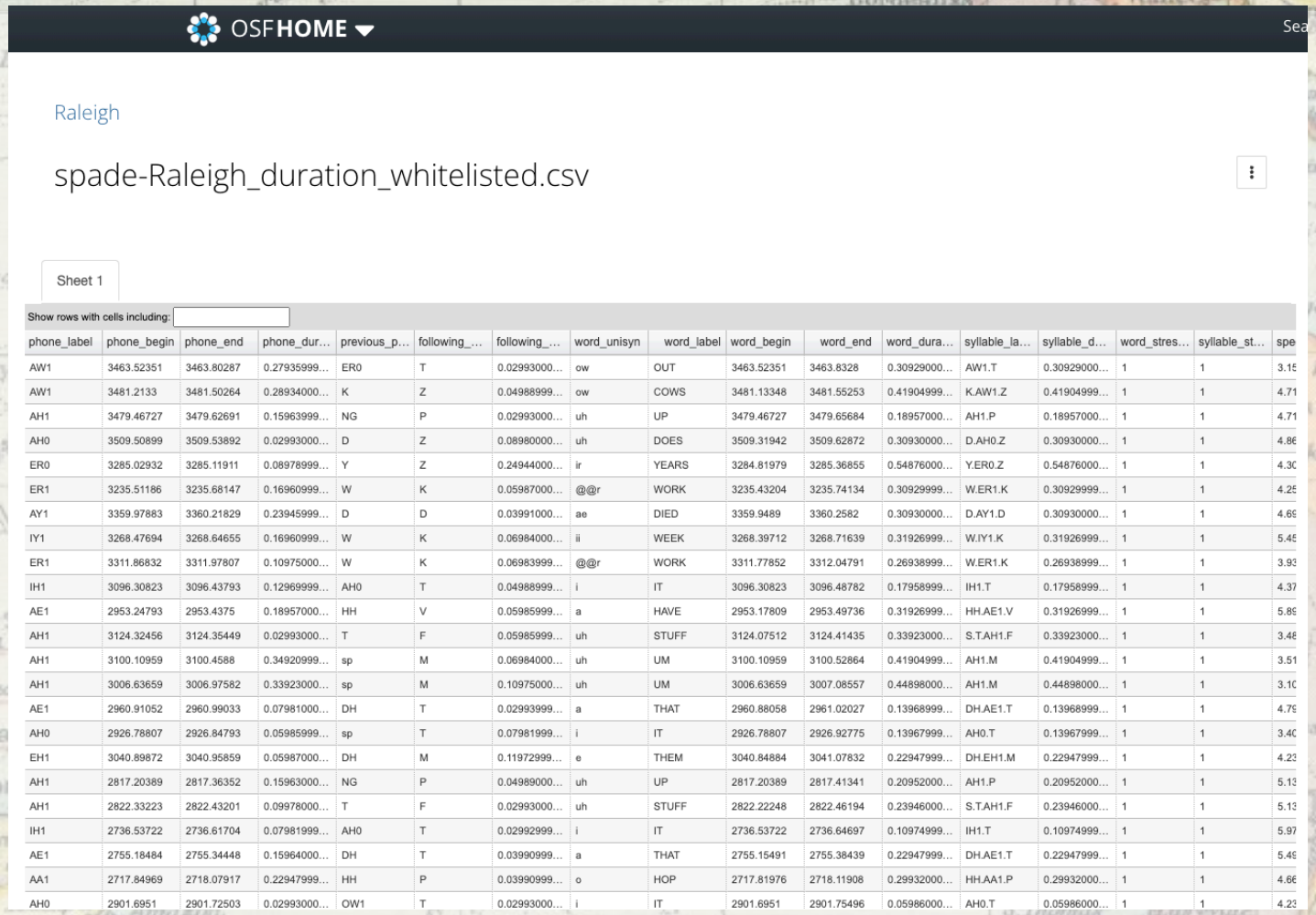
SPADE

SPeech Across Dialects of English

Datasets

https://osf.io/4jfrm/

- Measurements (CSV files):
  - Vowel formants (static)
  - vowel durations
  - sibilants



phone_label	phone_begin	phone_end	phone_dur…	previous_p…	following_…	following_…	word_unisyn	word_label	word_begin	word_end	word_dura…	syllable_la…	syllable_d…	word_stres…	syllable_st…	spe
AW1	3463.52351	3463.80287	0.27935999…	ER0	T	0.02993000…	ow	OUT	3463.52351	3463.8328	0.30929000…	AW1.T	0.30929000…	1	1	3.15
AW1	3481.2133	3481.50264	0.28934000…	K	Z	0.04988999…	ow	COWS	3481.13348	3481.55253	0.41904999…	K.AW1.Z	0.41904999…	1	1	4.71
AH1	3479.46727	3479.62691	0.15963999…	NG	P	0.02993000…	uh	UP	3479.46727	3479.65684	0.18957000…	AH1.P	0.18957000…	1	1	4.71
AH0	3509.50899	3509.53892	0.02993000…	D	Z	0.08980000…	uh	DOES	3509.31942	3509.62872	0.30930000…	D.AH0.Z	0.30930000…	1	1	4.86
ER0	3285.02932	3285.11911	0.08978999…	Y	Z	0.24944000…	ir	YEARS	3284.81979	3285.36855	0.54876000…	Y.ER0.Z	0.54876000…	1	1	4.30
ER1	3235.51186	3235.68147	0.16960999…	W	K	0.05987000…	@@r	WORK	3235.43204	3235.74134	0.30929999…	W.ER1.K	0.30929999…	1	1	4.25
AY1	3359.97883	3360.21829	0.23945999…	D	D	0.03991000…	ae	DIED	3359.9489	3360.2582	0.30930000…	D.AY1.D	0.30930000…	1	1	4.69
IY1	3268.47694	3268.64655	0.16960999…	W	K	0.06984000…	ii	WEEK	3268.39712	3268.71639	0.31926999…	W.IY1.K	0.31926999…	1	1	5.45
ER1	3311.86832	3311.97807	0.10975000…	W	K	0.06983999…	@@r	WORK	3311.77852	3312.04791	0.26938999…	W.ER1.K	0.26938999…	1	1	3.93
IH1	3096.30823	3096.43793	0.12969999…	AH0	T	0.04988999…	i	IT	3096.30823	3096.48782	0.17958999…	IH1.T	0.17958999…	1	1	4.37
AE1	2953.24793	2953.4375	0.18957000…	HH	V	0.05985999…	a	HAVE	2953.17809	2953.49736	0.31926999…	HH.AE1.V	0.31926999…	1	1	5.89
AH1	3124.32456	3124.35449	0.02993000…	T	F	0.05985999…	uh	STUFF	3124.07512	3124.41435	0.33923000…	S.T.AH1.F	0.33923000…	1	1	3.48
AH1	3100.10959	3100.4588	0.34920999…	sp	M	0.06984000…	uh	UM	3100.10959	3100.52864	0.41904999…	AH1.M	0.41904999…	1	1	3.51
AH1	3006.63659	3006.97582	0.33923000…	sp	M	0.10975000…	uh	UM	3006.63659	3007.08557	0.44898000…	AH1.M	0.44898000…	1	1	3.10
AE1	2960.91052	2960.99033	0.07981000…	DH	T	0.02993999…	a	THAT	2960.88058	2961.02027	0.13968999…	DH.AE1.T	0.13968999…	1	1	4.79
AH0	2926.78807	2926.84793	0.05985999…	sp	T	0.07981999…	i	IT	2926.78807	2926.92775	0.13967999…	AH0.T	0.13967999…	1	1	3.40
EH1	3040.89872	3040.95859	0.05987000…	DH	M	0.11972999…	e	THEM	3040.84884	3041.07832	0.22947999…	DH.EH1.M	0.22947999…	1	1	4.23
AH1	2817.20389	2817.36352	0.15963000…	NG	P	0.04989000…	uh	UP	2817.20389	2817.41341	0.20952000…	AH1.P	0.20952000…	1	1	5.13
AH1	2822.33223	2822.43201	0.09978000…	T	F	0.02993000…	uh	STUFF	2822.22248	2822.46194	0.23946000…	S.T.AH1.F	0.23946000…	1	1	5.13
IH1	2736.53722	2736.61704	0.07981999…	AH0	T	0.02992999…	i	IT	2736.53722	2736.64697	0.10974999…	IH1.T	0.10974999…	1	1	5.97
AE1	2755.18484	2755.34448	0.15964000…	DH	T	0.03990999…	a	THAT	2755.15491	2755.38439	0.22947999…	DH.AE1.T	0.22947999…	1	1	5.49
AA1	2717.84969	2718.07917	0.22947999…	HH	P	0.03990999…	o	HOP	2717.81976	2718.11908	0.29932000…	HH.AA1.P	0.29932000…	1	1	4.66
AH0	2901.6951	2901.72503	0.02993000…	OW1	T	0.02993000…	i	IT	2901.6951	2901.75496	0.05986000…	AH0.T	0.05986000…	1	1	4.23

SPADE

SPeech Across Dialects of English

Running PolyglotDB

https://polyglotdb.readthedocs.io/

- Under active development (RA employed at McGill)

Thank you!

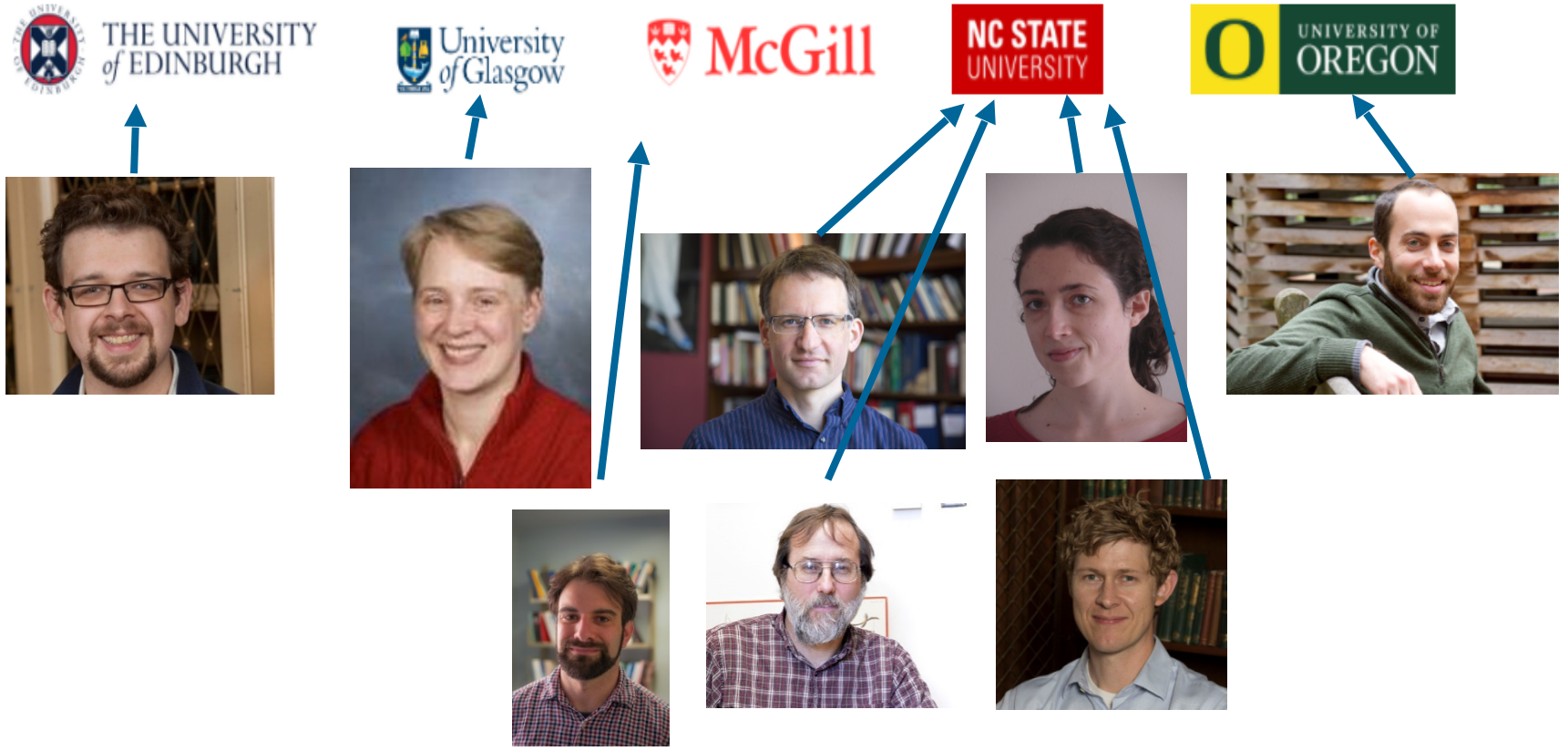
- Data Guardians
- SPADE Team

- SHINY tool: https://shiny.chass.ncsu.edu/spade/latest/
- Measures available: https://osf.io/4jfrm/

SPADE

SPeech Across Dialects of English

Investigators



http://spade.glasgow.ac.uk/

SPADE

SPeech Across Dialects of English

Postdocs

Rachel Macdonald
Project manager

Michael McAuliffe
Software development

James Tanner
PhD (2020), postdoc

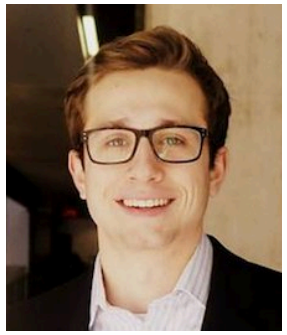
http://spade.glasgow.ac.uk/

Stacey Harkin
Kirsty McCahill
Mitchell McGee
Edward Marshall
Julia Moreno
Jo Pearce
Niamh Walker
Ewa Wanat

Jordan Holley
Peter Andrews
Kaylynn Gunter

and many more!

Arlie Coles (U. de Montréal)



Elias Stengel-Eskin (Johns Hopkins)

Michael Goodale, Sarah Mihuc (McGill)

Vanna Willerton (McGIll)

Ethics and credit

- For private datasets ethics complex: GDPR + US laws
- Data transfer agreement
  - data use in keeping with original permissions, as far as is possible
- Credit
  - When possible: ‘SPADE consortium’ author
  - Always: cite individual datasets used
- Datasets of measures → data guardians at end of project