SPADE

SPeech Across Dialects of English

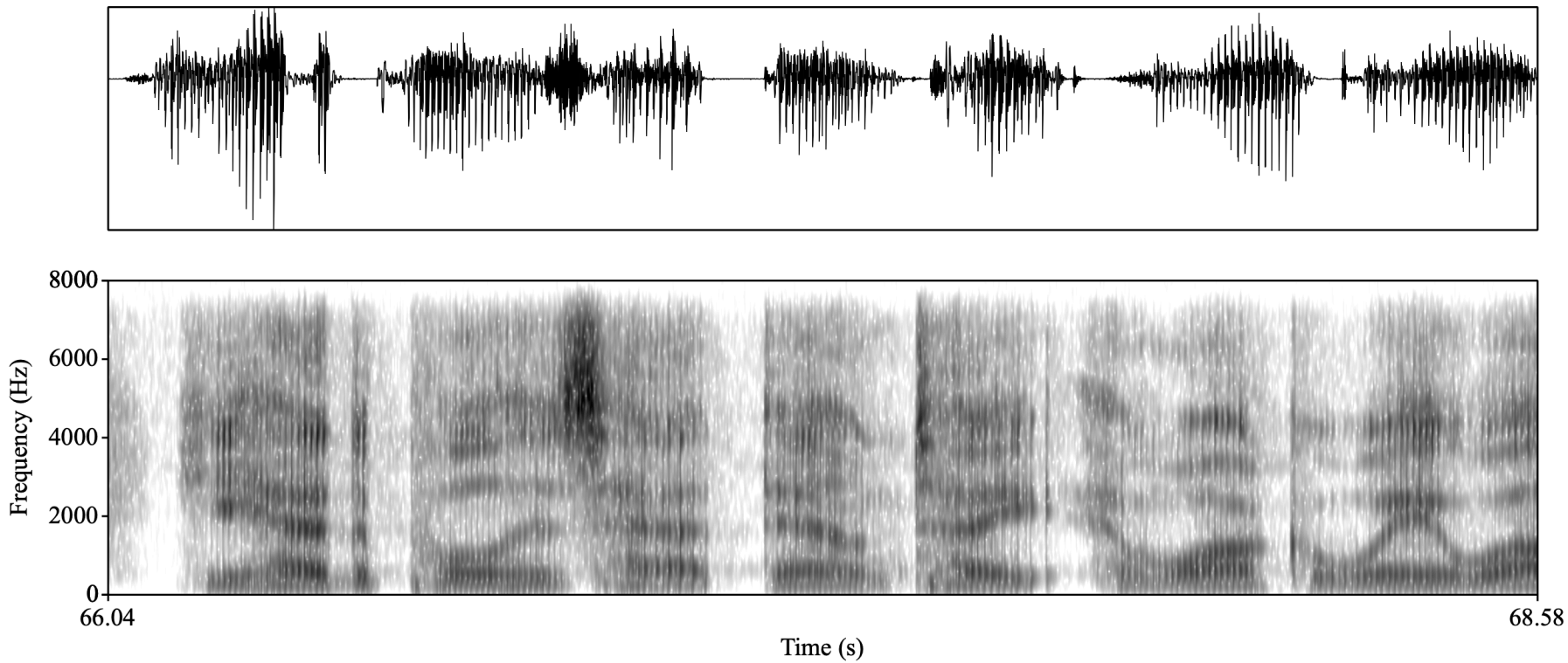
Vowel variation across English dialects: variation in formant dynamics and duration

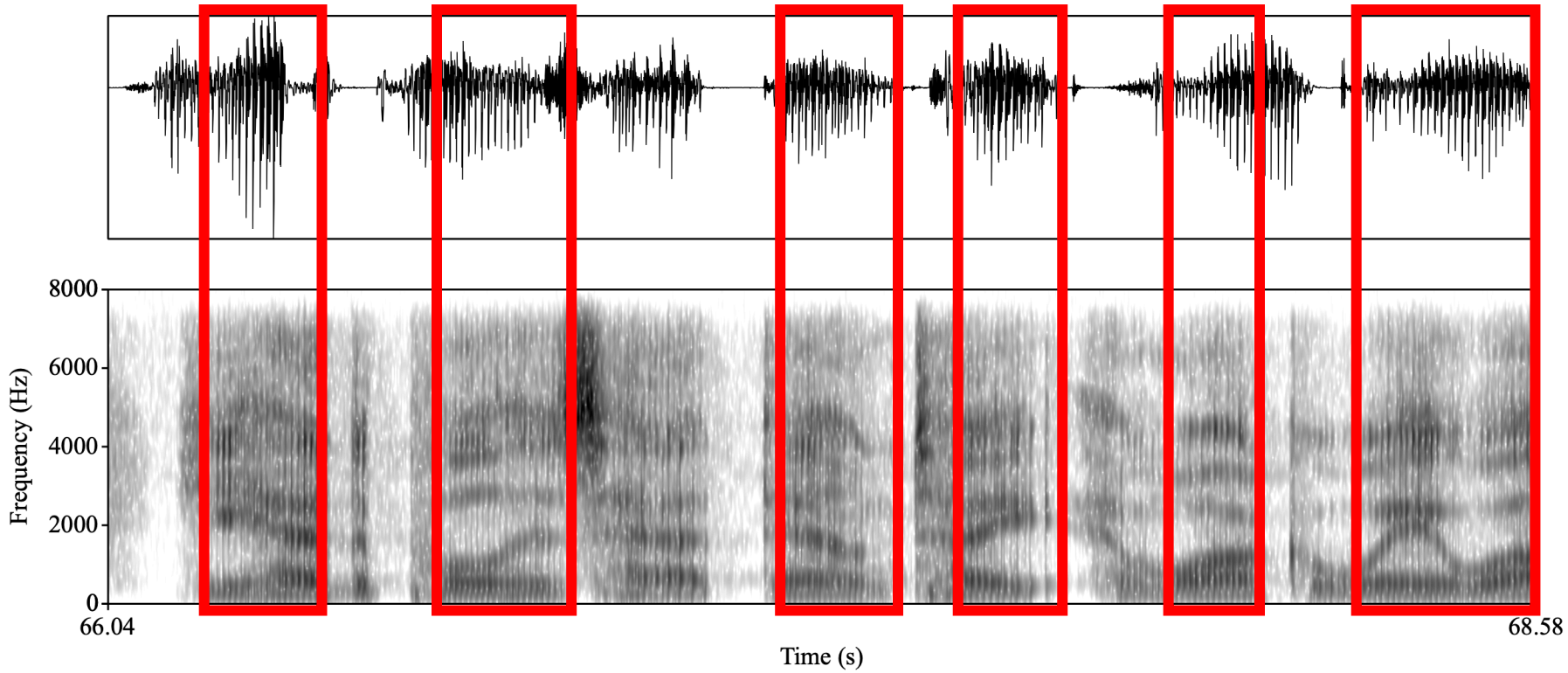
James Tanner, Jane Stuart-Smith, Morgan Sonderegger,
Jeff Mielke, Erik Thomas, Charles Boberg, Robin Dodsworth,
The SPADE Consortium

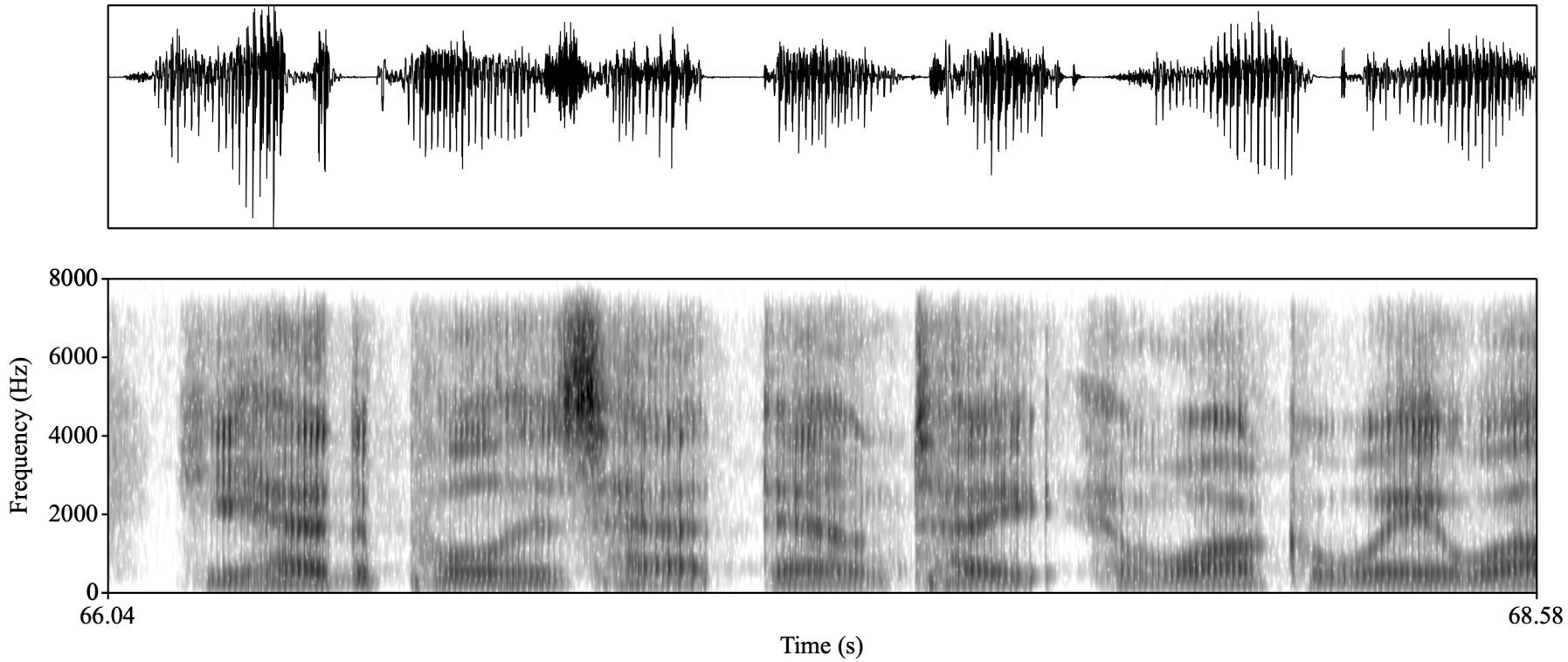
University of Glasgow

McGill University

NC STATE UNIVERSITY

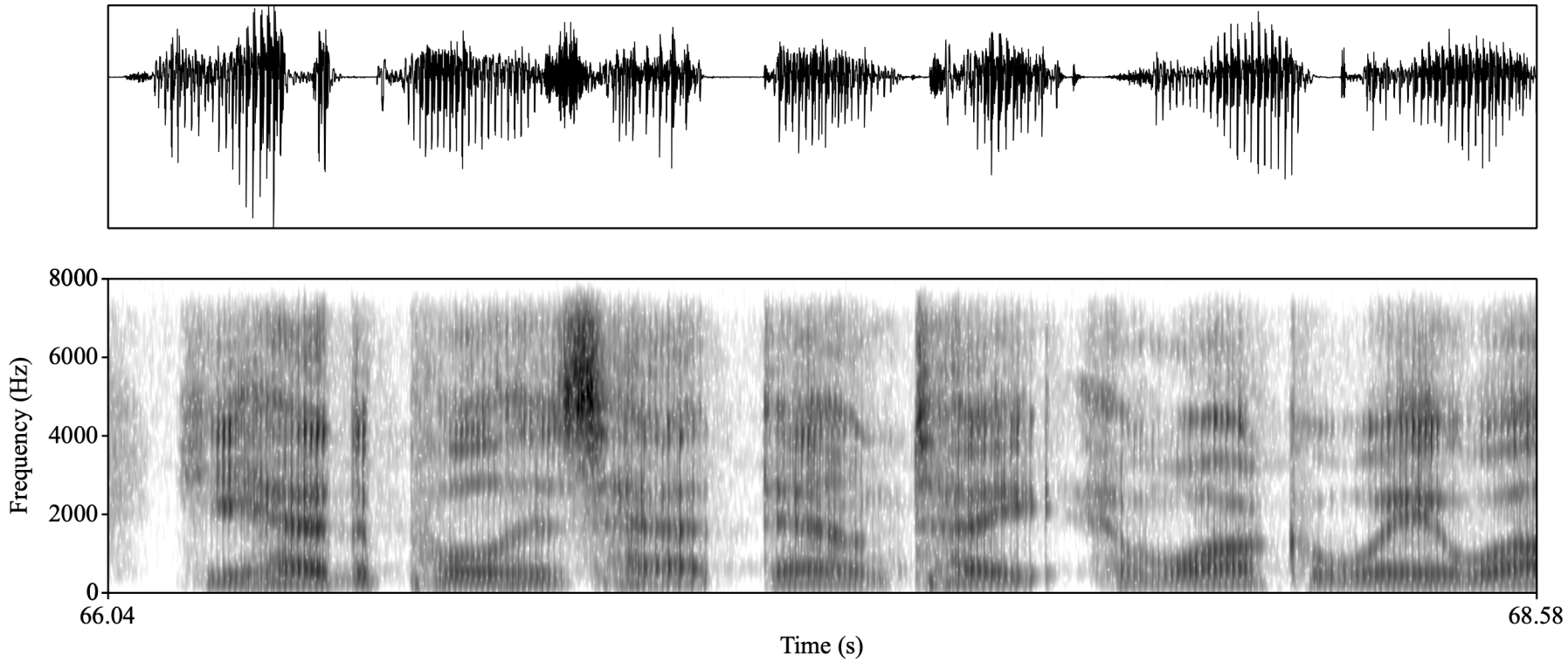

8000
6000
4000
2000
0
Frequency (Hz)
66.04
68.58
Time (s)


8000
6000
4000
2000
0
Frequency (Hz)
66.04
68.58
Time (s)



‘Vowel segments … are acoustically *multidimensional*’

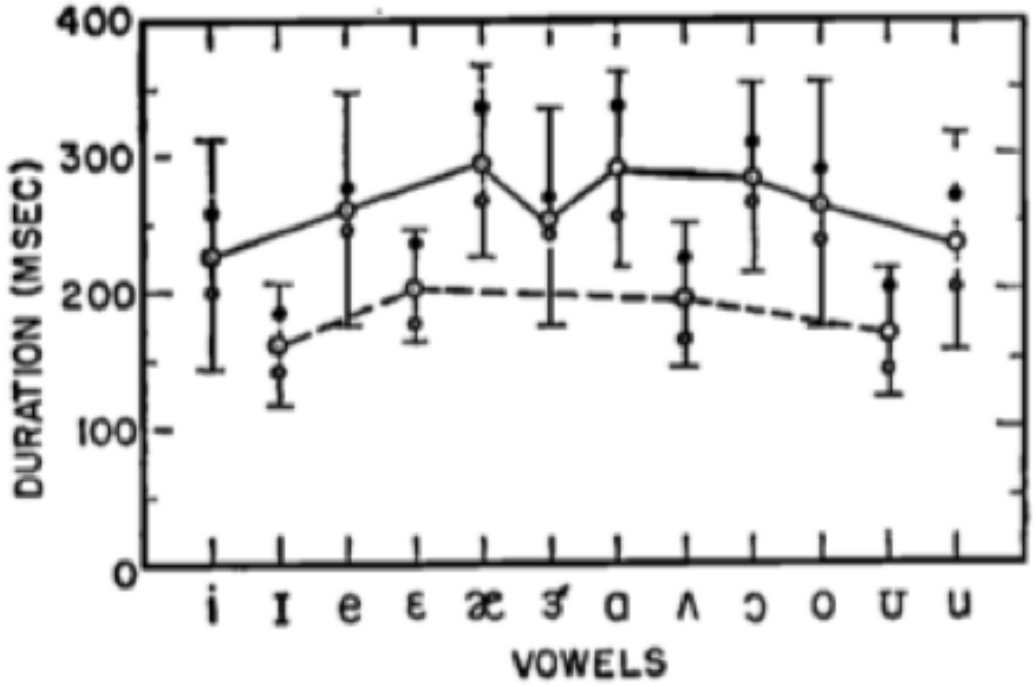
Williams et al (2019: 587)



What are the key acoustic dimensions that characterise the similarities and differences between vowels?

Key acoustic dimensions of vowels

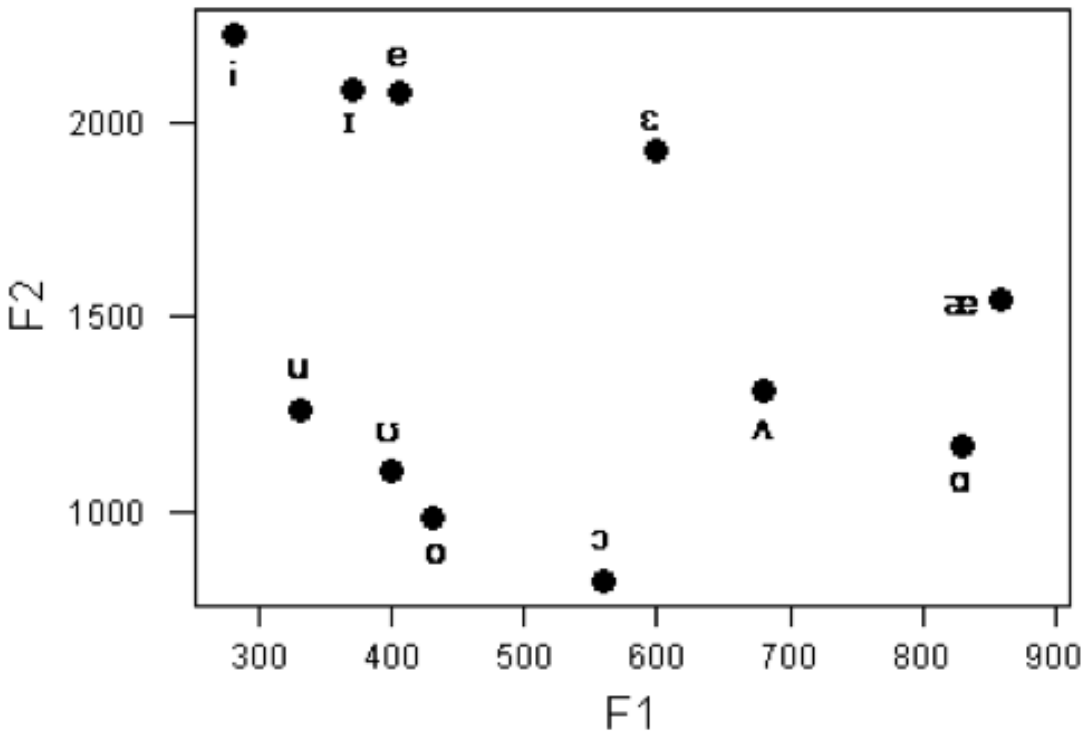
- Vowels differ in a number of dimensions
  - Duration (e.g., tense vs lax)



House (1961), Jacewicz & Fox (2007), Tauberer & Evanini (2009)

Key acoustic dimensions of vowels

- Vowels differ in a number of dimensions
  - Duration (e.g., tense vs lax)
  - F1 x F2



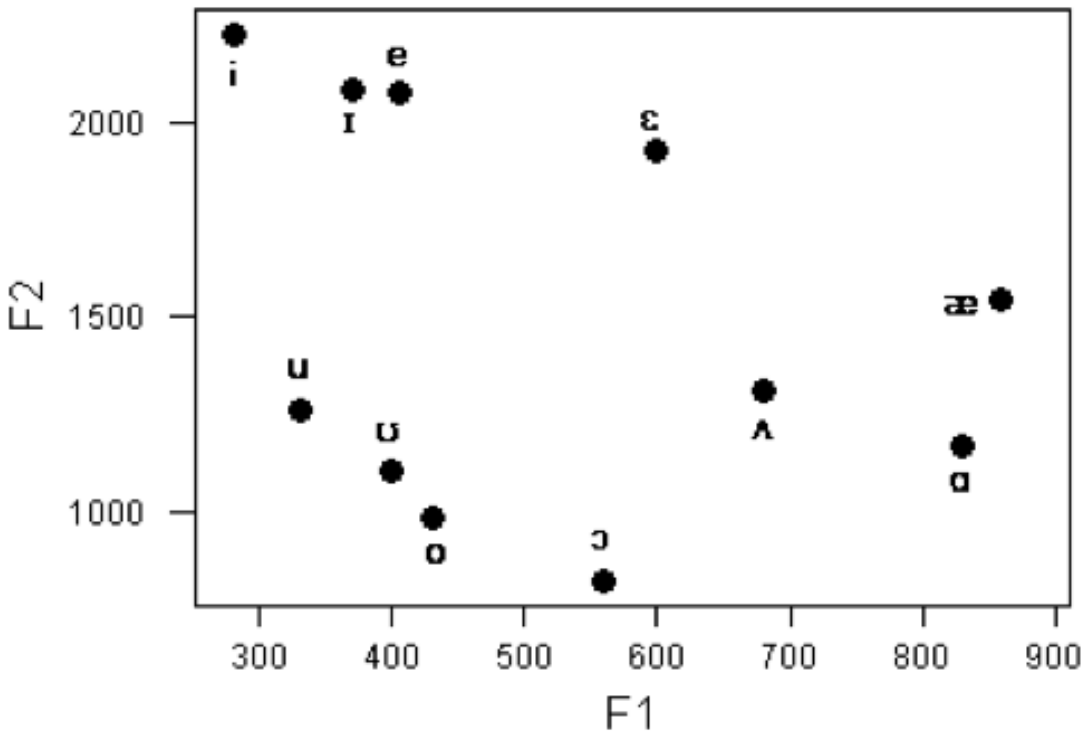
Peterson & Barney (1952), Fant (1960), Ladefoged (1963)

Key acoustic dimensions of vowels

- Vowels differ in a number of dimensions
  - 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.’

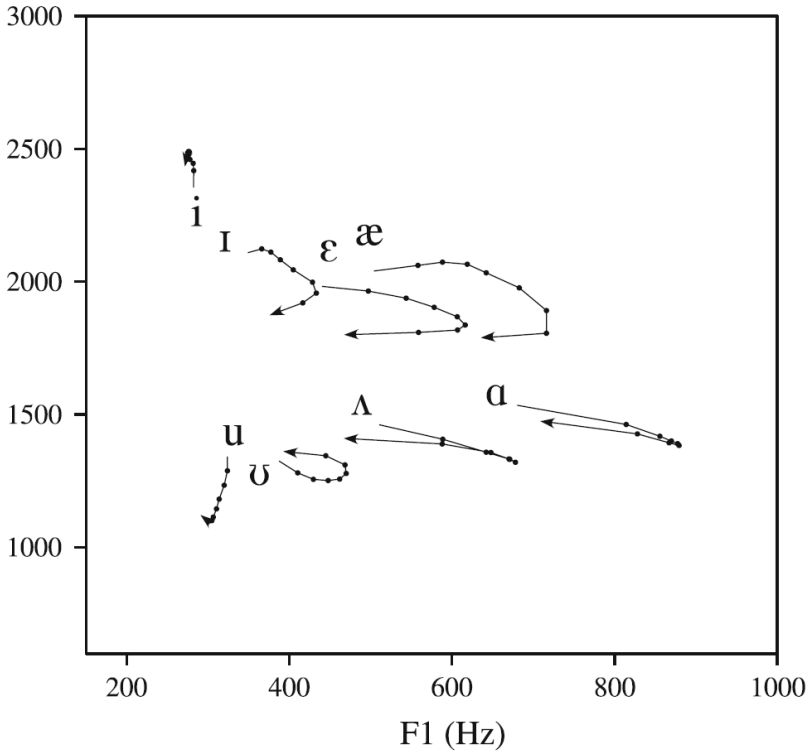
Peterson & Barney (1952: 184)



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

Key acoustic dimensions of vowels

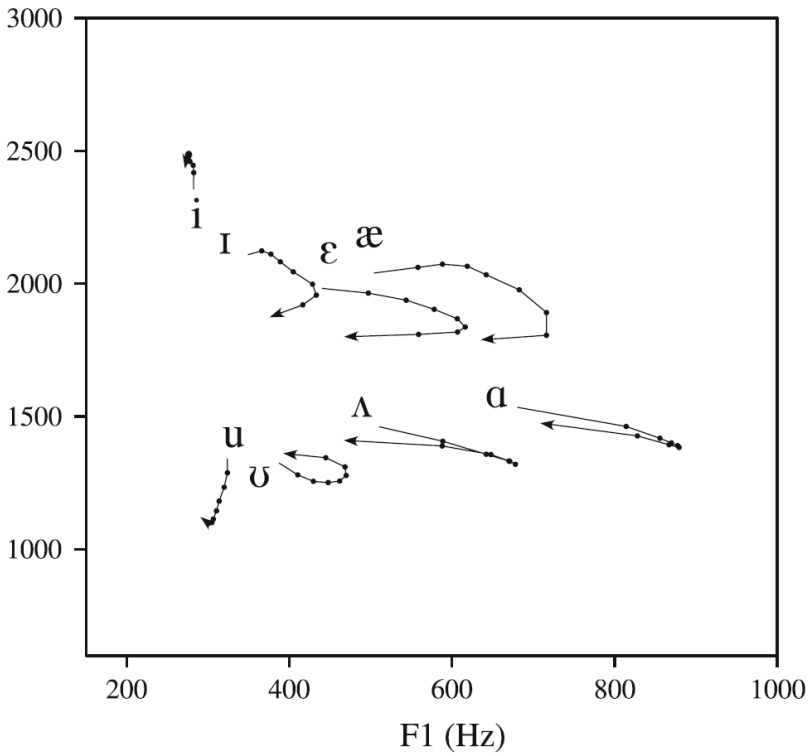
- Vowels differ in a number of dimensions
  - Duration (e.g., tense vs lax)
  - F1 x F2
  - more or less movement in F1/ F2



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

Key acoustic dimensions of vowels

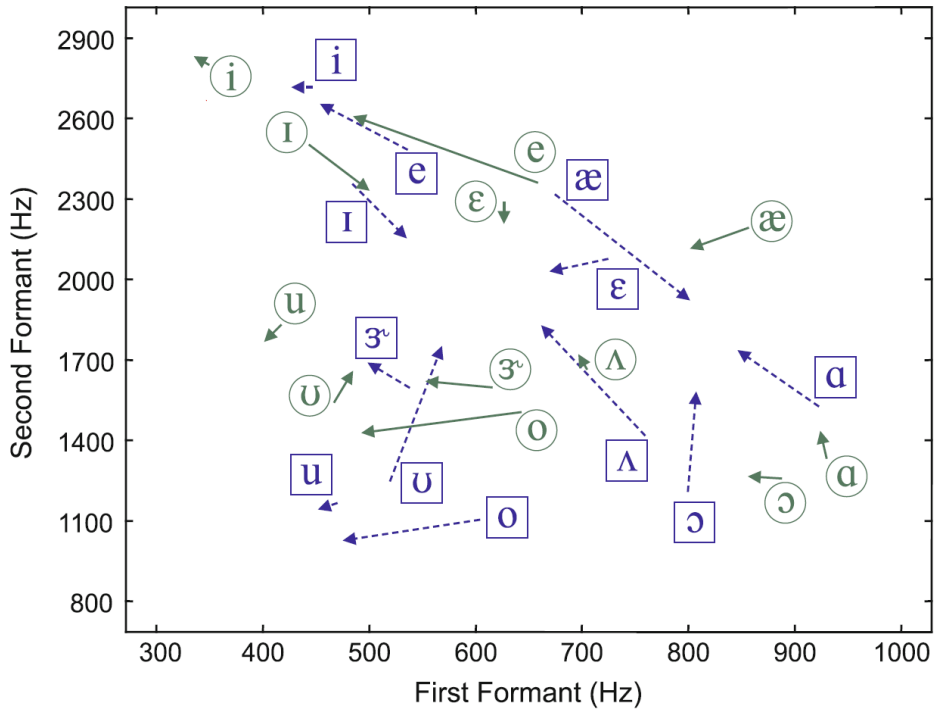
- Vowels differ in a number of dimensions
  - Duration (e.g., tense vs lax)
  - F1 x F2
  - more or less movement in F1/ F2

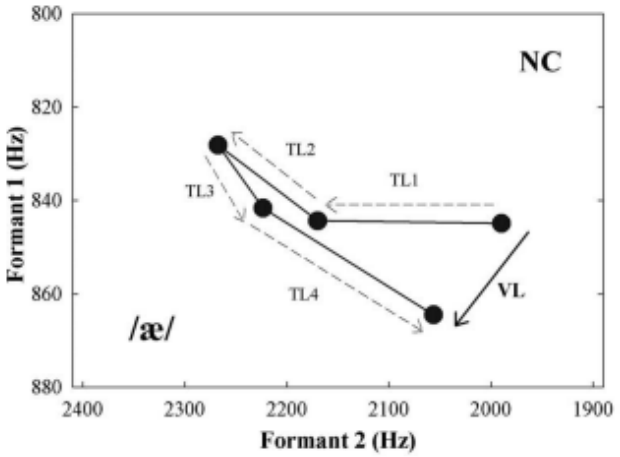


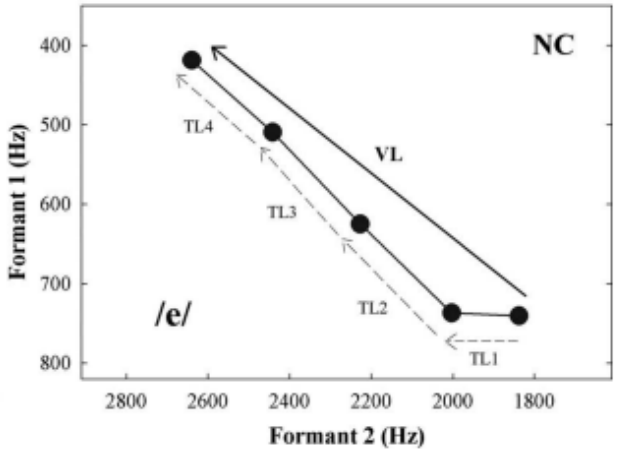
➢ What combination of these dimensions best describes how vowels vary across varieties?

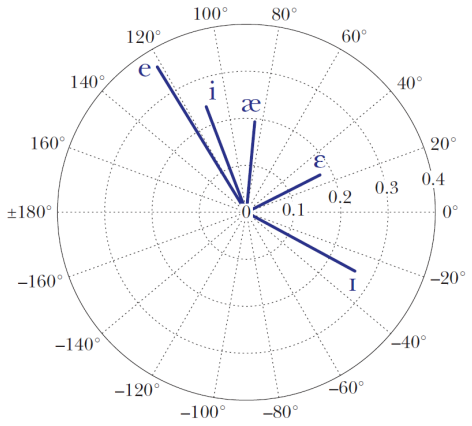
Nearey & Assmann (1986), Watson & Harrington (1999), Fox & Jacewicz (2009), Williams et al (2014)

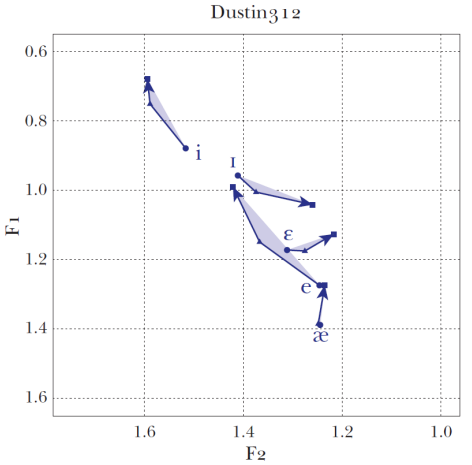
How do we capture ‘more or less movement of F1/F2’?







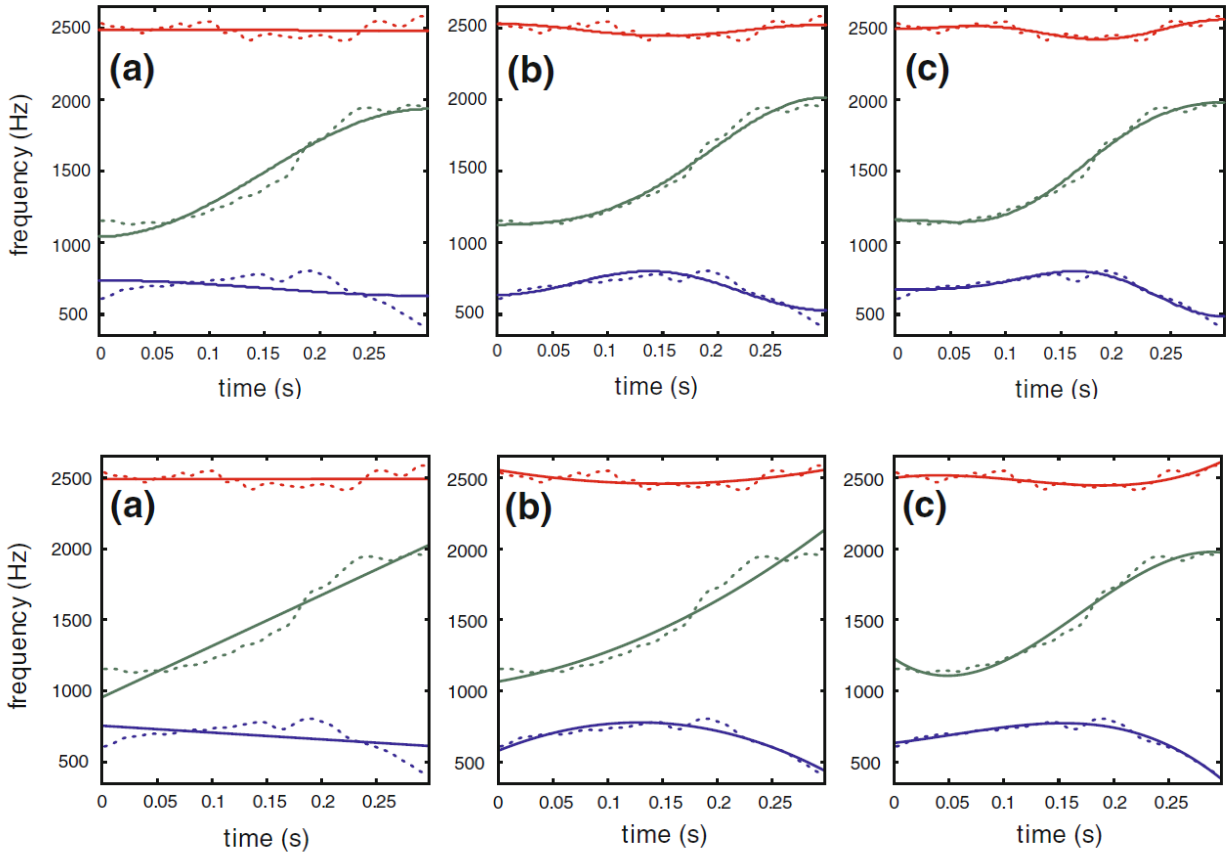


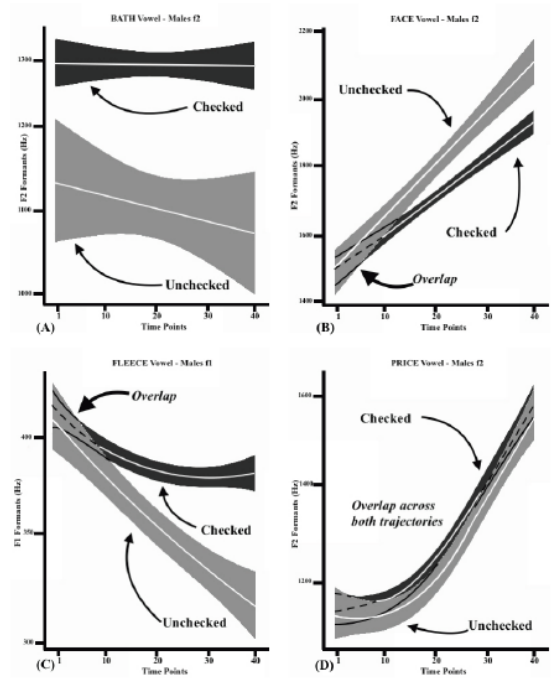


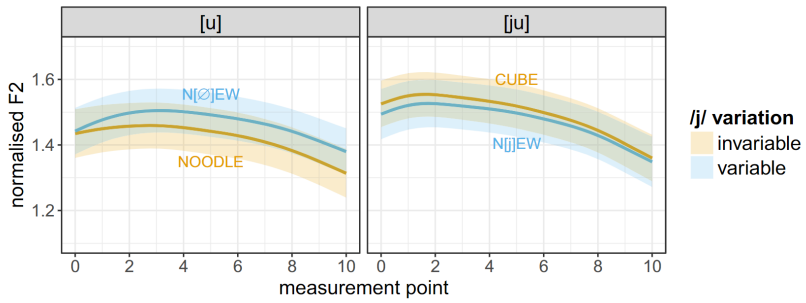
multi-point parameterization
Vector Length, Trajectory Length
Vector Angle …

Fox & Jacewicz (2009), Morrison (2013), Farrington et al (2018)

How do we capture ‘more or less movement of F1/F2’?







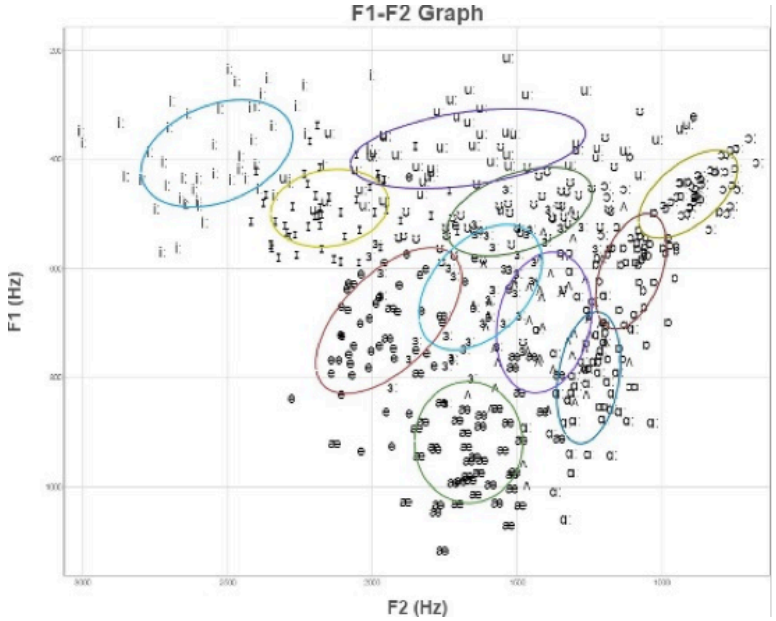
curve parameterization

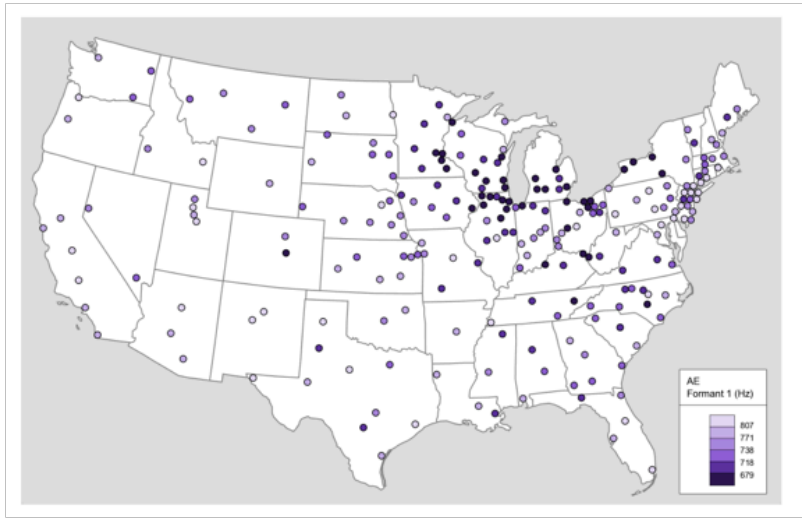
DCT, FDA, SS-ANOVA, GAMMs…

Watson & Harrington (1999), Morrison (2013), Risdal & Kohn (2014)
Docherty et al (2015), Soskuthy et al (2018)

Vowels in English dialects

- differences within + across dialects captured in terms of *static quality* -- position in F1 x F2 space



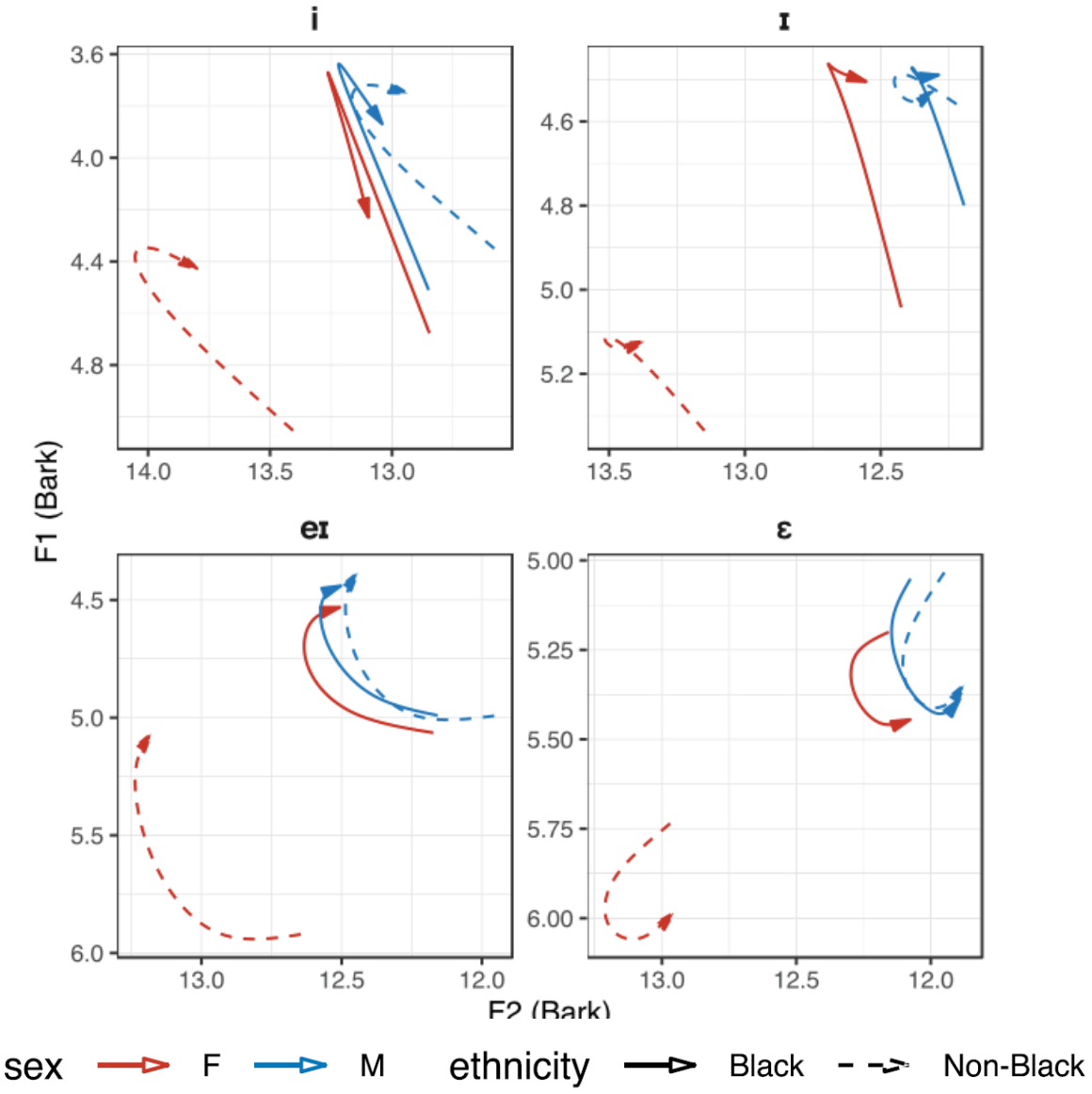


Labov (1972, 1991), Thomas (2001), Clopper et al. (2005), Labov et al. (2006), Boberg (2018)

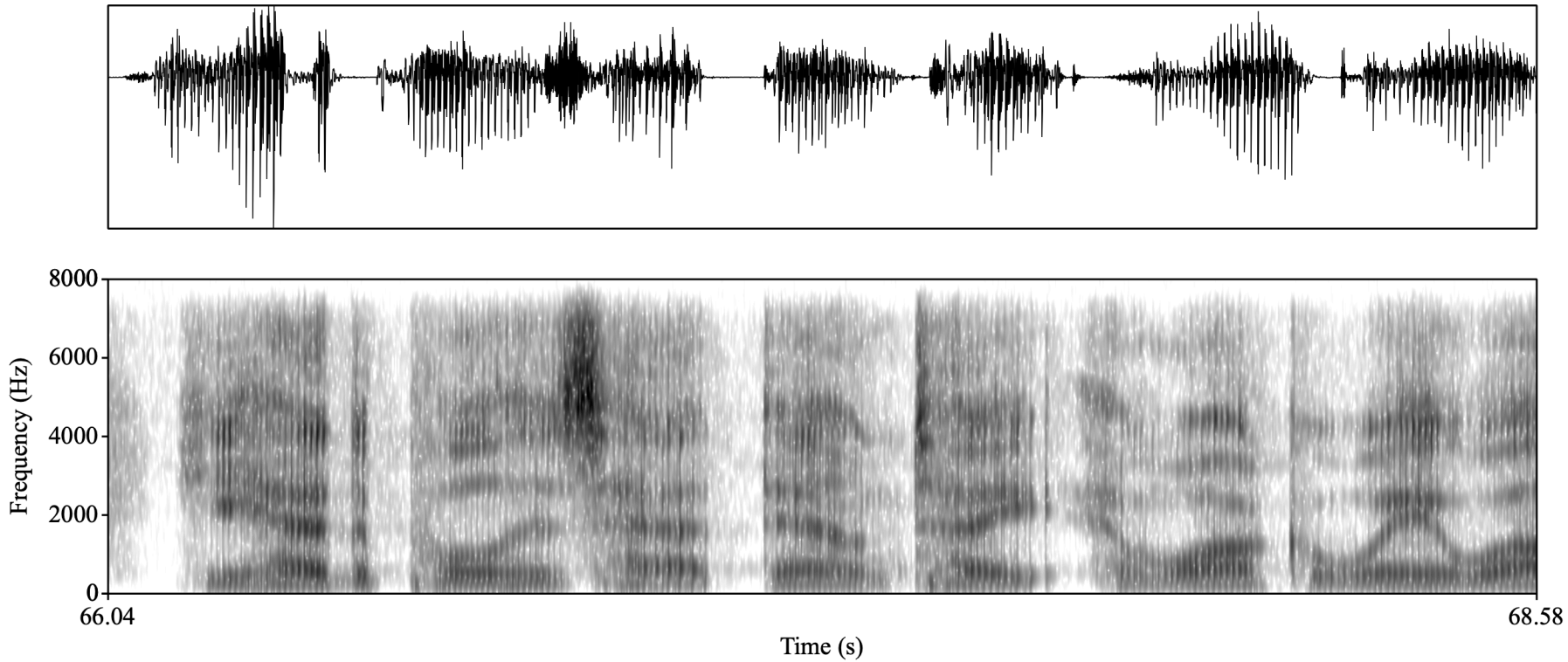
Vowels in English dialects

- 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)



‘duration, likely along with spectral change over time, may be a part of a package of acoustic distinctions that signals both dialect and vowel category information’

Fridland et al (2014: 348)

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?


Data

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 from 21 points in each vowel (5% increments)
  1. Generate 'candidate' formants (with 8-14 LPC coefficients)
  2. Choose candidates closest to manually pruned 'prototype' datasets
- Central 60% of vowel: first & last 20% excluded to avoid consonant effects
- Z-normalized

Mielke et al. (2019)

SPADE

SPeech Across Dialects of English

- 40+ public and private corpora from 4 countries
- **This study:** 30 dialects, ~4.6k speakers, ~1.3m tokens

McAuliffe et al. (2019), Sonderegger et al. (2022)

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

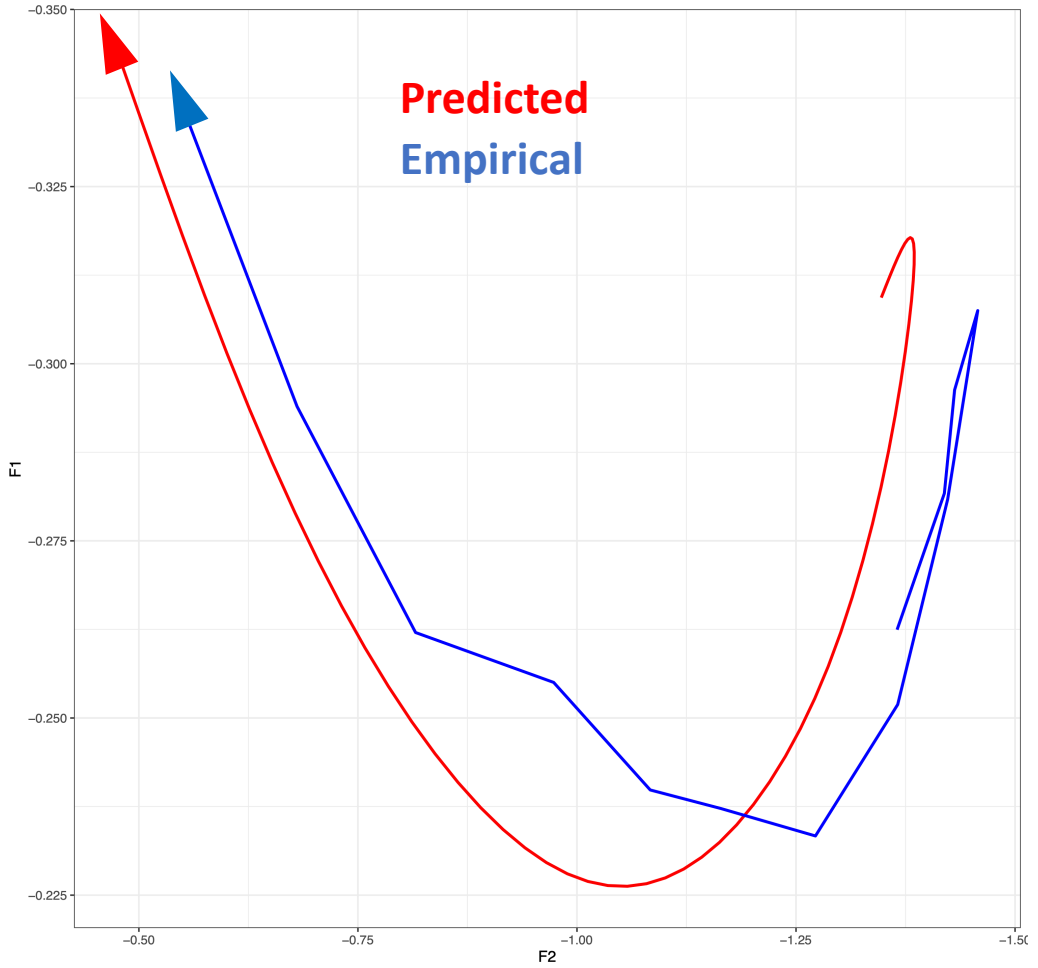
African American
Latino American

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)

Data: models

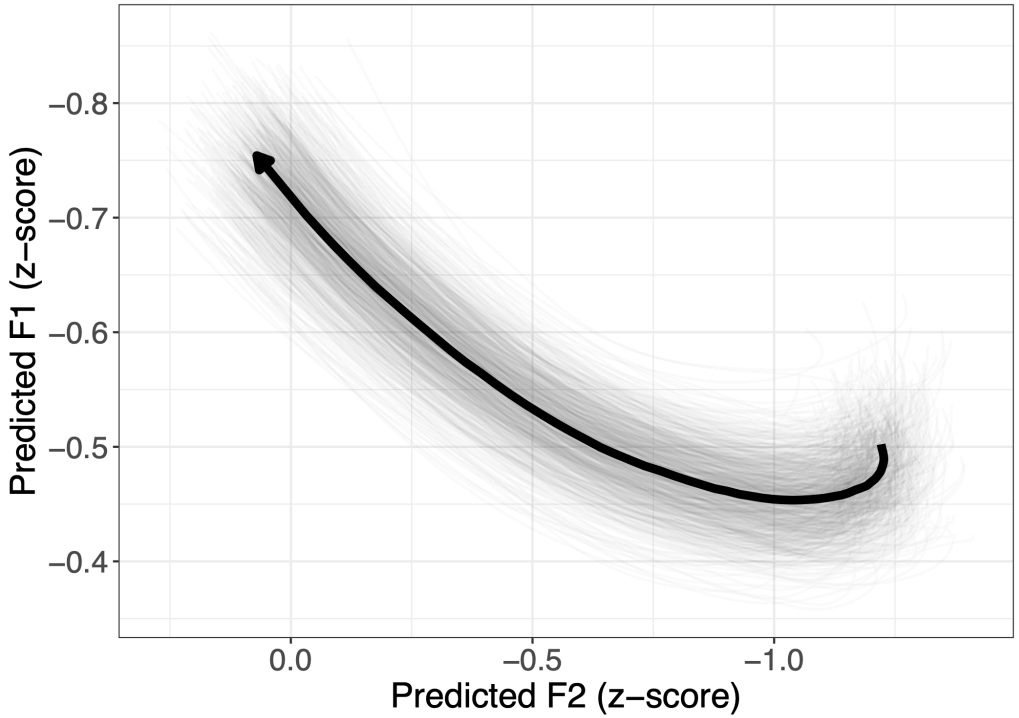
- Generalised Additive (Mixed) Models
  - fit non-linear formant trajectory over sampled timepoints
  - Speakers and words = random smooths



Wood (2011, 2017), Sóskuthy (2017, 2021), Tanner (2023), Bates et al. (2015)

Data: models

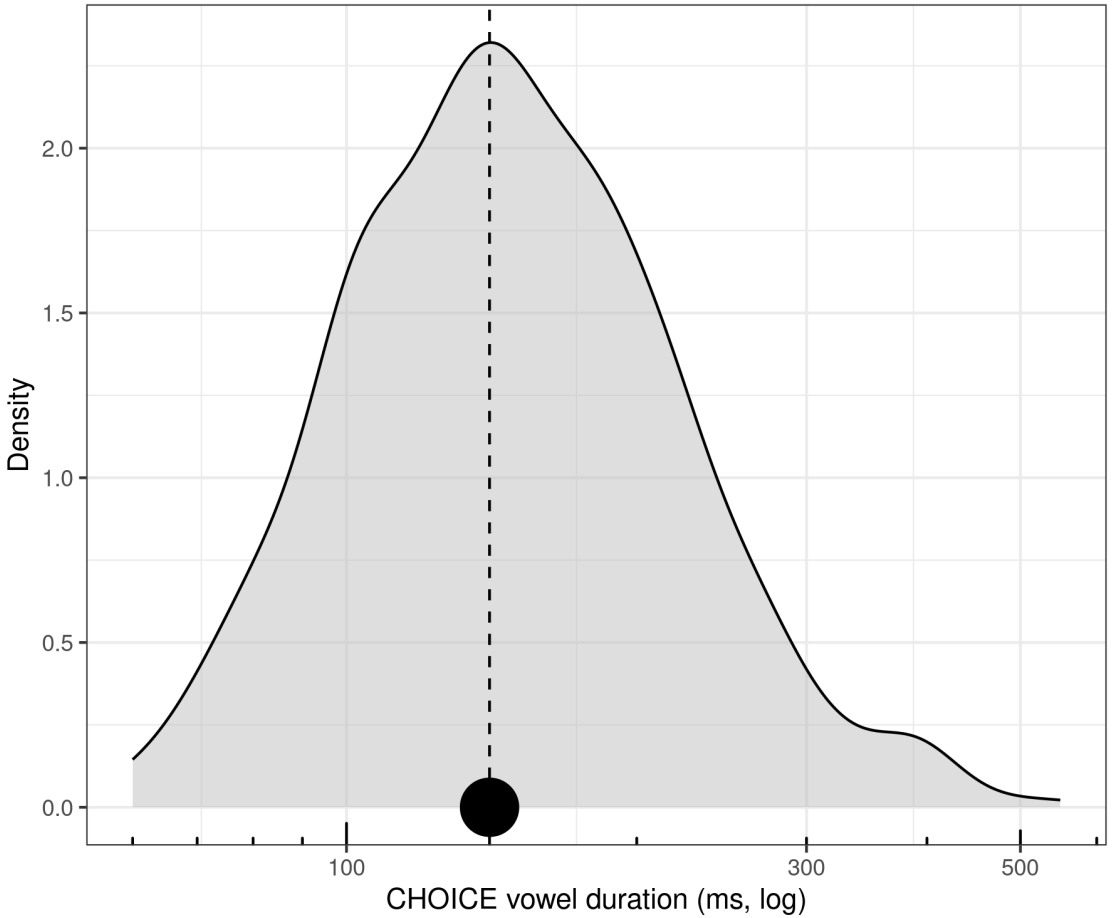
- Generalised Additive (Mixed) Models
  - fit non-linear formant trajectory over sampled timepoints
  - Speakers and words = random smooths
  - Predict trajectory from model
  - simulate trajectories from model and take median
  - reflects 'average' trajectory for 'average' speaker (and ‘average’ word) for that dialect



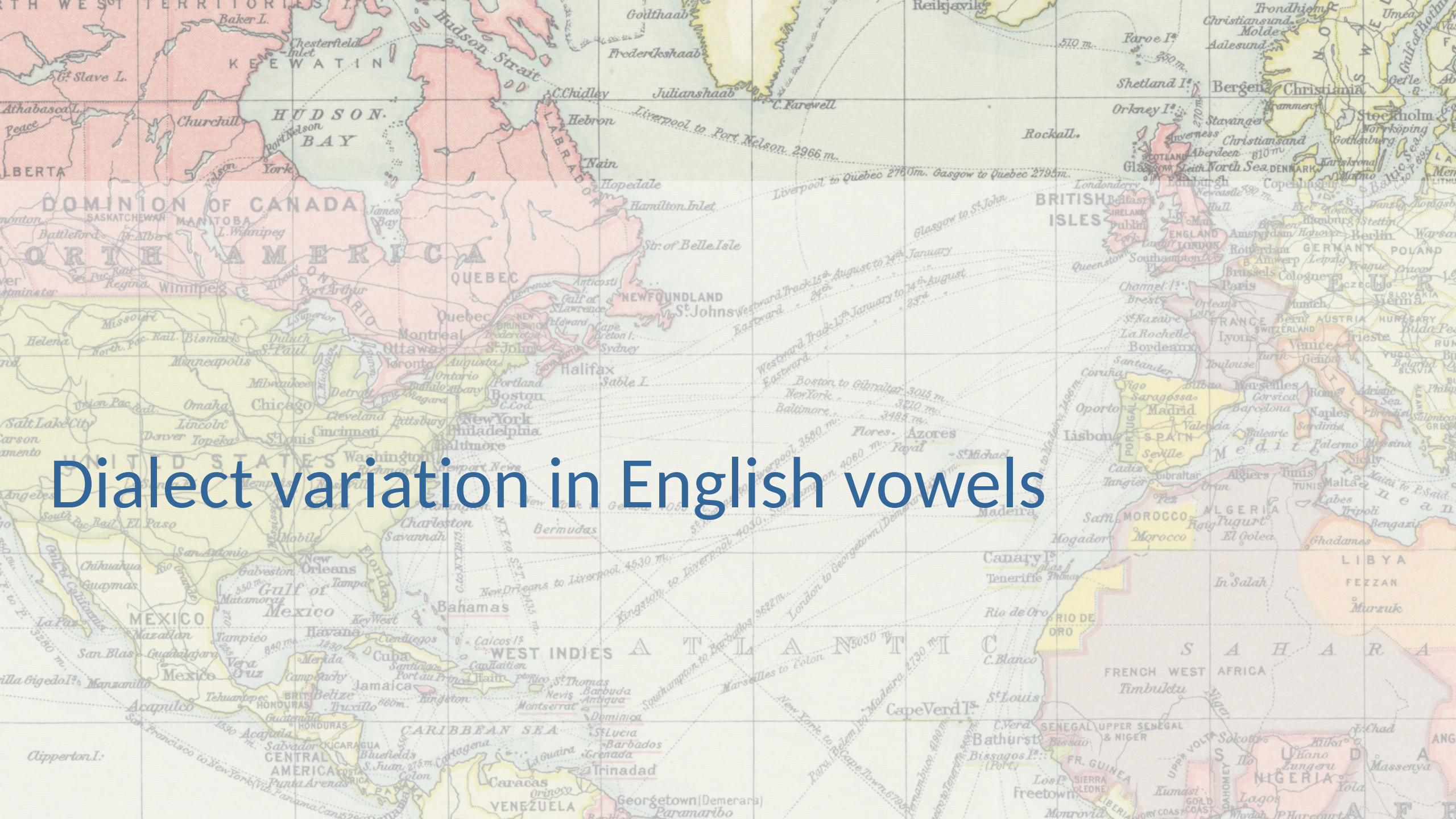
Wood (2011, 2017), Sóskuthy (2017, 2021), Tanner (2023), Bates et al. (2015)

Data: models

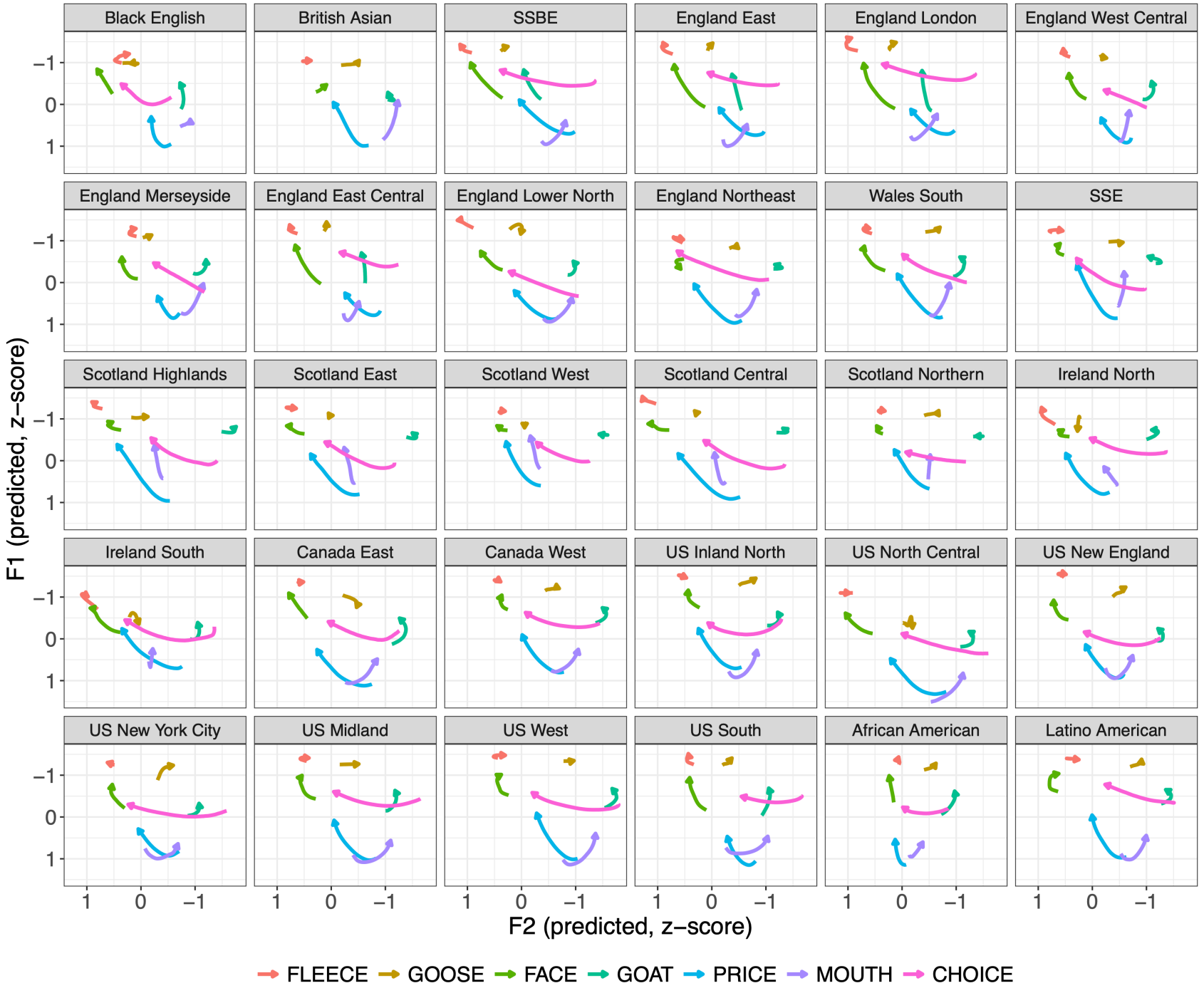
- Duration
  - Linear mixed-effects model of log(duration)
  - Speakers and words = random intercepts
- take model intercept as dialect-predicted duration

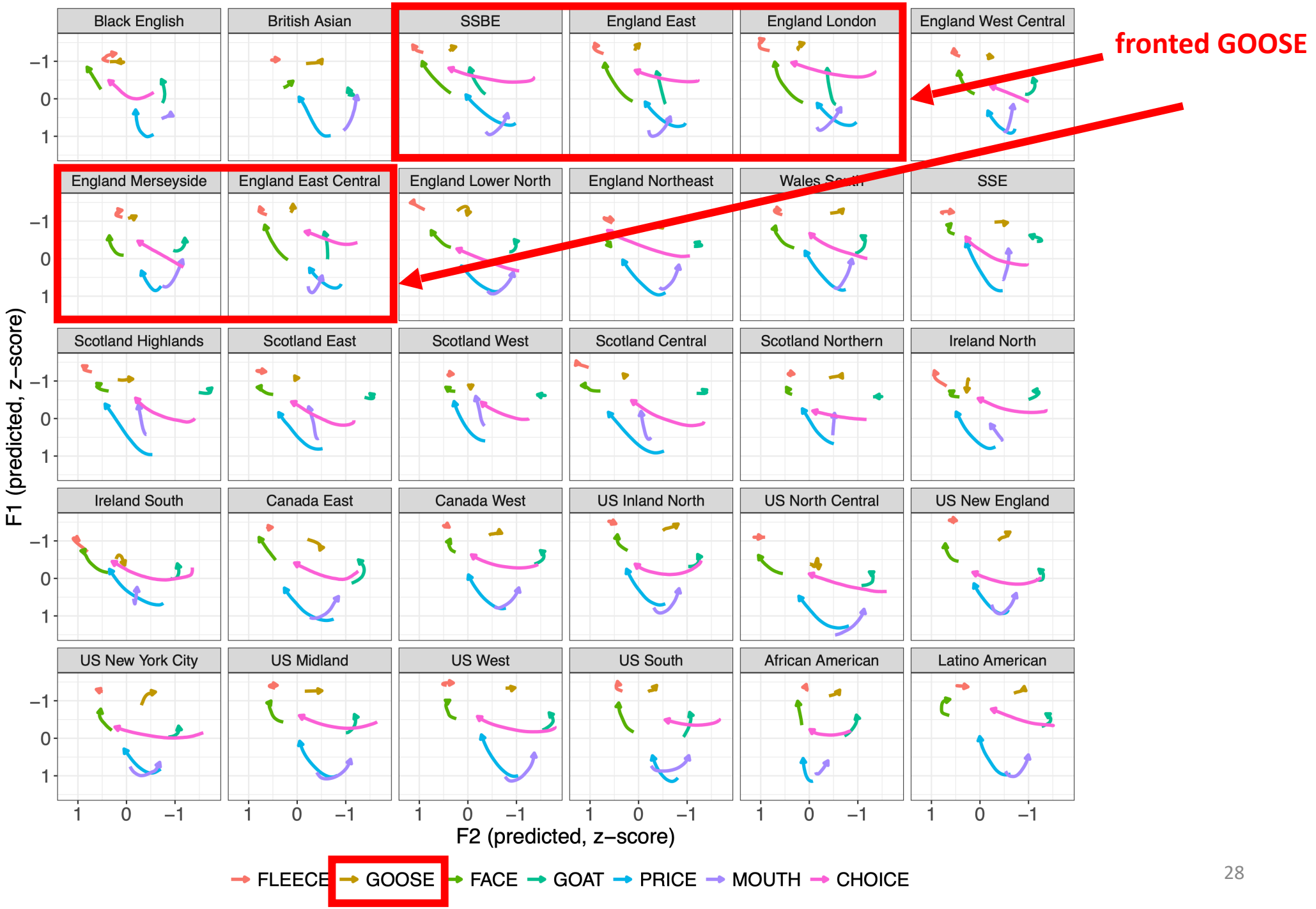


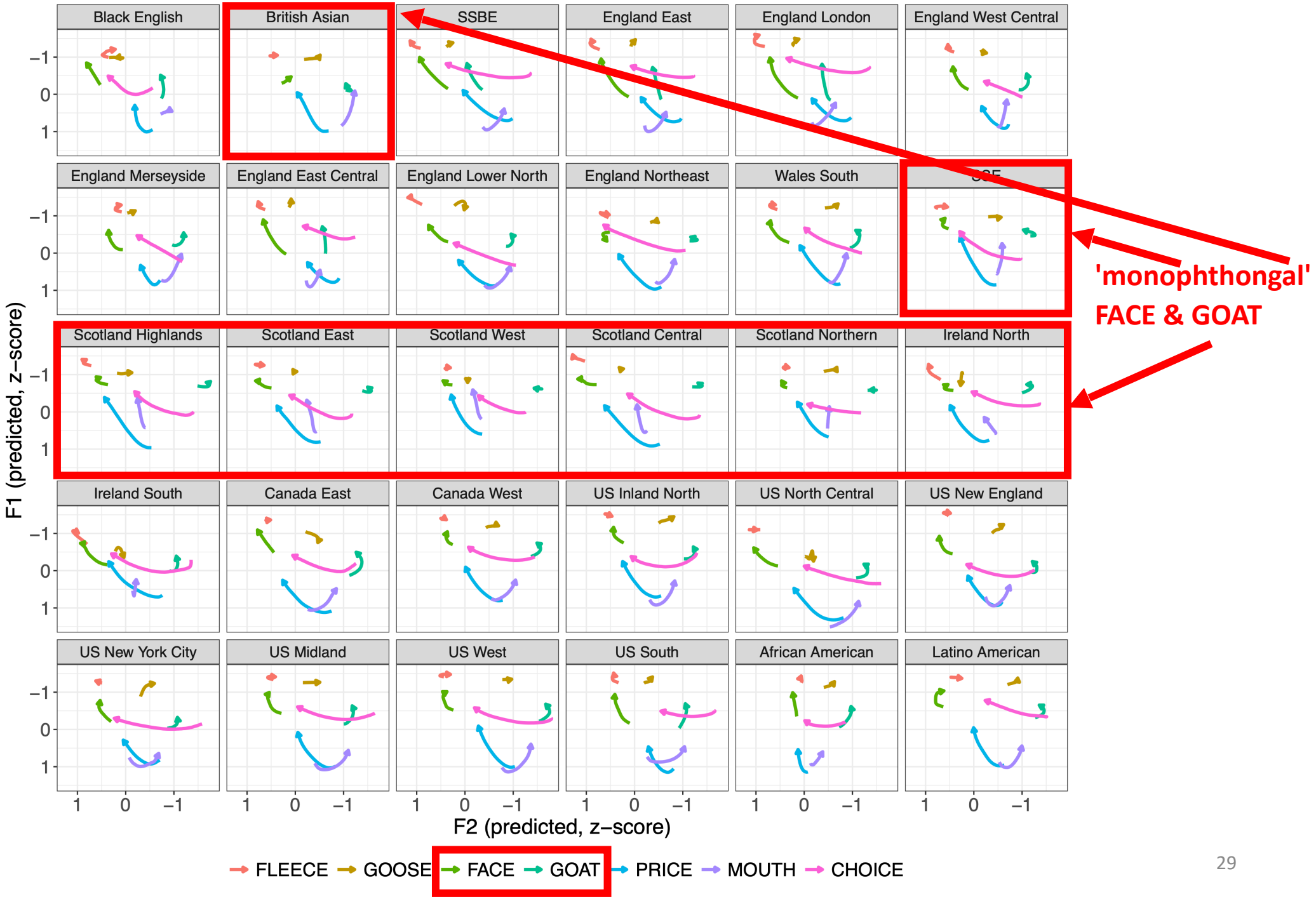
Wood (2011, 2017), Sóskuthy (2017, 2021), Tanner (2023), Bates et al. (2015)

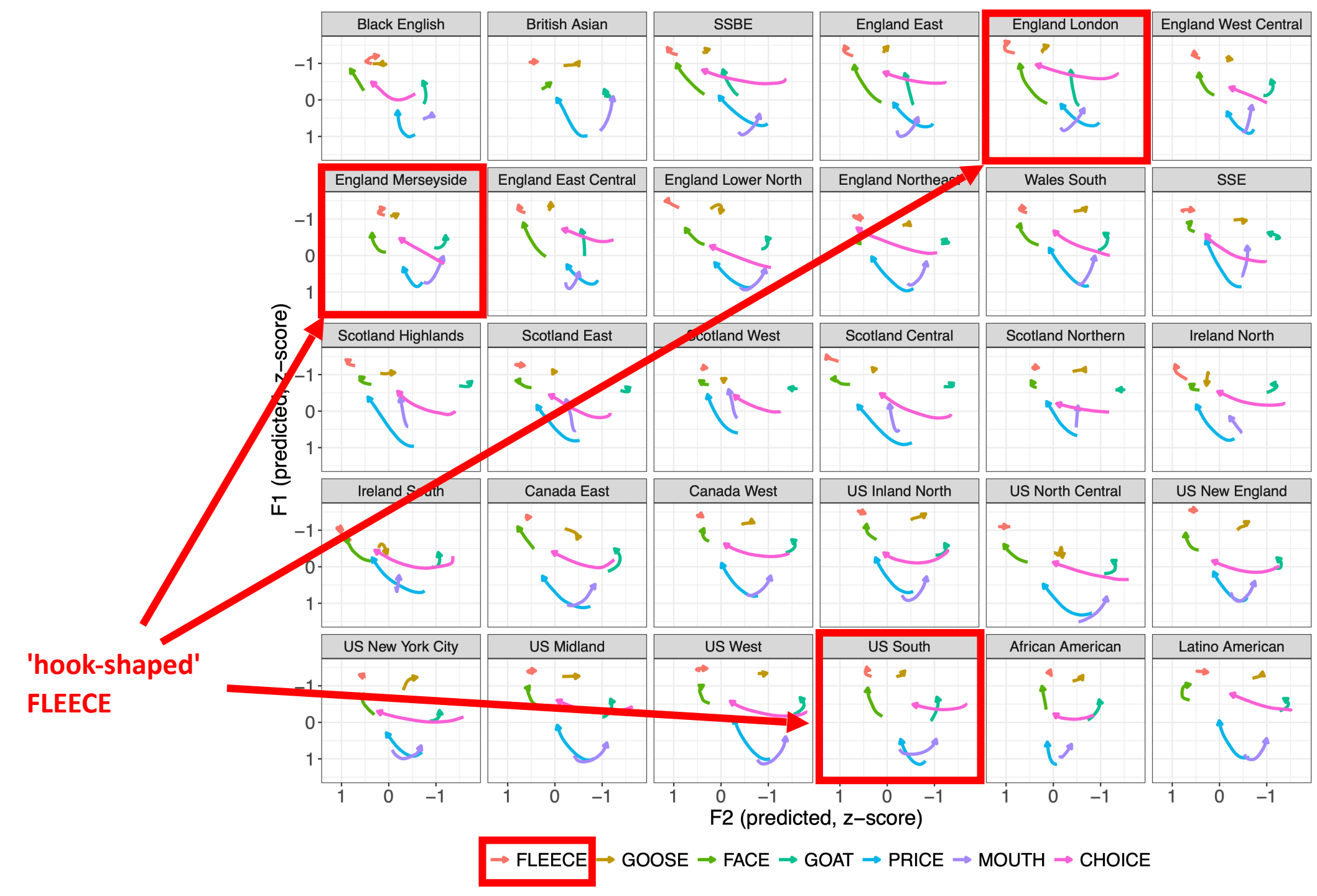


Dialect variation in English vowels

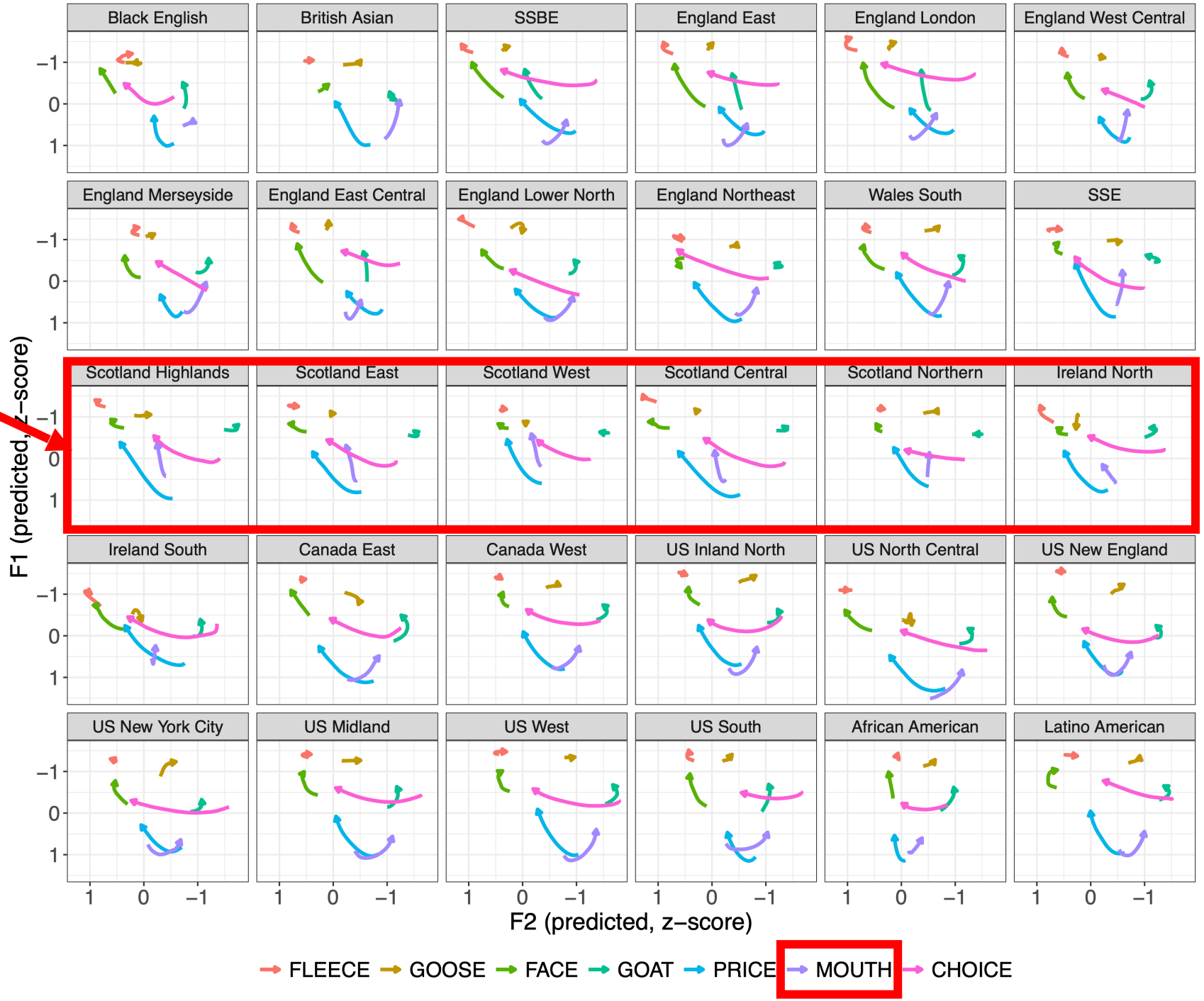

Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada 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

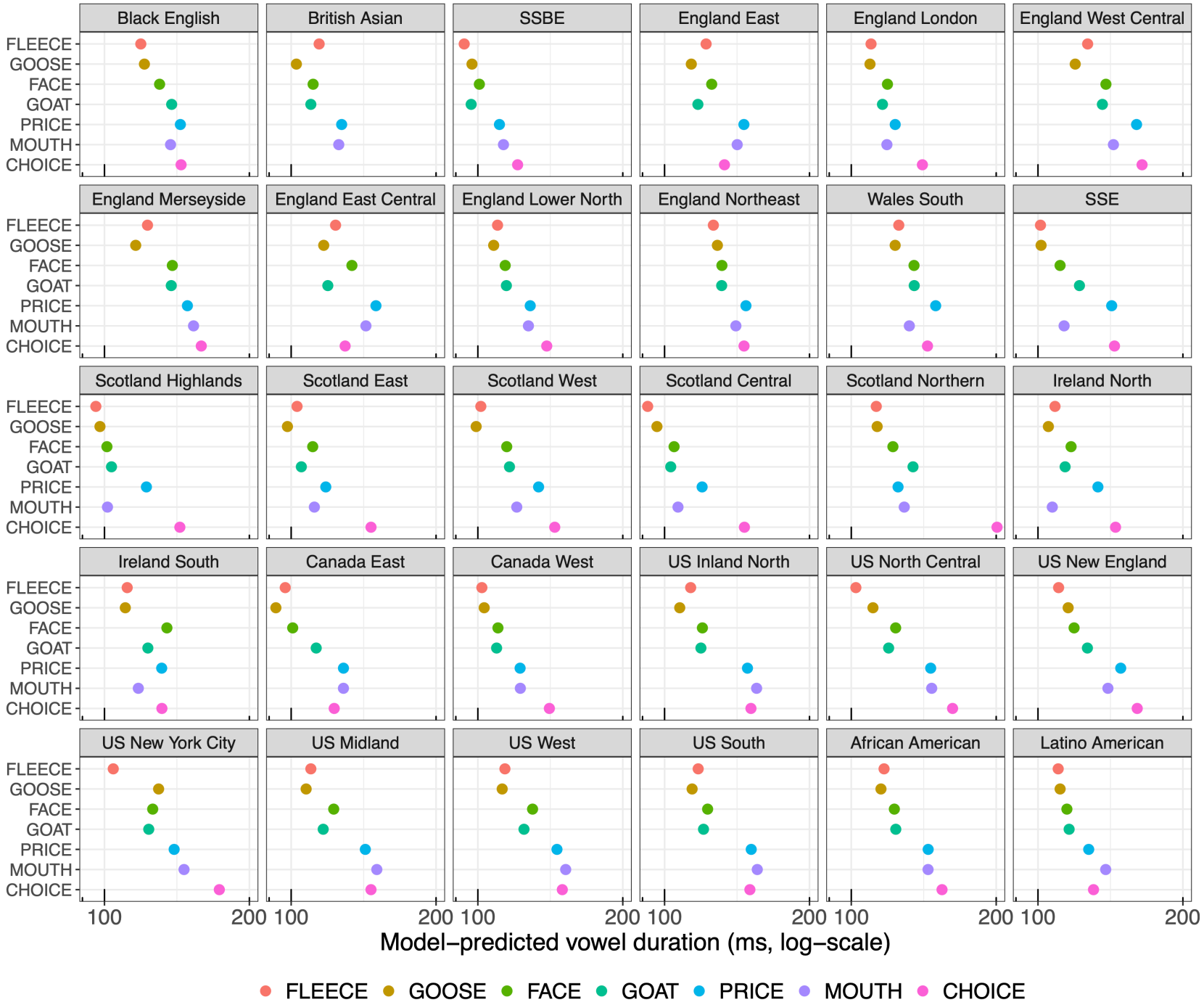

fronted GOOSE
Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada 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

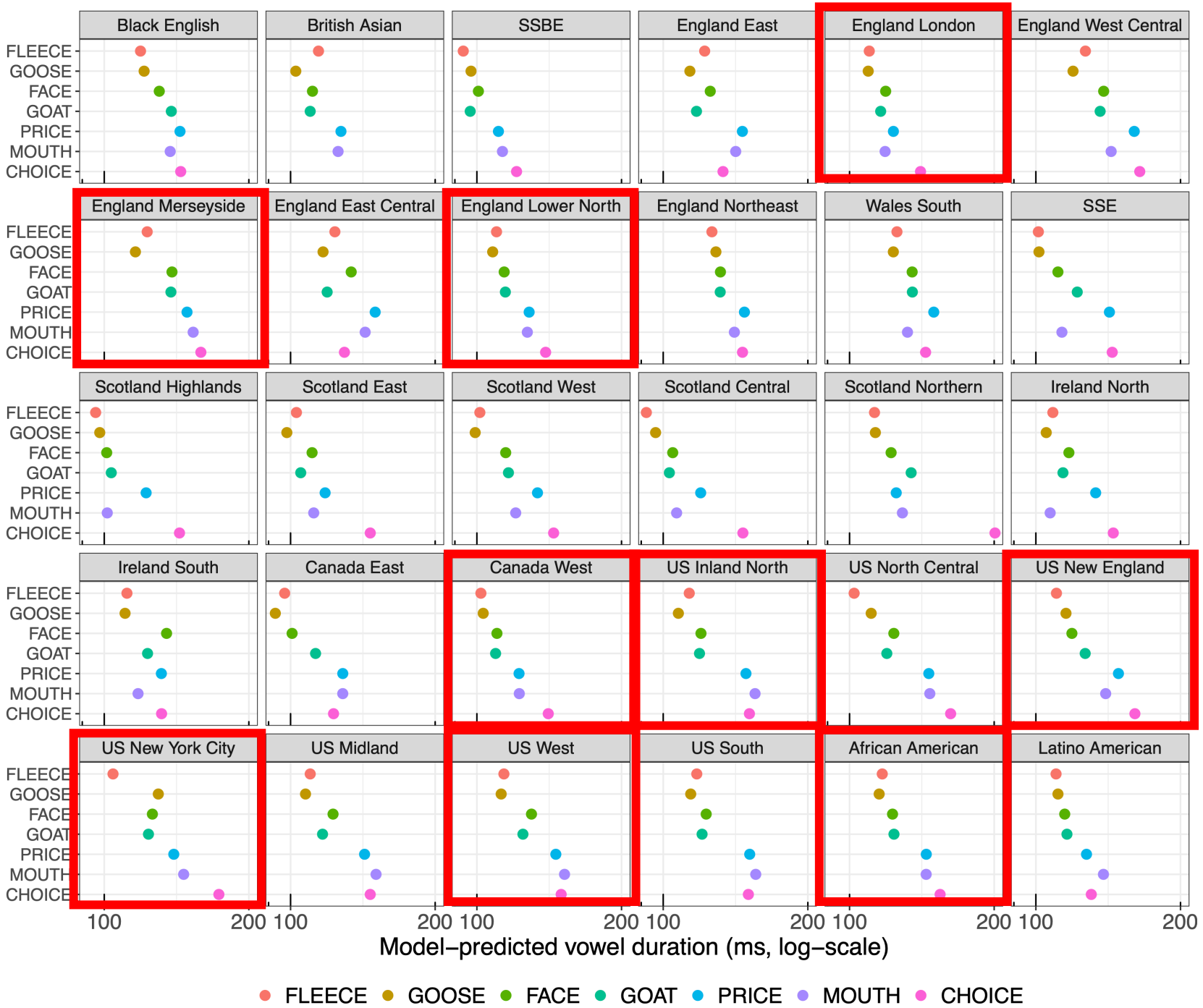
front-raising MOUTH

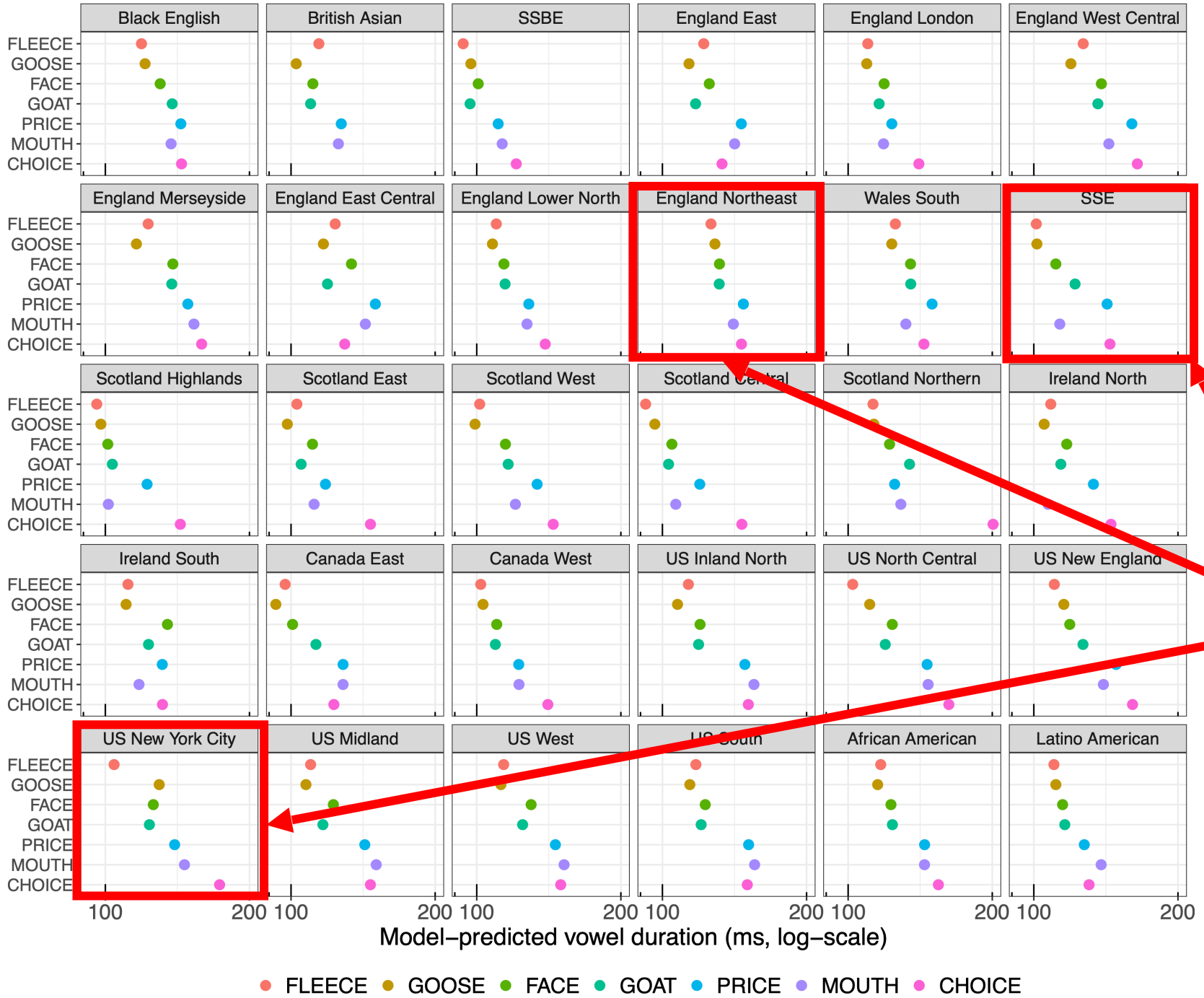



Black English
British Asian
SSBE
England East
England London
England West Central
England Merseyside
England East Central
England Lower North
England Northeast
Wales South
SSE
Scotland Highlands
Scotland East
Scotland West
Scotland Central
Scotland Northern
Ireland North
Ireland South
Canada East
Canada West
US Inland North
US North Central
US New England
US New York City
US Midland
US West
US South
African American
Latino American
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE
100
200
Model–predicted vowel duration (ms, log–scale)

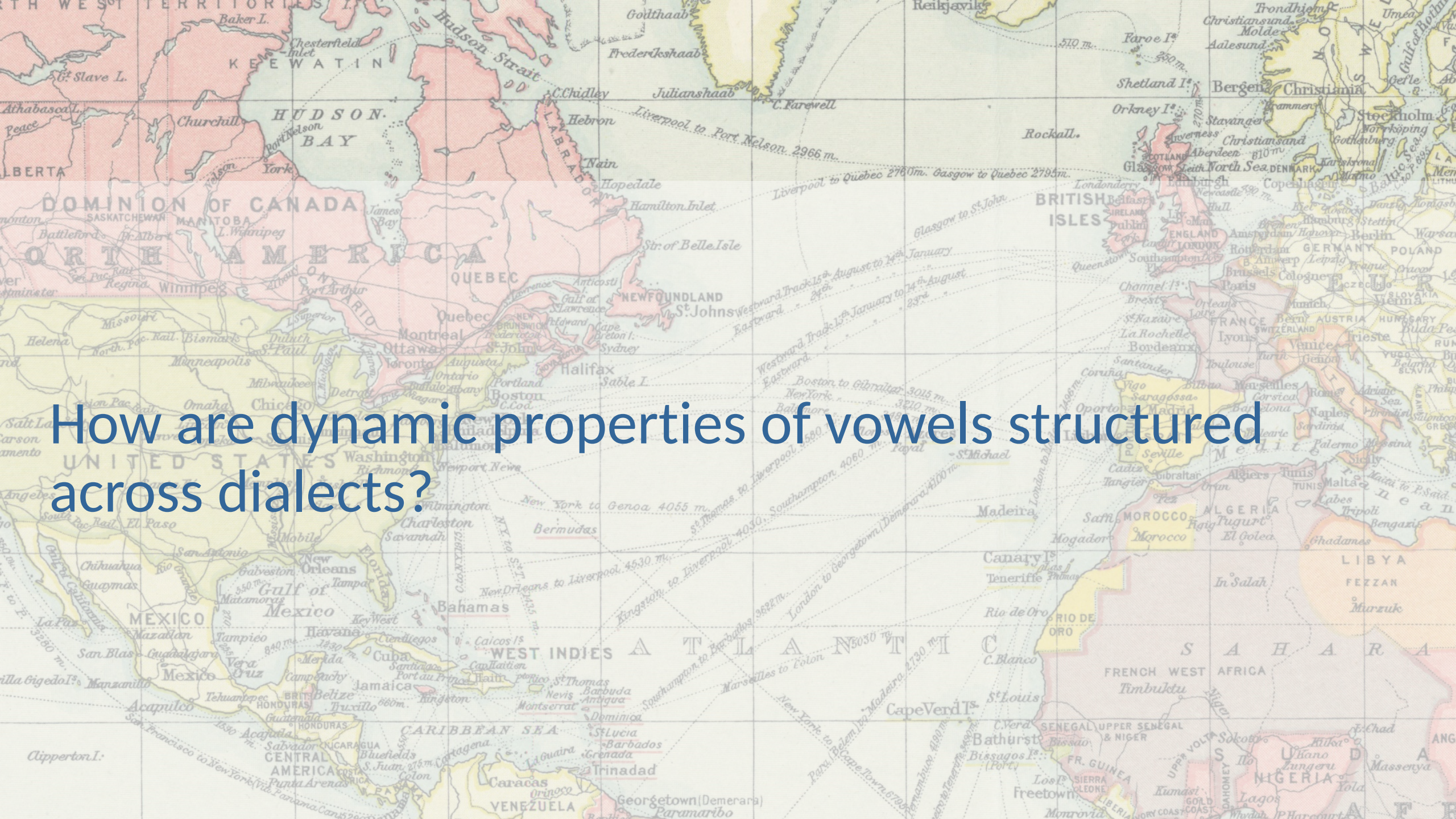
'Intrinsic' duration: Similar pattern across many dialects

{CHOICE, MOUTH, PRICE} > {FACE, GOAT} > {FLEECE, GOOSE}





But also *differences* between dialects



How are dynamic properties of vowels structured across dialects?

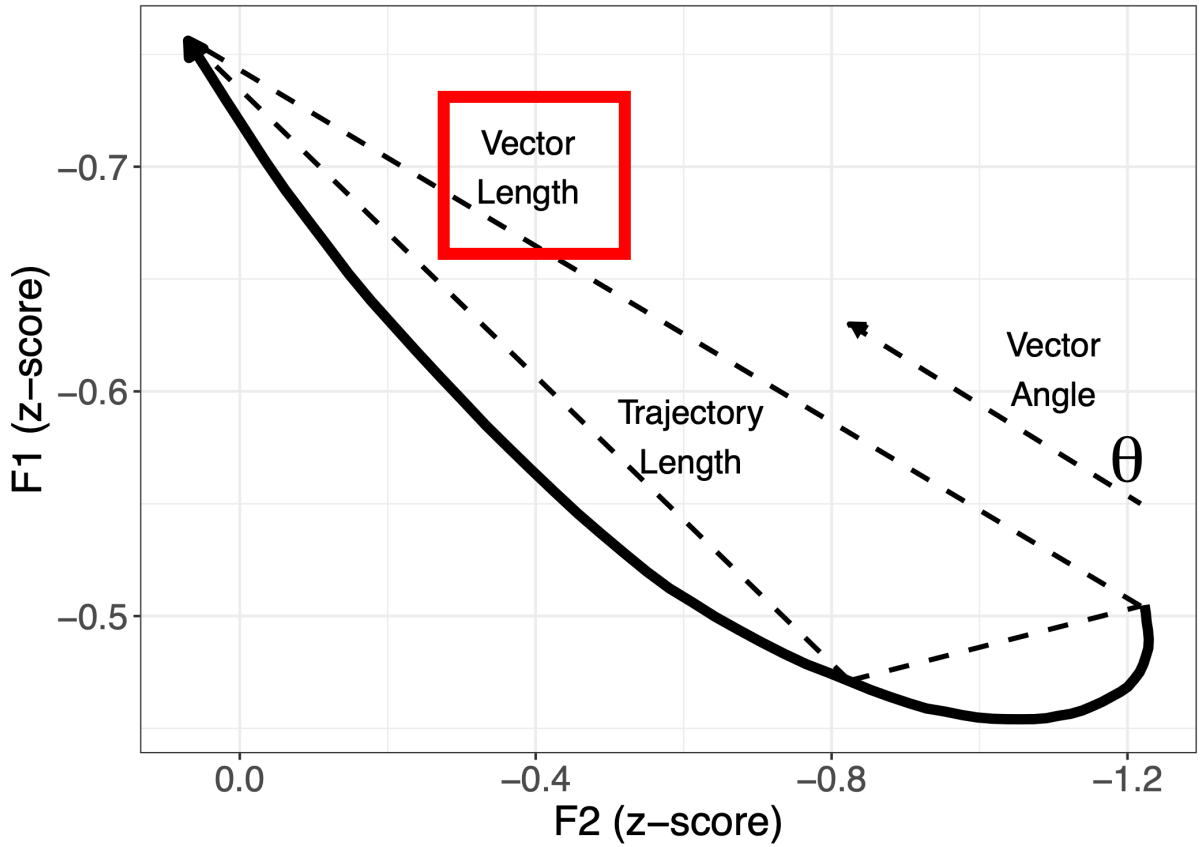
Multi-point parameterization

- *lower dimensional* representation needed to capture essential properties of trajectories across vowels and dialects
- Here: set of measures based around Euclidean distance between formant points

MacDougall & Nolan (2007), Morrison (2013), Watson & Harrington (1999), Williams et al. (2019), Fabricius 2008), Fox & Jacewicz (2009), Farrington et al. (2018)

Capturing movement in F1/F2

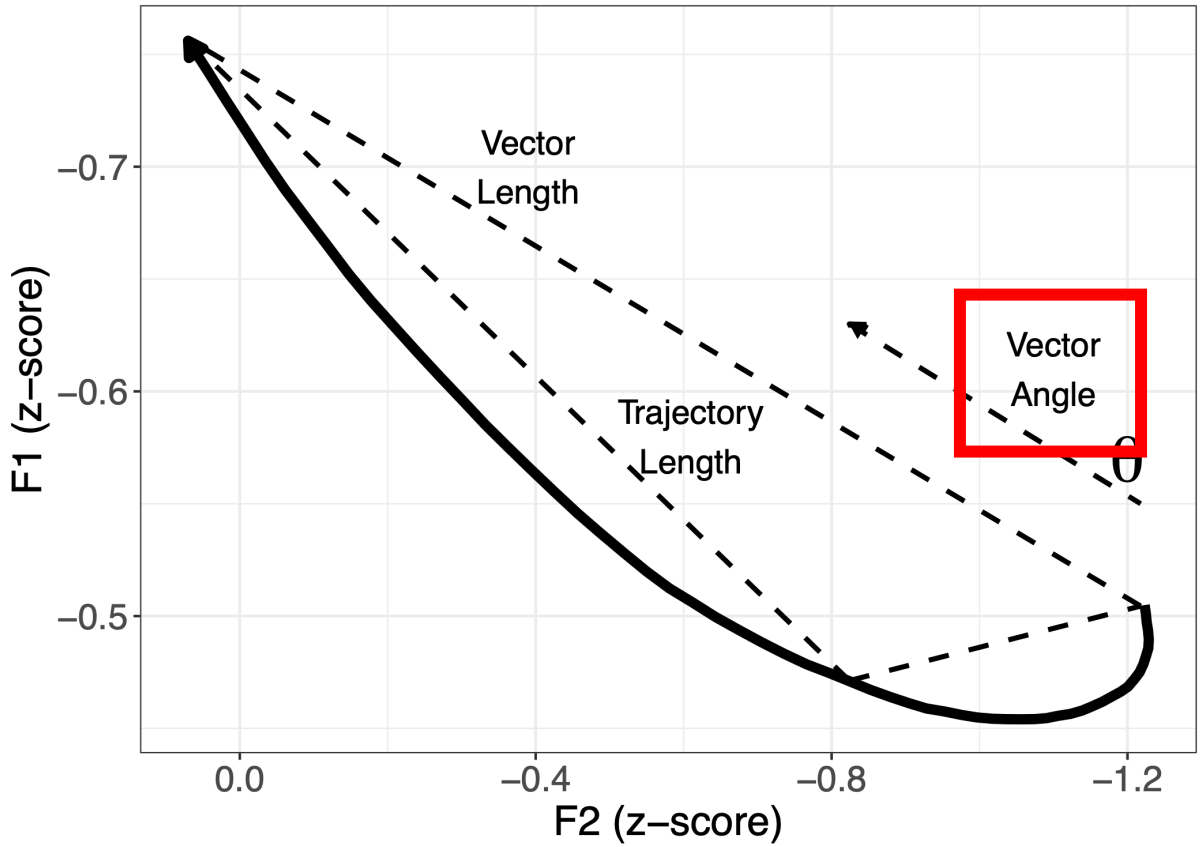
- Vector Length
  - distance between first and last F1,F2 coordinates
  - reflects overall degree of *linear* formant change



MacDougall & Nolan (2007), Morrison (2013), Watson & Harrington (1999), Williams et al. (2019), Fabricius (2008), Fox & Jacewicz (2009), Farrington et al. (2018), Renwick & Stanley (2020)

Capturing movement in F1/F2

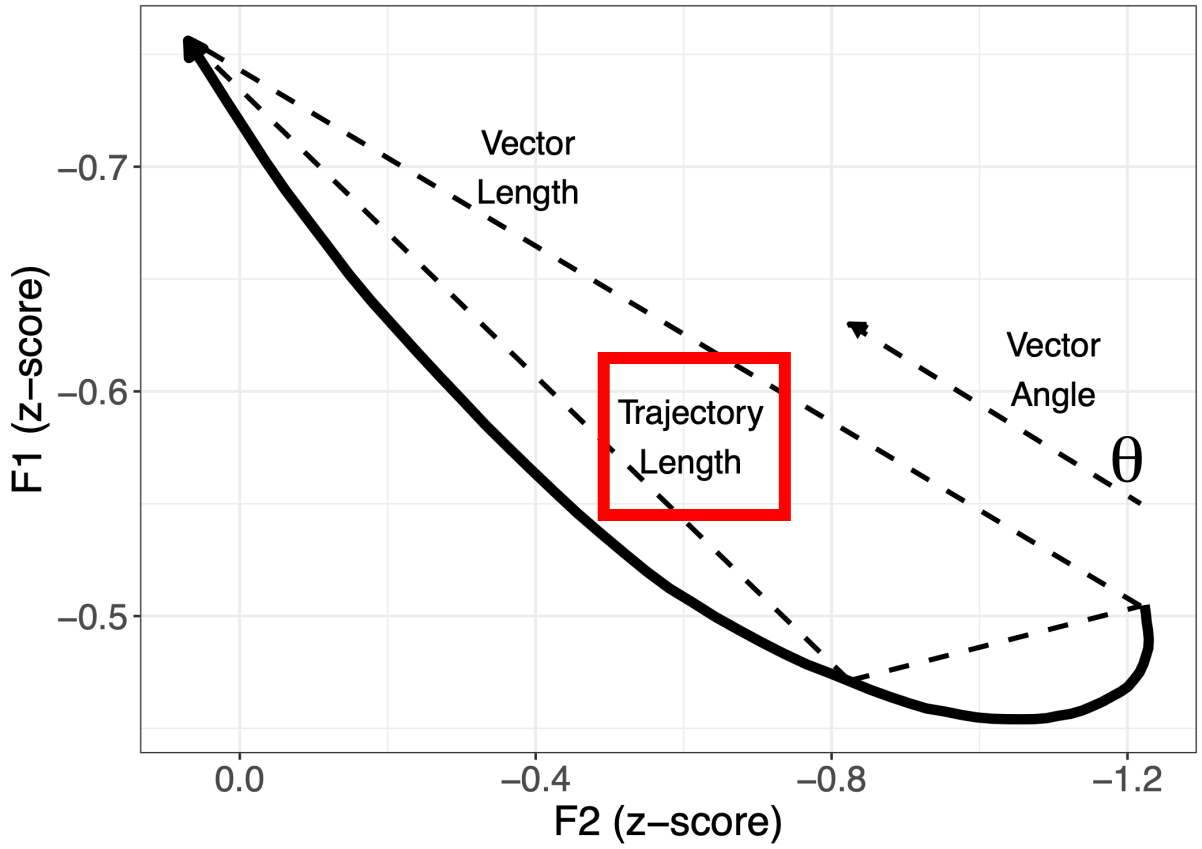
- Vector Angle
  - reflects *direction* of change in 360° space



MacDougall & Nolan (2007), Morrison (2013), Watson & Harrington (1999), Williams et al. (2019), Fabricius (2008), Fox & Jacewicz (2009), Farrington et al. (2018), Renwick & Stanley (2020)

Capturing movement in F1/F2

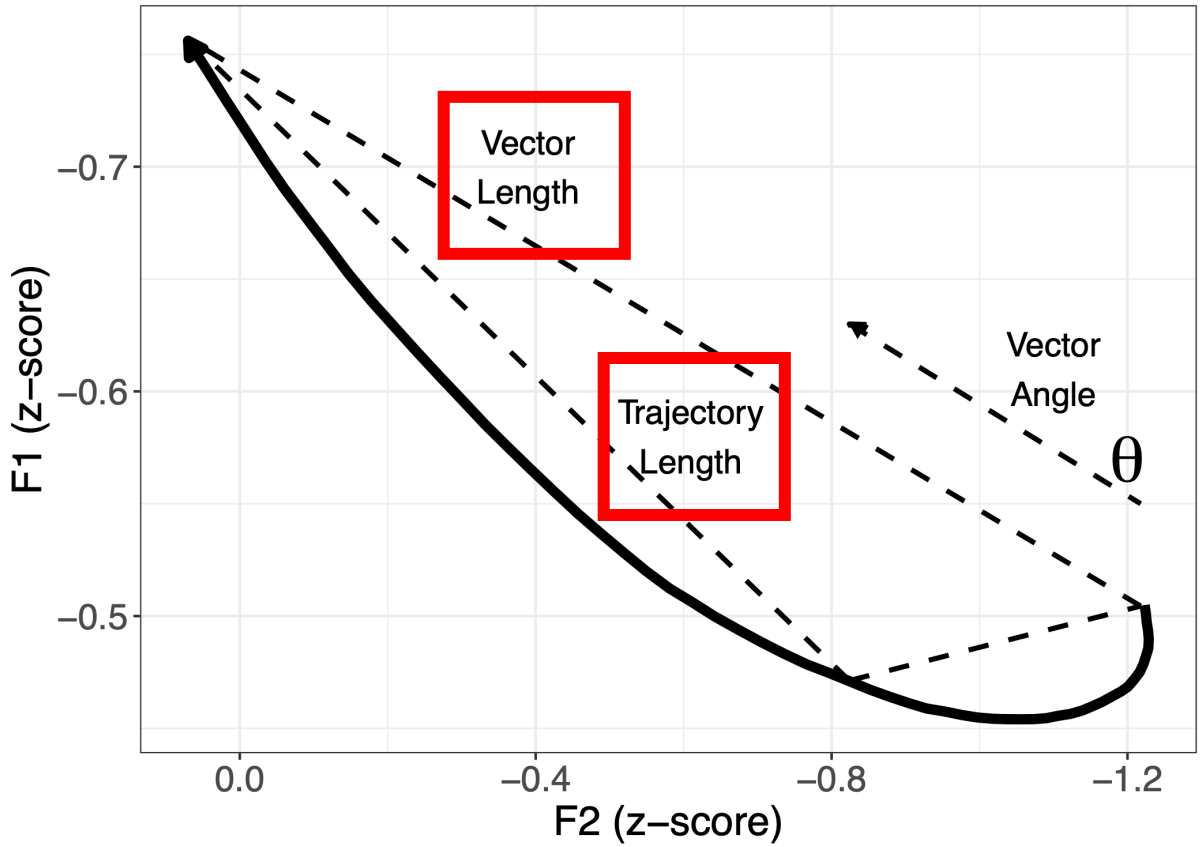
- Trajectory Length
  - sum of Vector Lengths from start-mid & mid-end
- intended to capture *non-linear* formant change, but…
  - highly correlated with Vector Length
  - can't distinguish between linear & non-linear change



MacDougall & Nolan (2007), Morrison (2013), Watson & Harrington (1999), Williams et al. (2019), Fabricius (2008), Fox & Jacewicz (2009), Farrington et al. (2018), Renwick & Stanley (2020)

Capturing movement in F1/F2

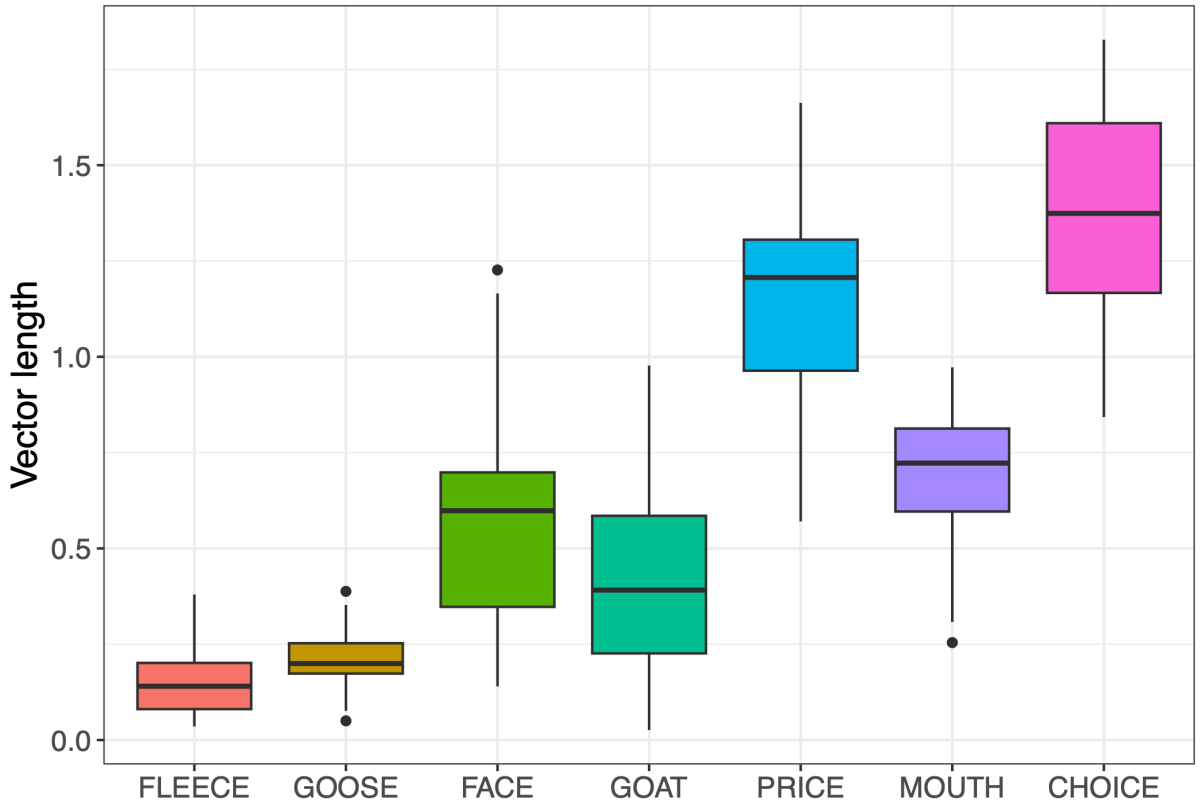
- Vector Length Ratio
  - Vector Length / Trajectory Length
  - expresses *non-linearity* as a *ratio* of overall change



MacDougall & Nolan (2007), Morrison (2013), Watson & Harrington (1999), Williams et al. (2019), Fabricius (2008), Fox & Jacewicz (2009), Farrington et al. (2018), Renwick & Stanley (2020)

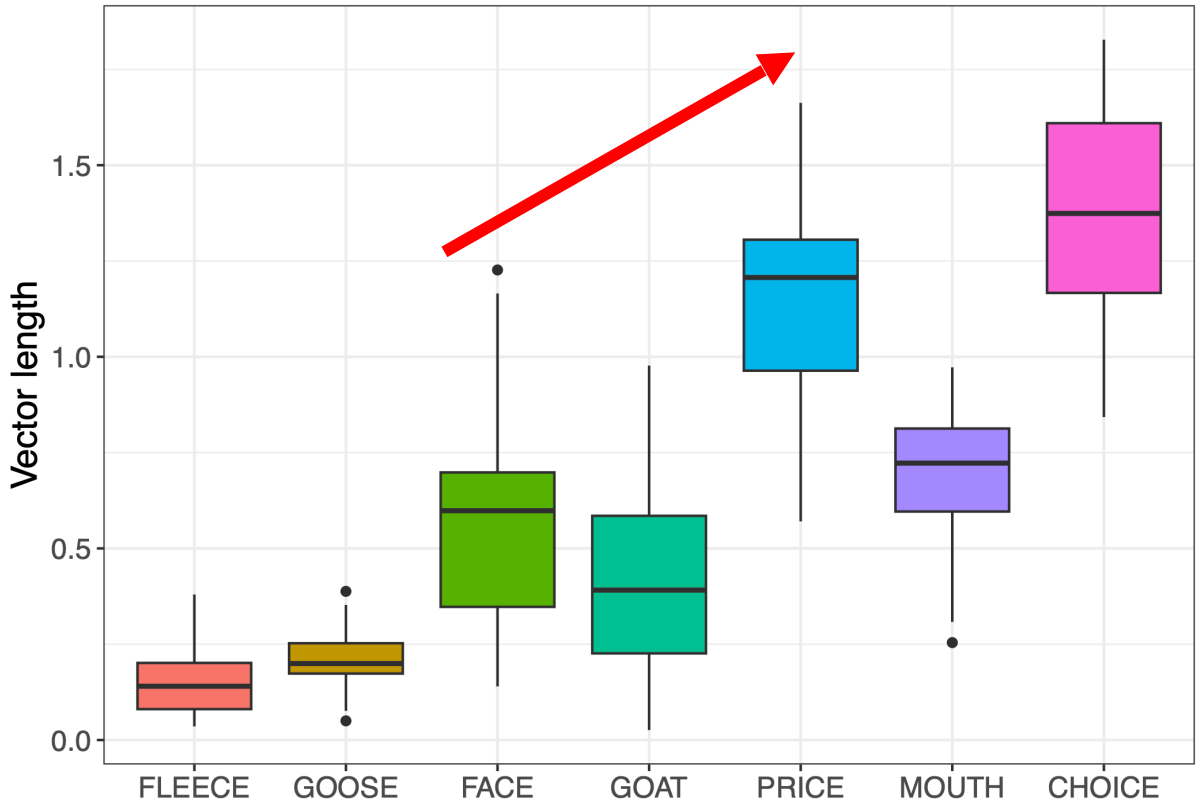
Results: Vector Length

- *broadly* follows monophthong-diphthong continuum



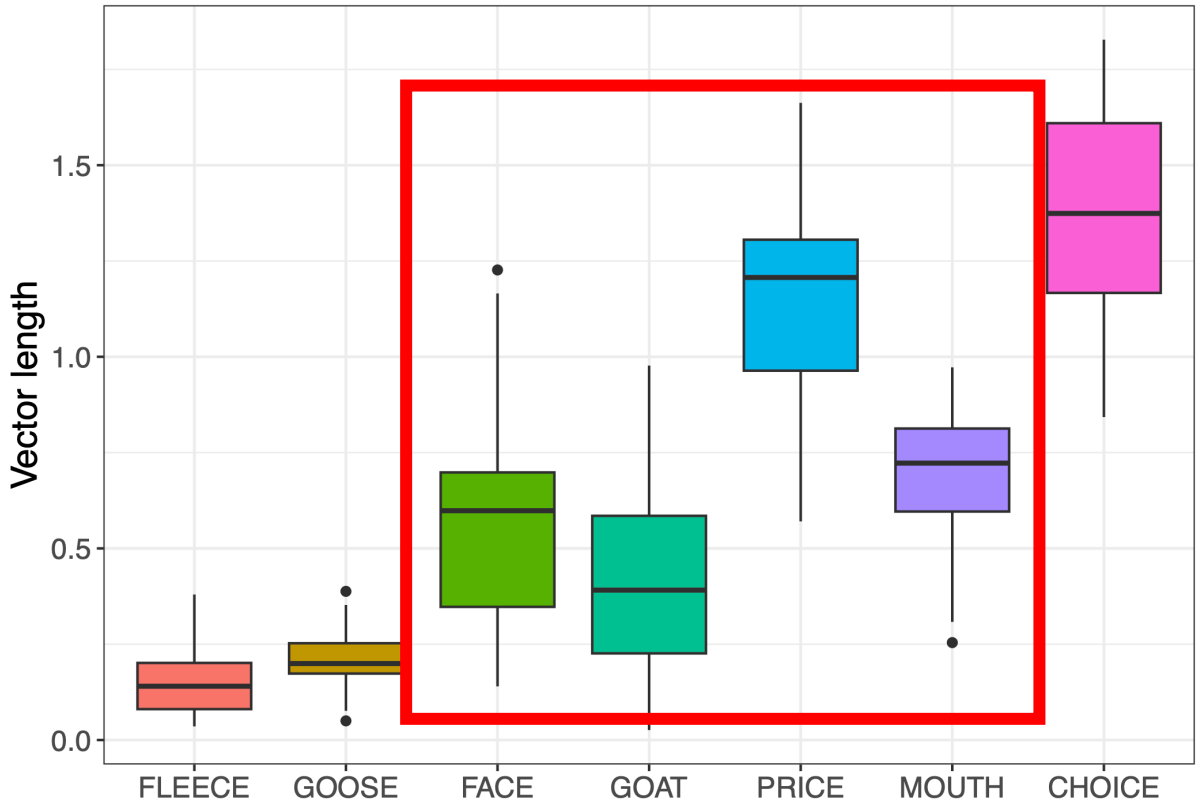
Results: Vector Length

- *broadly* follows monophthong-diphthong continuum
- dynamicity is *gradient*



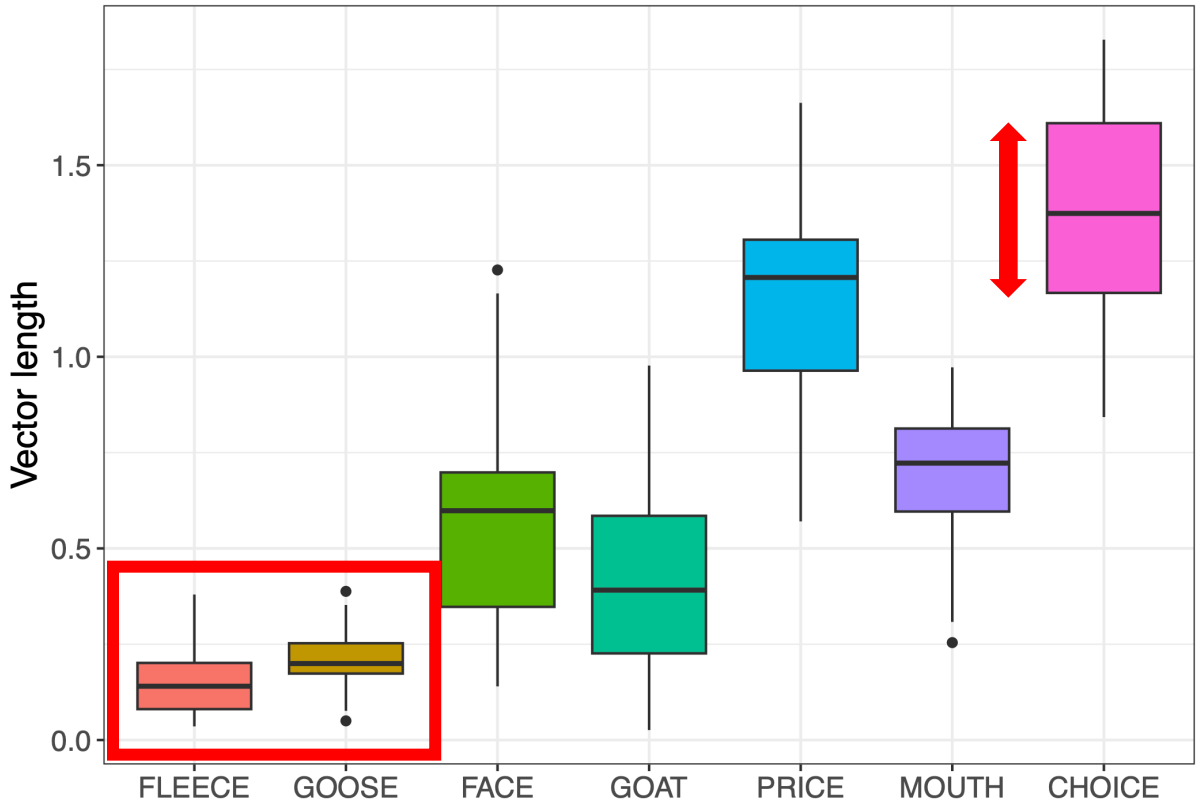
Results: Vector Length

- *broadly* follows monophthong-diphthong continuum
- dynamicity is *gradient*
  - dialectally-variable vowels have intermediate VL

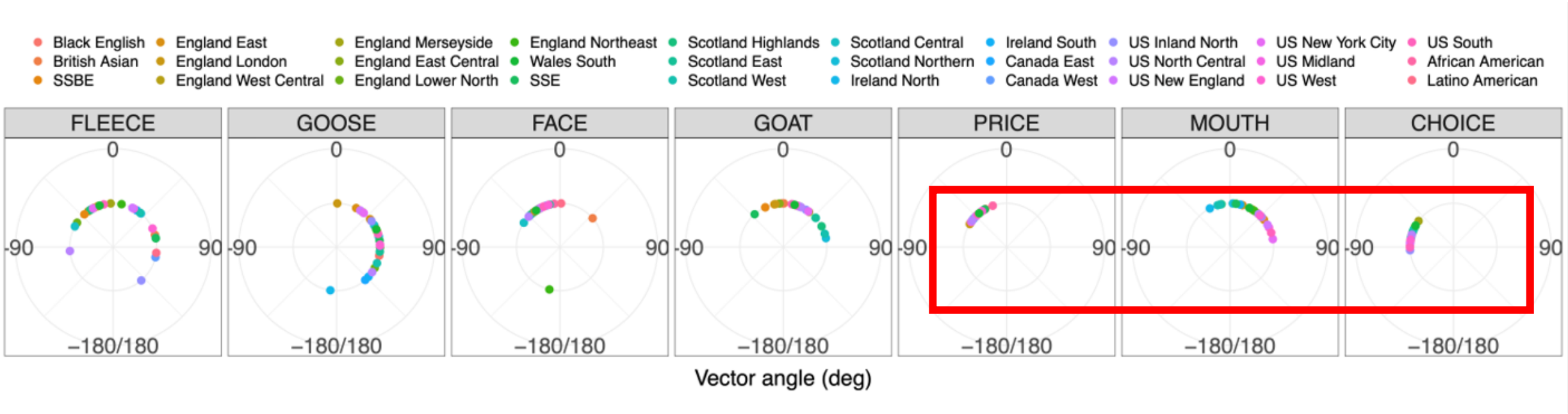


Results: Vector Length

- *broadly* follows monophthong-diphthong continuum
- dynamicity is *gradient*
  - dialectally-variable vowels have intermediate VL
- dialect variation in VL (least for monophthongs)

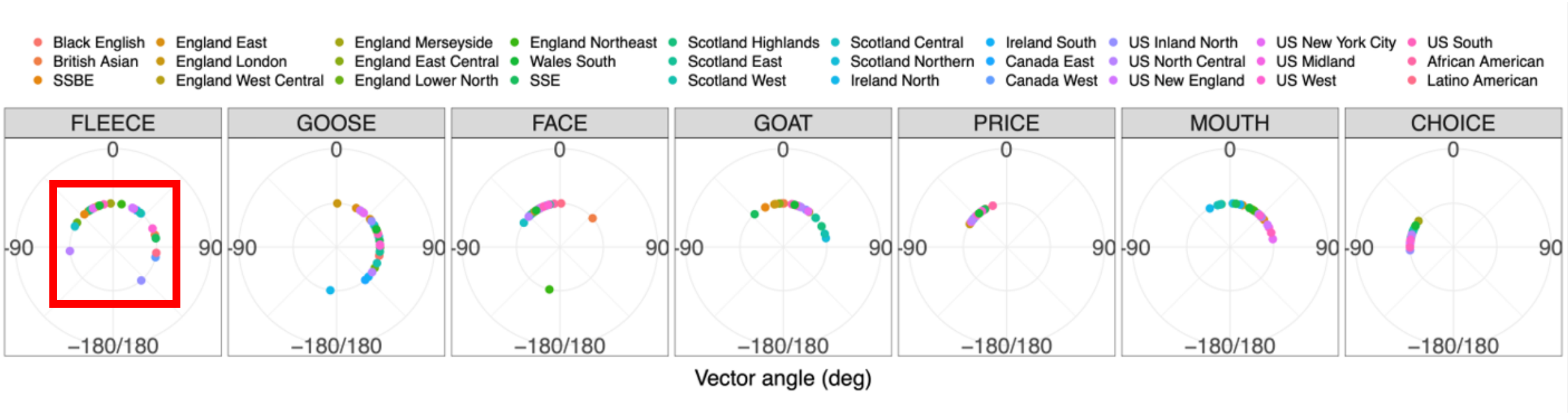


Results: Vector Angle



- *direction* of change less variable for more diphthongal vowels

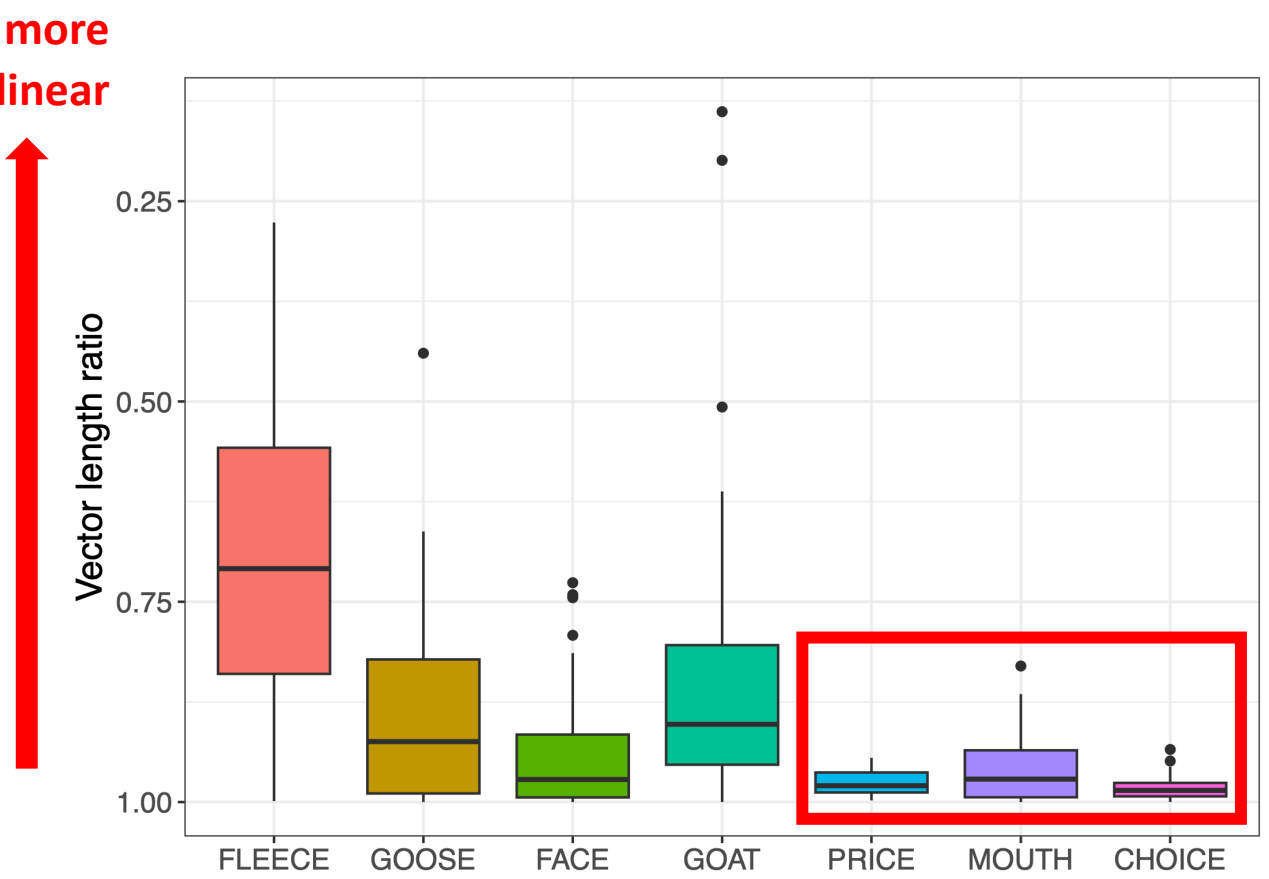
Results: Vector Angle



- *direction* of change less variable for more diphthongal vowels
- wide range of angles for FLEECE and GOOSE reflects less-defined change

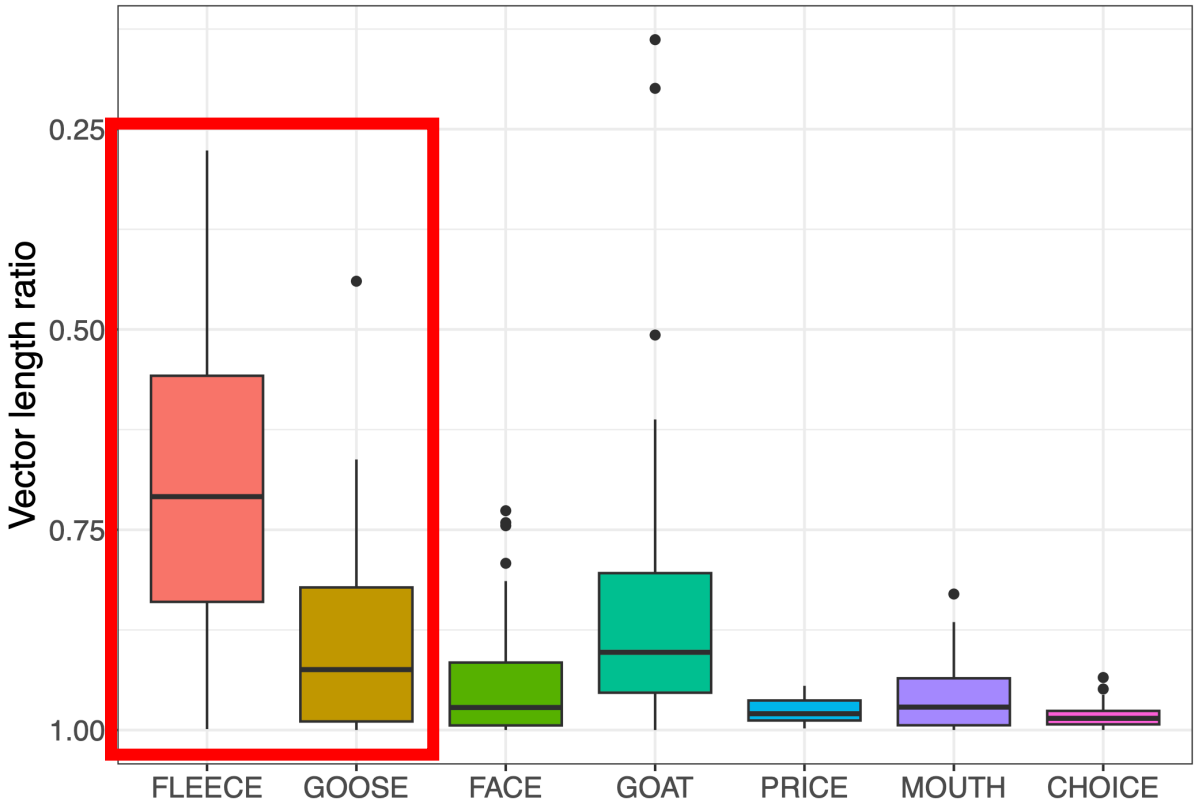
Results: Vector Length Ratio

- opposite direction from Vector Length
- vowels with more *linear* change exhibit less *non-linear* change



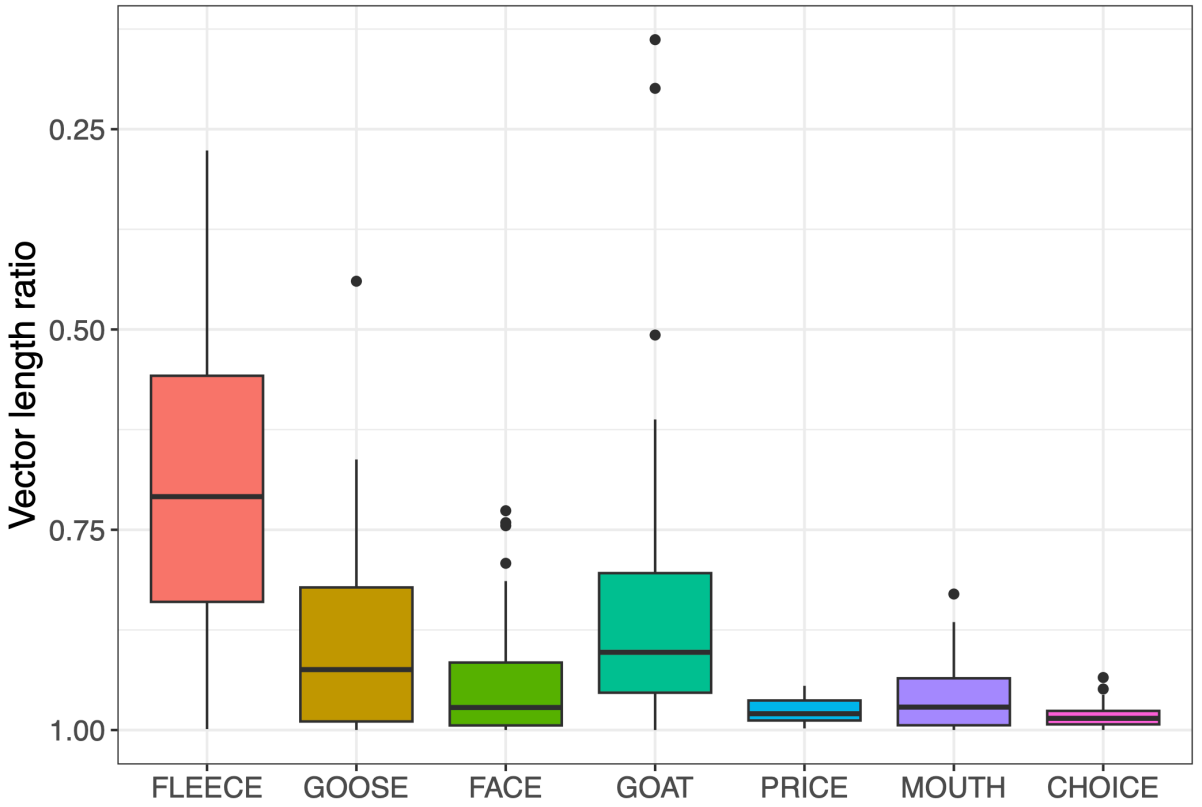
Results: Vector Length Ratio

- inverse of Vector Length
- vowels with more *linear* change exhibit less *non-linear* change
- FLEECE distinct from GOOSE
  - shows difference despite little linear change



Results: Vector Length Ratio

- inverse of Vector Length
- vowels with more *linear* change exhibit less *non-linear* change
- FLEECE distinct from GOOSE
- non-linear change is a *separate dimension* to linear formant change



Interim summary (RQ1)

- dynamic variation in vowels across dialects broadly reflects the monophthong-diphthong continuum
- dynamicity is gradient across both vowels and dialects
  - CHOICE exhibits most *linear* (VL) and least *non-linear* (VLR) variation
  - FLEECE exhibits most *non-linear* (VLR) and least *linear* (VL) variation
  - vowels can independently vary in linear and non-linear formant change

Watt (1999, 2002), Haddican et al. (2013), Fridland et al. (2014)

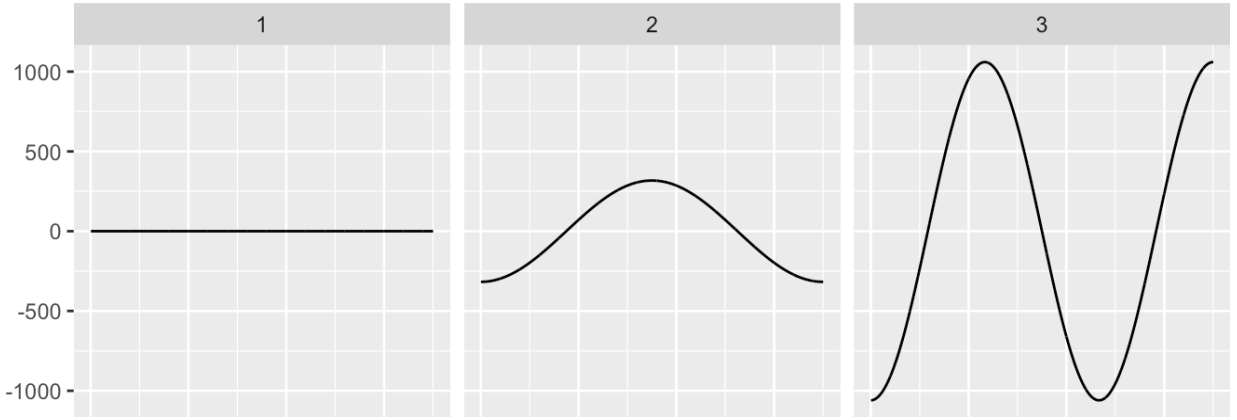
How do dialects vary in the dynamic properties of their vowels?

Dialect classification

- *Quantify* the role of dynamic information in distinguishing dialects beyond static F1/F2
- Can a model trained with static + dynamic information classify dialects better than one with just static?

Dialect classification: Measures

- Discrete Cosine Transform (DCT)
  - Represents a signal as set of cosine functions at different frequencies



- Higher DCT coefficients reflect more complex elements of the formant trajectory
  - DCT C_0 = mean formant value
  - DCT C_1 = amount & direction of linear change
  - DCT C_2 = amount of non-linearity in that change

Zahorian & Jagharghi (1993), Watson & Harrington (1999), Hillenbrand et al. (2001), Morrison (2013)

Dialect classification: Method

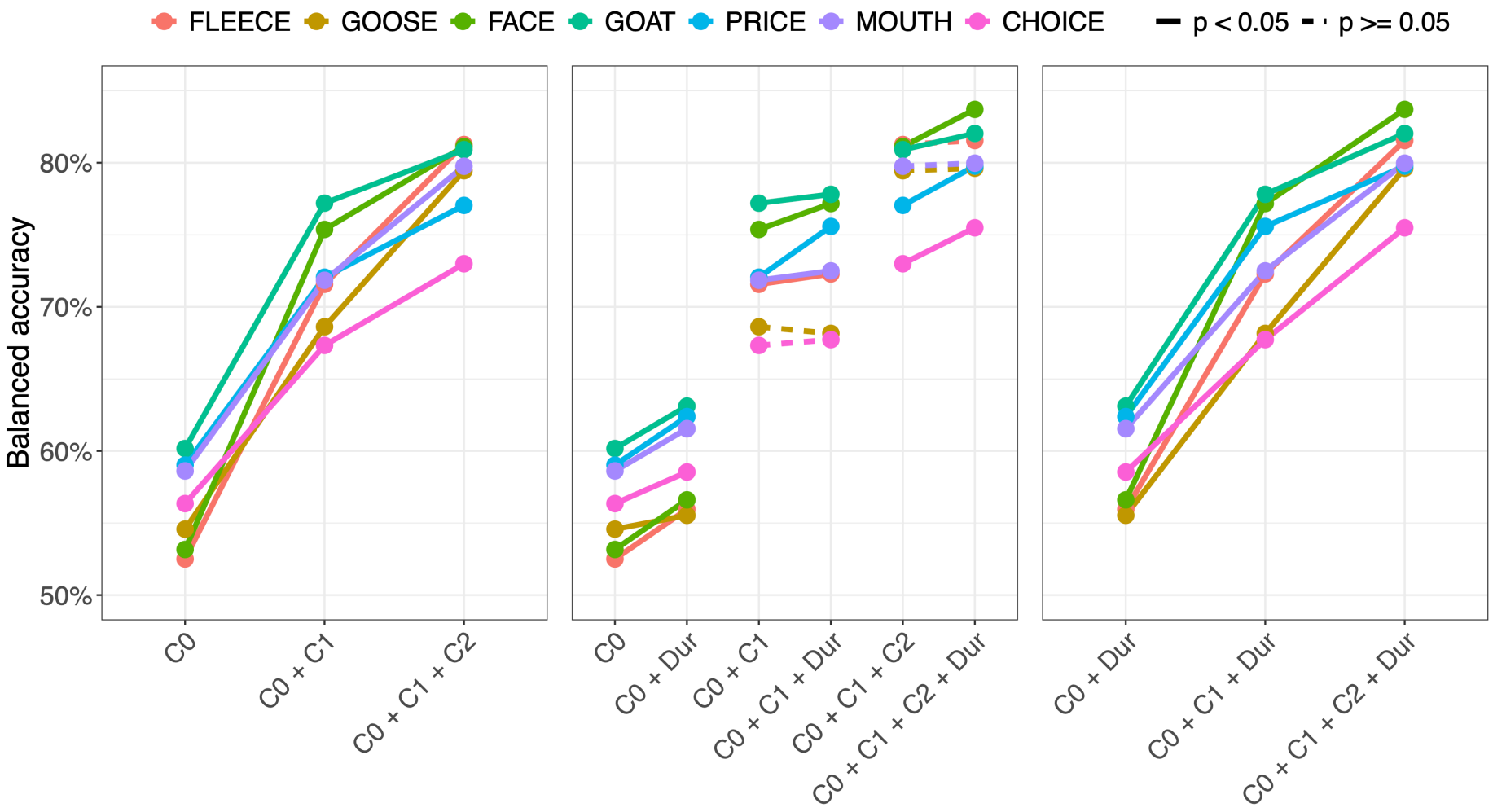
- Random Forests (RFs)
  - Supervised learning algorithm based on aggregation of decision trees
- train models to predict a speaker's dialect label (n=30) based on different combinations of measures:
  - C0 (GAMM-predicted F1 + F2)
  - C0 + Duration
  - C0 + C1
  - C0 + C1 + Duration
  - C0 + C1 + C2
  - C0 + C1 + C2 + Duration
- 6 models x 7 vowels = 42 models

Breiman (2001), Altmann et al. (2011), Sonderegger & Sóskuthy (2024)

Dialect classification: Method

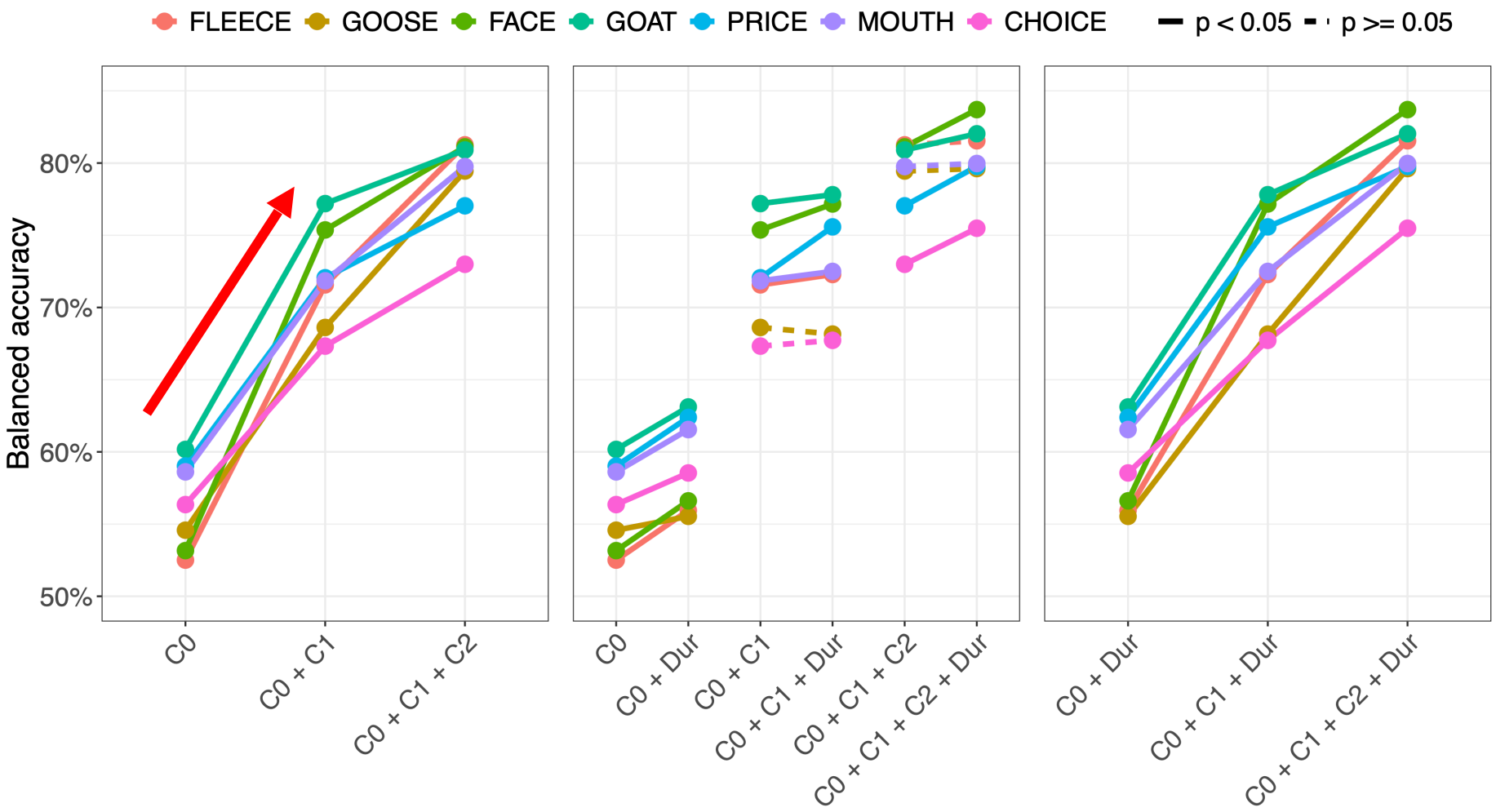
- train on 80% of speakers (n=3733) and predict dialect of remaining 20% (n = 933)
- Does adding dynamic information aid in distinguishing dialects?
  - What changes classification **accuracy** between models?
  - Which acoustic features are **most important** in distinguishing dialects?

Dialect classification: Accuracy



Dialect classification: Accuracy

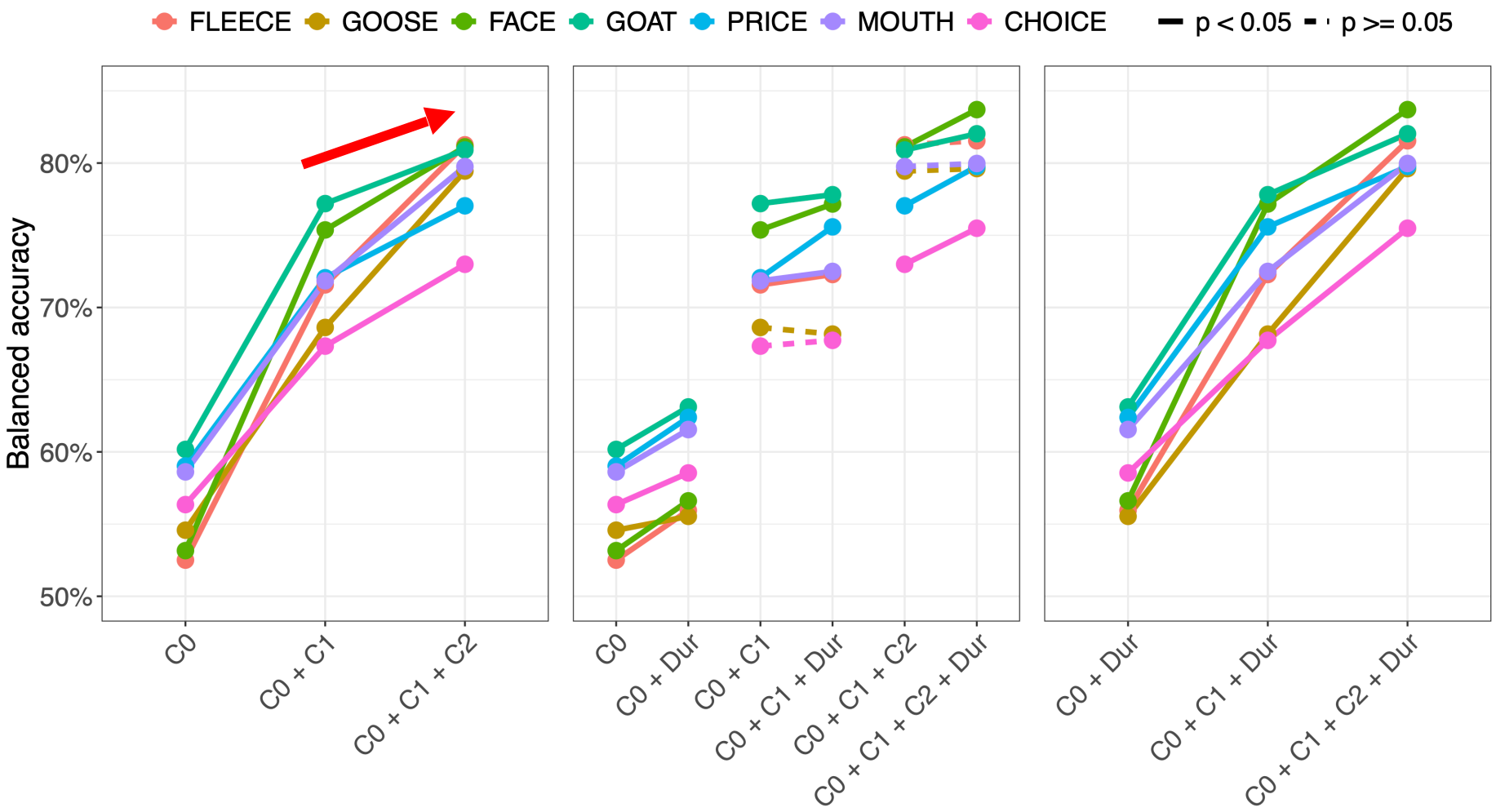
C_1 greatly improves accuracy



Dialect classification: Accuracy

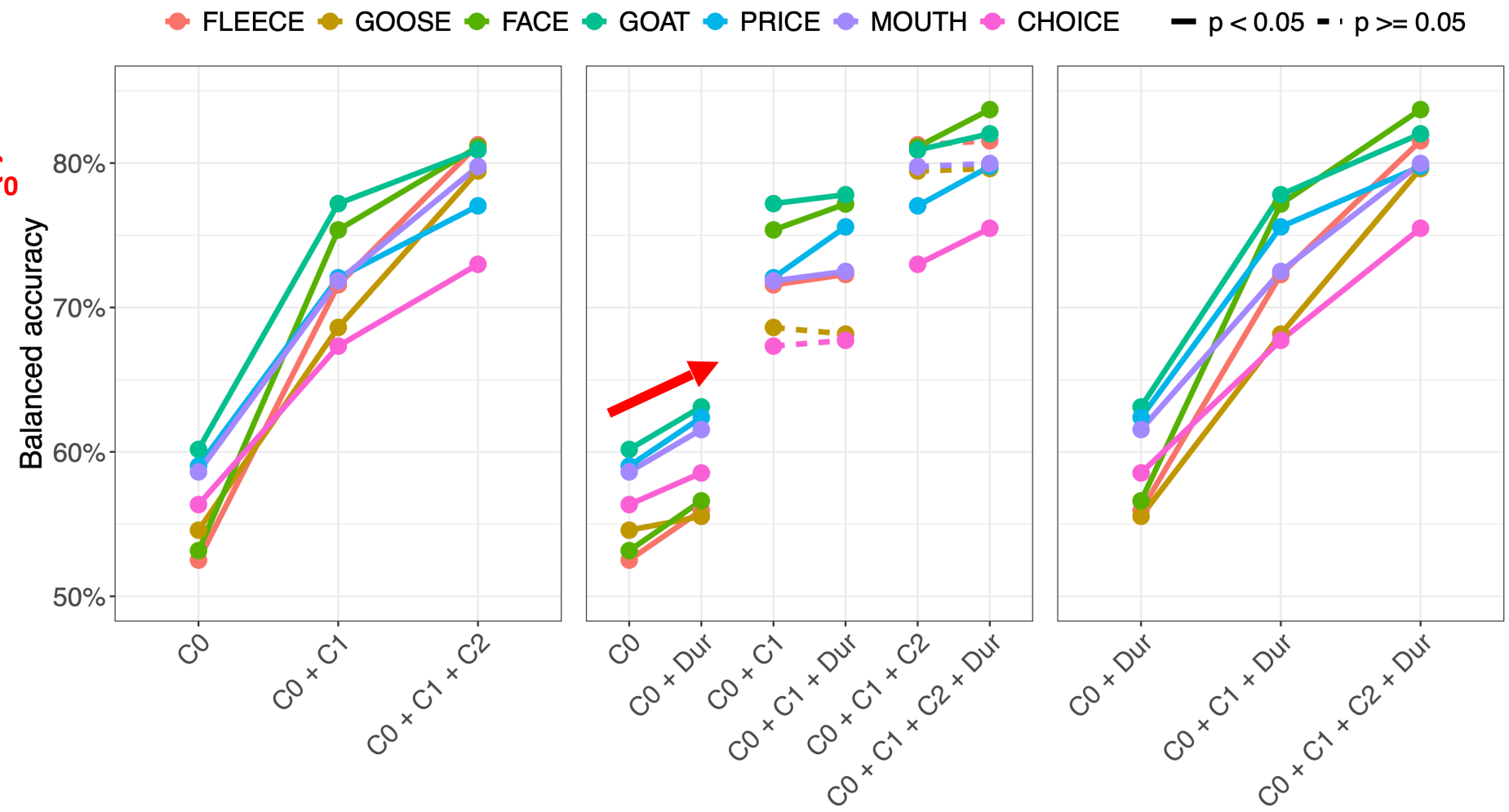
C_2 also improves accuracy

adds more for FLEECE and GOOSE



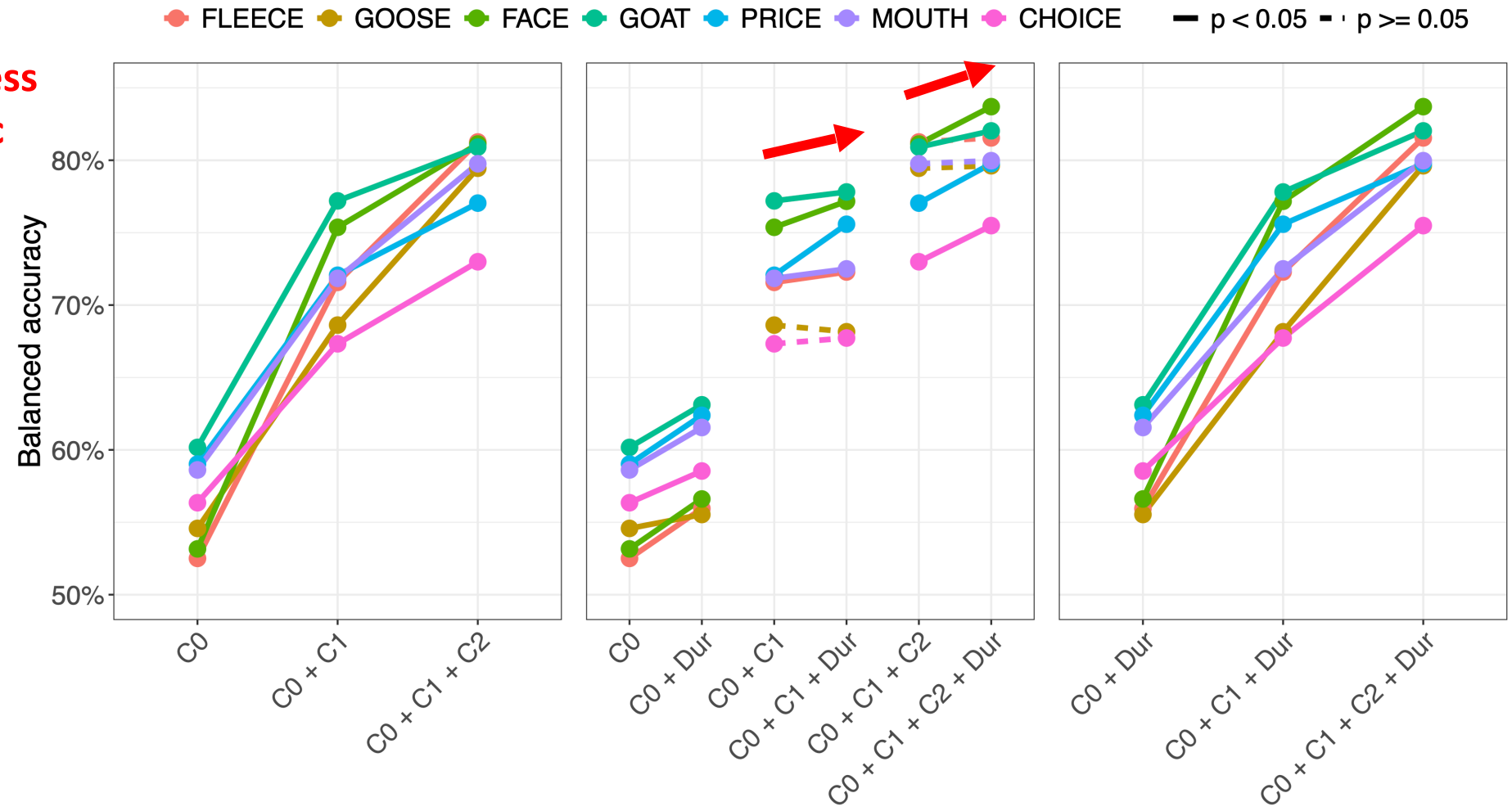
Dialect classification: Accuracy

including duration alongside C_0 improves accuracy



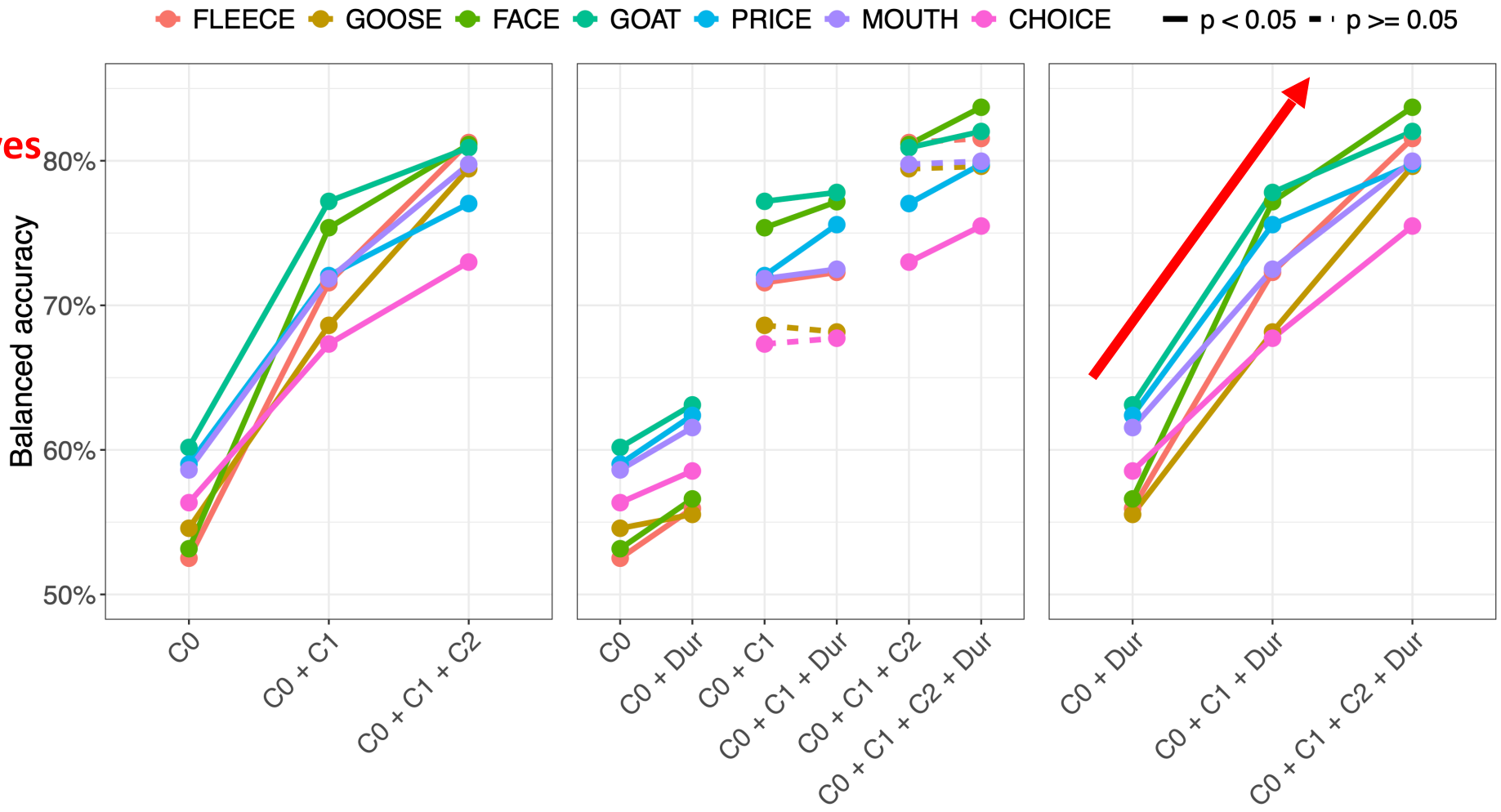
Dialect classification: Accuracy

duration contributes less once dynamic (C_1 & C_2) information included

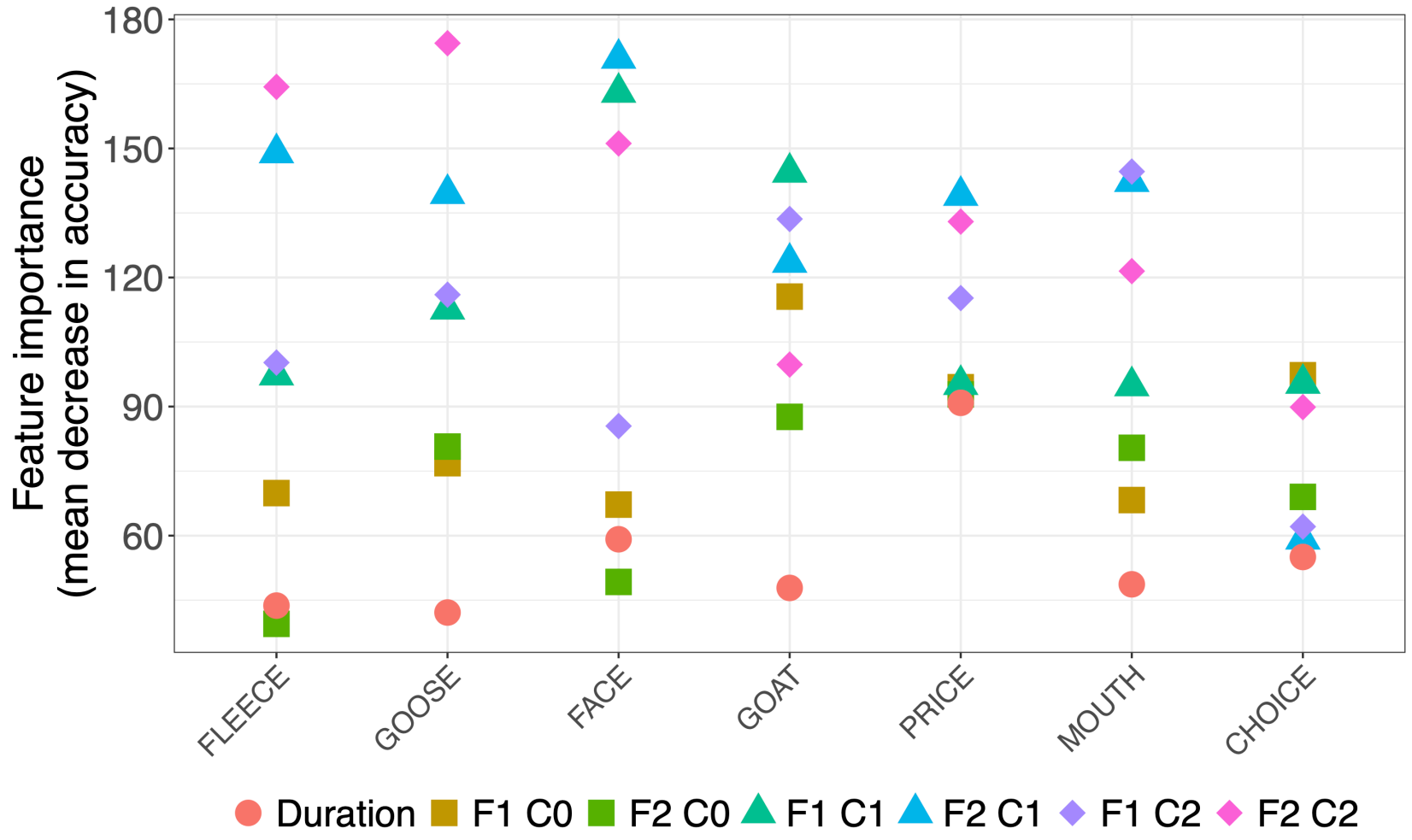


Dialect classification: Accuracy

dynamic information greatly improves accuracy – even with duration included

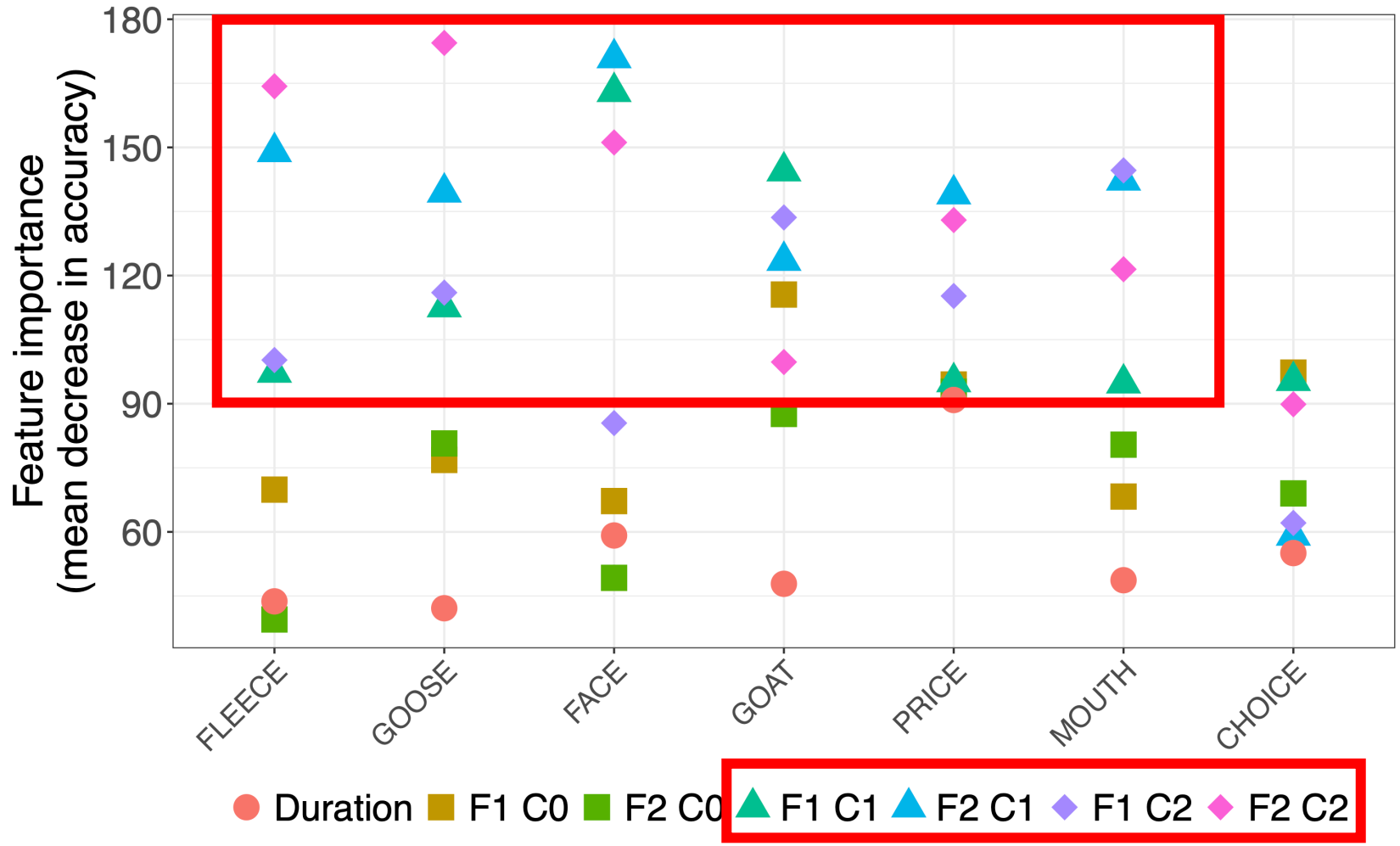


Dialect classification: Feature importance



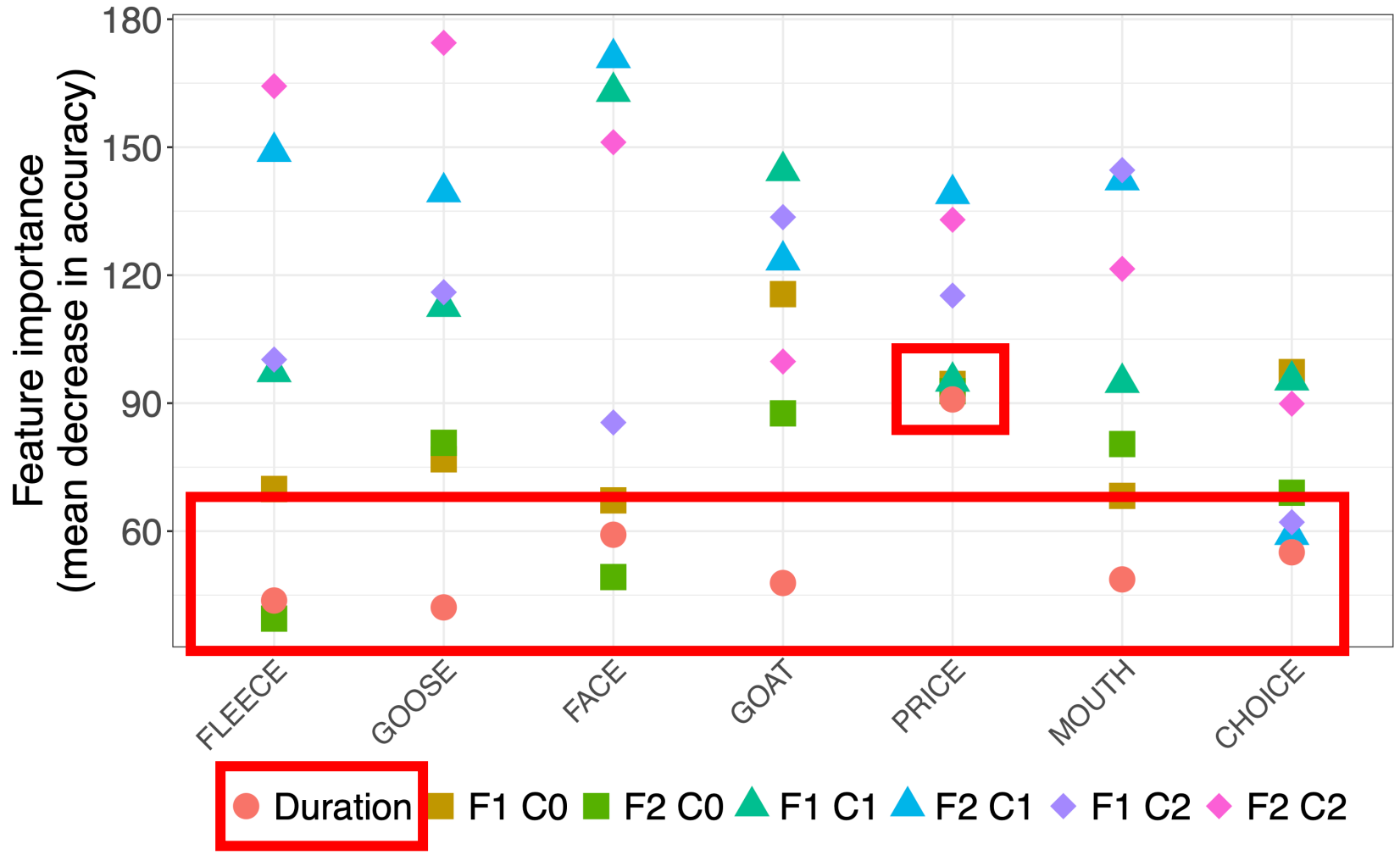
Dialect classification: Feature importance

dynamic information (C_1 & C_2) is most important for all vowels (except CHOICE)



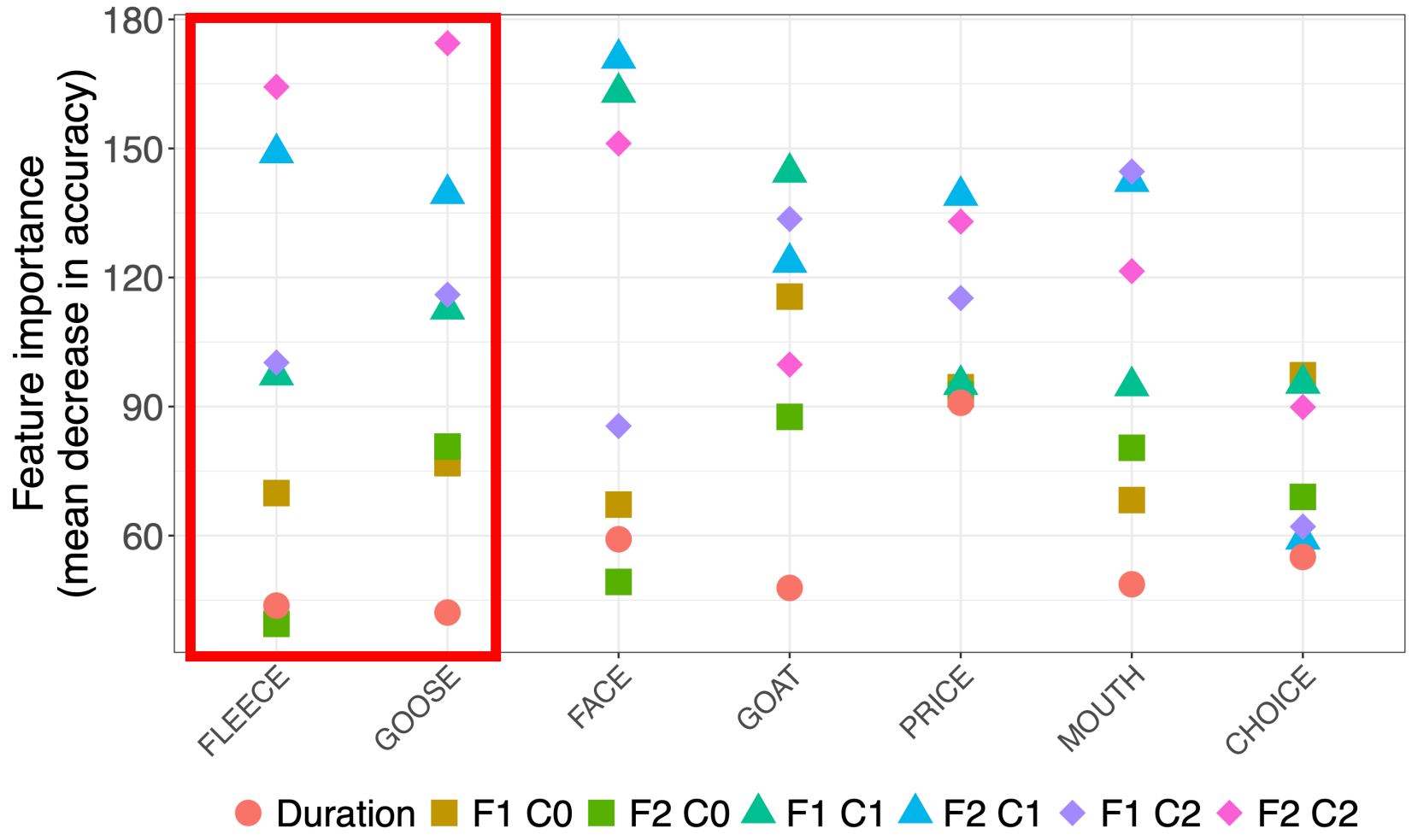
Dialect classification: Feature importance

duration is a less important measure



Dialect classification: Feature importance

FLEECE and GOOSE share similar feature importance

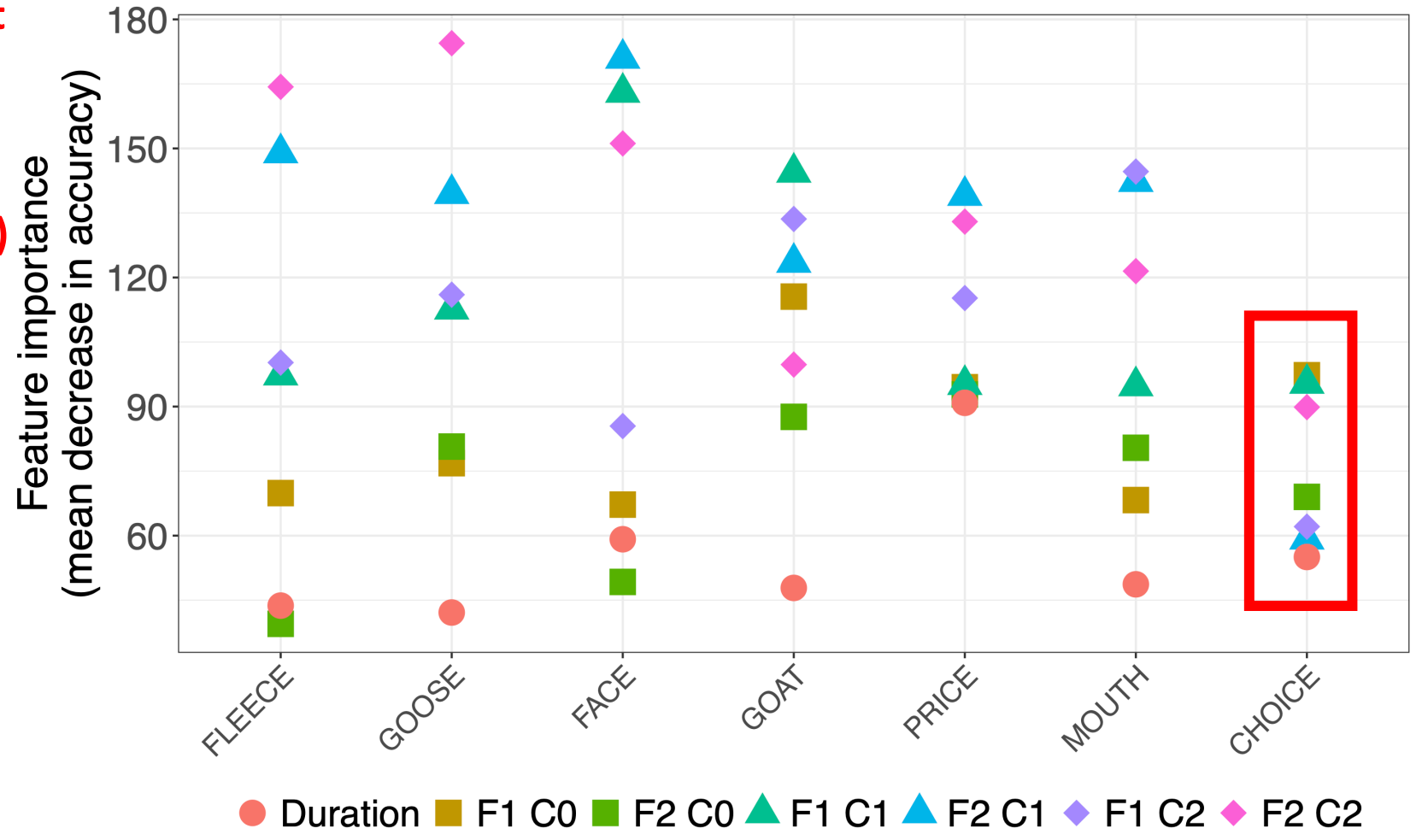


Dialect classification: Feature importance

CHOICE – lowest classification accuracy

no single (set of) measures most important

highest importance related to position & length of F1 trajectory



Summary (RQ1 + RQ2)

- dynamic properties of English vowels broadly reflects the monophthong-diphthong continuum
- dynamicity is gradient across both vowels and dialects
  - FLEECE most *non-linear* (VLR) and least *linear* (VL)
  - CHOICE most *linear* (VL) and least *non-linear* (VLR)
  - vowels can independently vary in linear and non-linear formant change
- dynamic properties of vowels very important in distinguishing dialects
  - both in the **size** and **curvature** of the dynamic trajectory
  - duration is also informative, but contributes less when dynamic properties are known

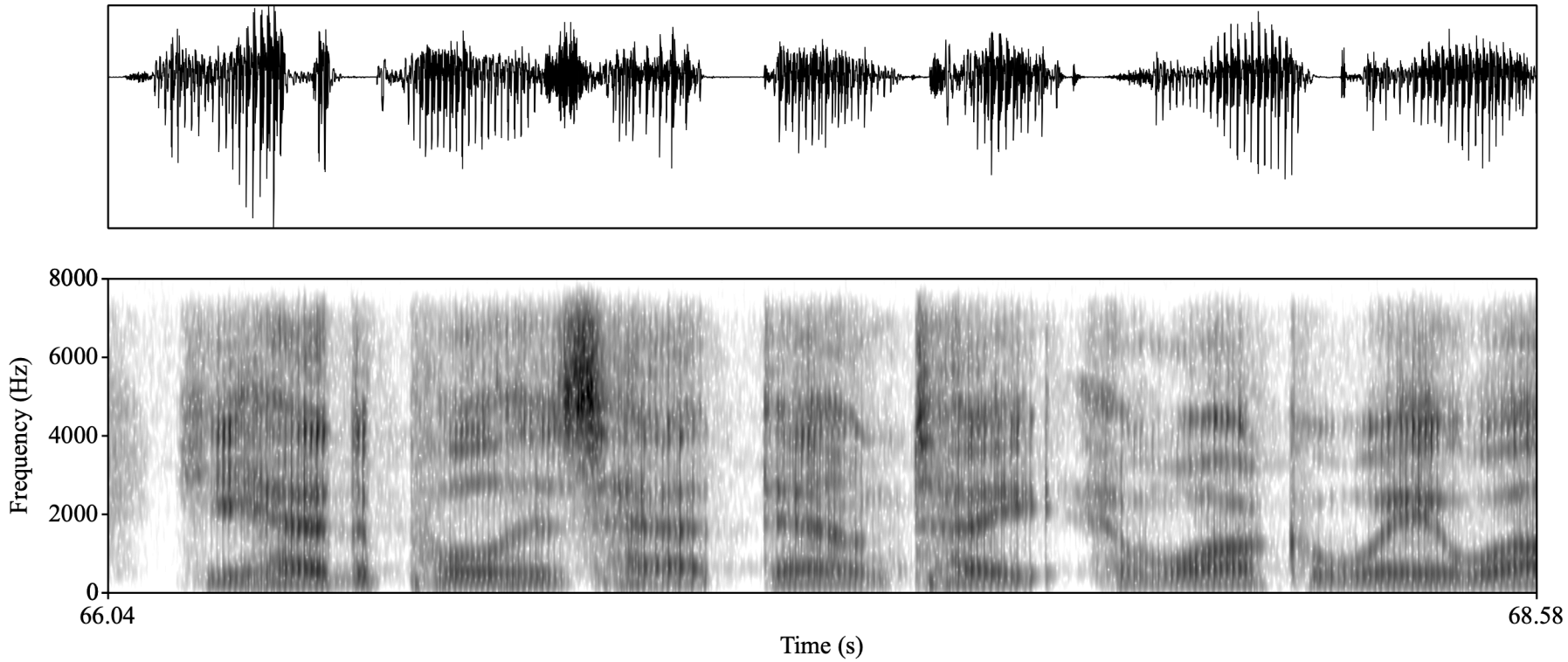
Watt (1999, 2002), Haddican et al. (2013), Fridland et al. (2014), Tanner (2023)

Discussion

1. How are dynamic properties of vowels structured across dialects?
   - looking across many dialects shows that these vowels are consistent in their dynamic properties, pointing to shared structures (Stanley et al 2021)
   - at same time, there is dialect variation
2. How do dialects vary in the dynamic properties of their vowels?
   - dynamic properties of vowels are very important in distinguishing between dialects (cf Williams et al 2014; Renwick and Stanley 2020)

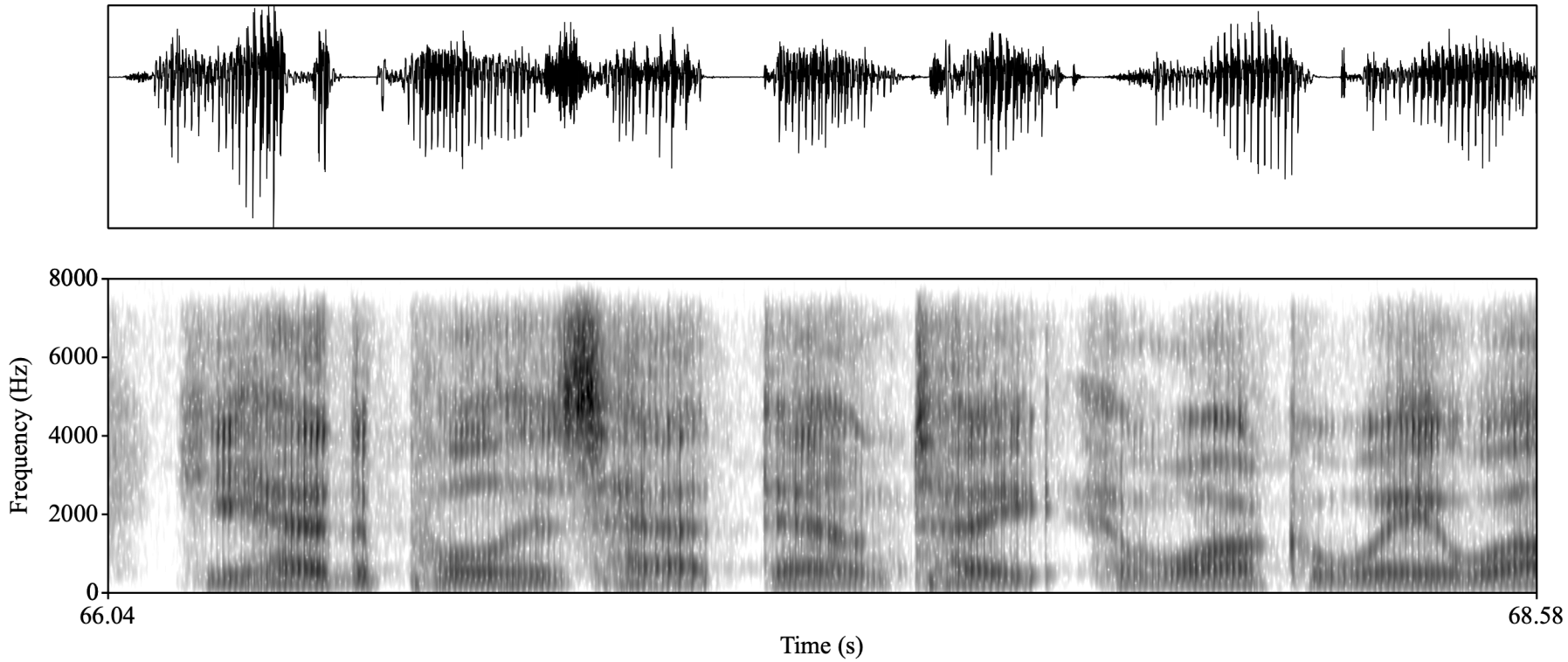
Discussion

- What kinds of information is encoded through vowel dynamics?
  - previous studies using DCT only needed C_1 (slope) to distinguish vowels within a single dialect, or across a few dialects
  - looking across many more dialects, we also find C_2 (curvature) is helpful
  - additional complexity (C_3) has been found necessary for individual speaker identification (e.g. Morrison 2013)
- How much do these low dimensional dynamic properties matter for listeners? (cf e.g. Nearey and Assmann 1986; Shaw et al 2024)
- Do speakers vary more in dynamicity than dialects do?



‘duration, likely along with spectral change over time, may be a part of a package of acoustic distinctions that signals both dialect and vowel category information’

Fridland et al (2014: 348)

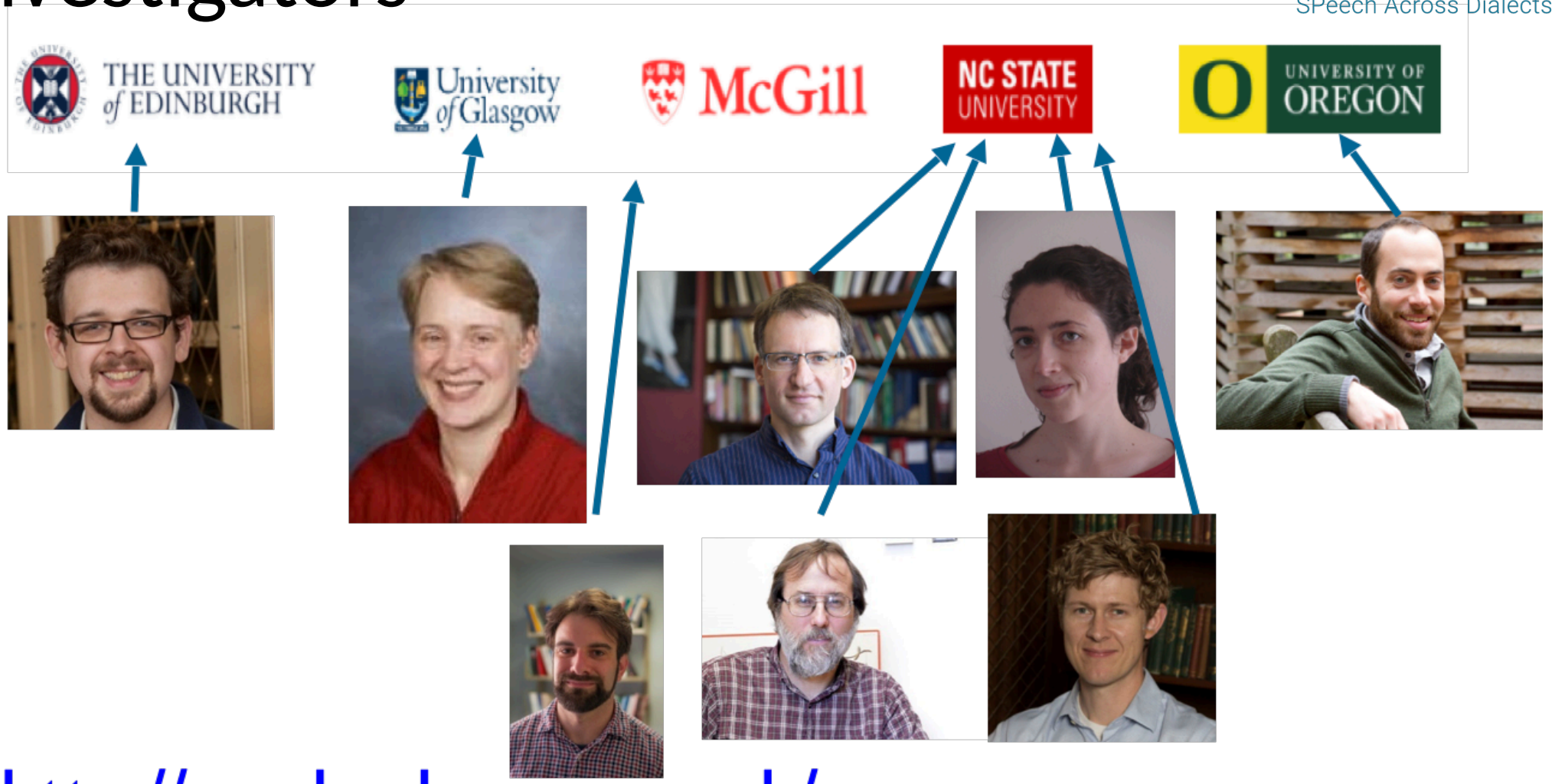


spectral change over time, along with duration, is part of a package of acoustic distinctions that signals both dialect and vowel category information'

Fridland et al (2014: 348)

Investigators

SPADE
SPeech Across Dialects of English



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

SPADE

SPeech Across Dialects of English

James Tanner
Postdoc

Rachel Macdonald
Project Manager

Michael McAuliffe
Software Development

The British Academy

Trans-Atlantic Platform
Social Sciences and Humanities

DIGGING INTO DATA CHALLENGE

Calcul Québec

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

SPADE
SPeech Across Dialects of English

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

Arlie Coles (U. de Montréal)

Elias Stengel-Eskin (Johns Hopkins)

Michael Goodale (ENS), Sarah Mihuc (McGill)

Jordan Holley
Peter Andrews
Kaylynn Gunter

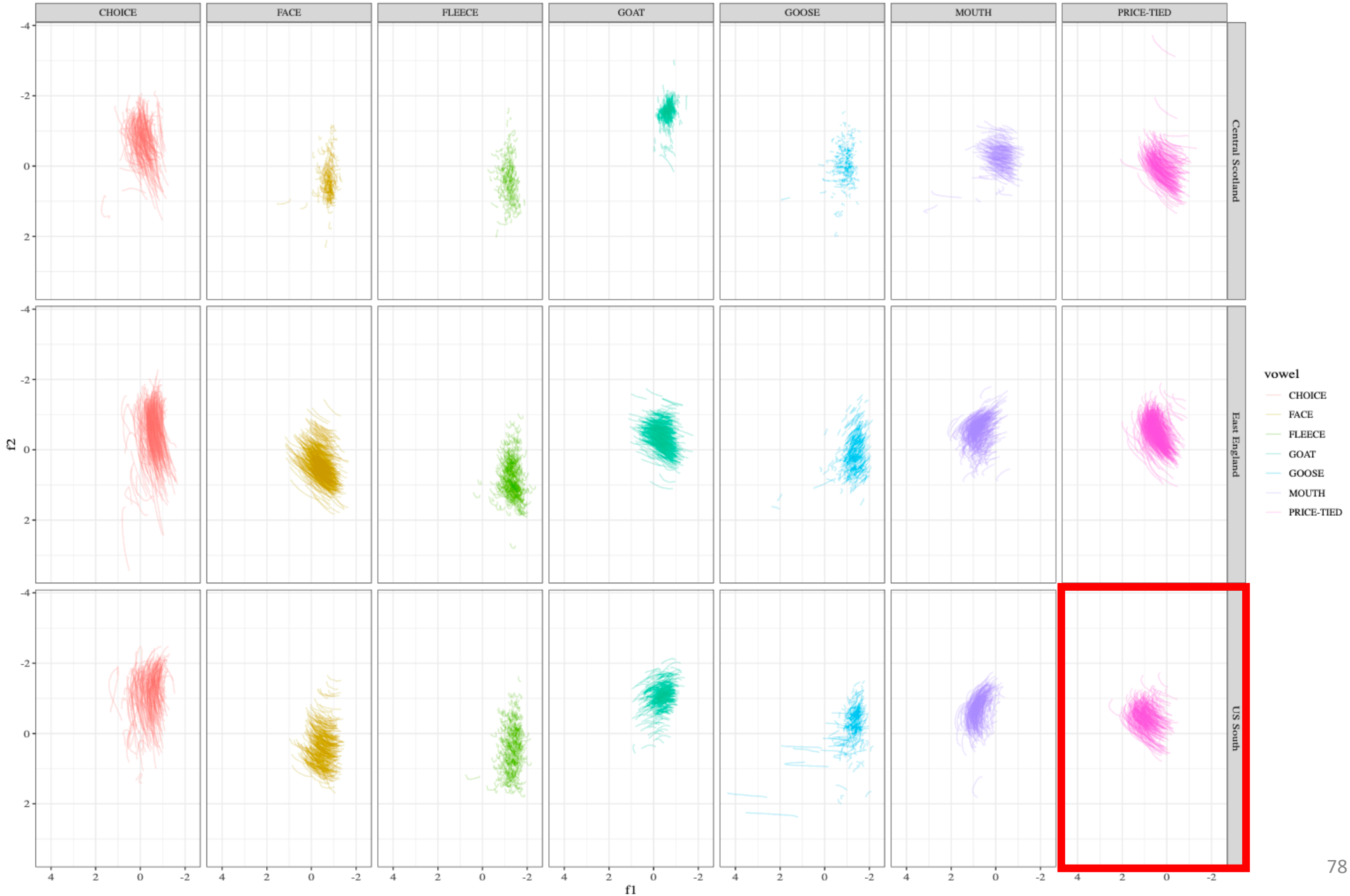
Vanna Willerton (McGIll)


Thank you!


Extra slides

Vowel category decisions

- Vowel labels from UNISYN (Fitt 2001): cross-dialect pronunciation lexicon
- Some categories merged for 'broad' comparisons
  - FLEECE: single category
  - GOOSE: GOOSE + GHOUL (pre-/l/) + BREWED (SVLR)
  - FACE: FACE + WAIST (monophthong in Wales English)
  - GOAT: GOAT + KNOW (diphthong in Wales English)
  - PRICE: TIED + PRICE (raising)
  - MOUTH: LOUD + MOUTH (raising)
  - CHOICE: single category


CHOICE
FACE
FLEECE
GOAT
GOOSE
MOUTH
PRICE-TIED
Central Scotland
East England
US South
f2
f1
vowel
CHOICE
FACE
FLEECE
GOAT
GOOSE
MOUTH
PRICE-TIED

region	speakers	tokens
African_American	267	217206
Black_English	17	1680
British_Asian	18	1476
Canada_East	9	1765
Canada_West	305	125594
England_East	353	105289
England_East_Central	101	20334
England_London	96	14681
England_Lower_North	266	62432
England_Merseyside	62	12702
England_Northeast	156	7672
England_West_Central	129	23131
Ireland_North	94	11218
Ireland_South	167	12702
Latino_American	91	32331
Scotland_Central	32	11898
Scotland_East	29	9286
Scotland_Highlands	7	2943
Scotland_Northern	35	6834
Scotland_West	247	116417
Standard_Scottish_English	311	18072
Standard_Southern_British_English	169	54015
US_Inland_North	120	20018
US_Midland	280	117535
US_New_England	471	34842
US_New_York_City	72	3047
US_North_Central	98	852
US_South	368	140585
US_West	220	110782
Wales_South	76	15998

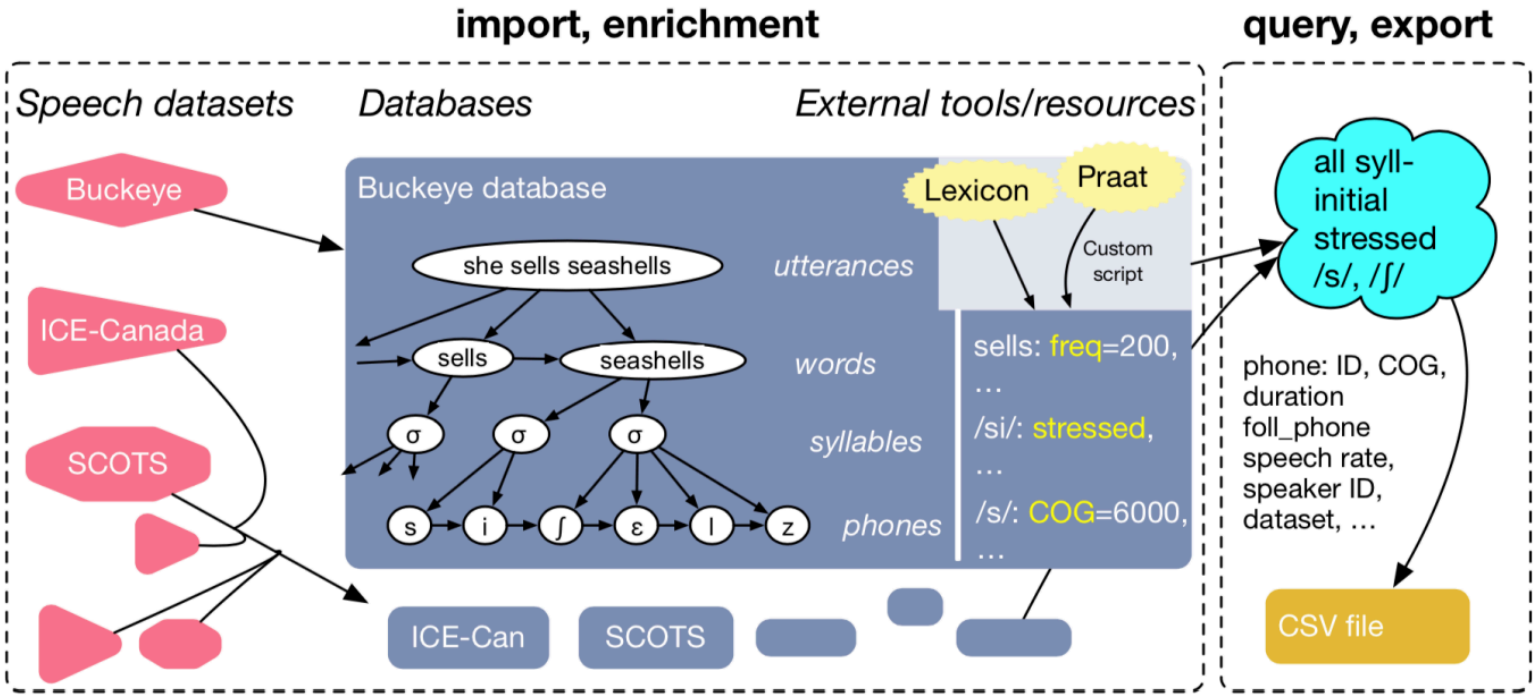
region	CHOICE	FACE	FLEECE	GOAT	GOOSE	MOUTH	PRICE.TIED
African_American	1428	37930	38829	30543	26050	19900	62526
Black_English	12	225	352	234	210	150	497
British_Asian	NA	155	317	228	234	138	404
Canada_East	14	398	437	257	184	157	318
Canada_West	608	20909	23960	17934	12023	10383	39777
England_East	2039	21017	23306	14009	8705	10099	26114
England_East_Central	218	3330	3748	3040	1746	1934	6318
England_London	202	2197	2719	1895	1380	1242	5046
England_Lower_North	1130	9191	11036	12605	5630	6573	16267
England_Merseyside	55	1980	2279	2432	1128	1422	3406
England_Northeast	57	1345	1685	1477	677	697	1734
England_West_Central	149	3983	4299	4525	2222	2494	5459
Ireland_North	121	2321	2650	1708	917	897	2604
Ireland_South	47	2512	2741	1752	975	1067	3608
Latino_American	174	6762	3543	2831	1871	1644	15506
Scotland_Central	463	2958	2415	1689	946	904	2523
Scotland_East	134	2099	2011	1421	882	567	2172
Scotland_Highlands	34	651	623	441	331	164	699
Scotland_Northern	145	1261	1522	1334	764	305	1503
Scotland_West	1453	22501	24390	17123	10966	9664	30320
Standard_Scottish_English	273	3436	3204	2644	3225	1475	3815
Standard_Southern_British_English	659	7598	8927	9479	4471	6592	16289
US_Inland_North	436	2670	3575	3002	1874	2426	6035
US_Midland	1046	20175	23484	18412	10528	9329	34561
US_New_England	860	2694	6298	4753	3002	4935	12300
US_New_York_City	77	363	590	429	289	368	931
US_North_Central	40	147	313	137	78	34	103
US_South	1014	25426	27885	18610	11079	14815	41756
US_West	745	27023	9804	7308	5239	4906	55757
Wales_South	309	3228	3330	2579	1736	1418	3398

Vowels for this study

	SHORT		LONG					
			Upgliding				Ingliding	
			Front upgliding		Back upgliding			
	V		Vy		Vw		Vh	
nucleus	front	back	front	back	front	back	unrounded	rounded
high	KIT	FOOT	FLEECE			GOOSE		
mid	DRESS	STRUT	FACE	CHOICE		GOAT	NURSE	THOUGHT
low	TRAP			PRICE	MOUTH		PALM, LOT	

Table 2.4. Labov et al. (2006)

Data Processing



https://github.com/MontrealCorpusTools/PolyglotDB

Data

- Each dialect: X tokens from Y speakers
- How do we represent each dialect's 'average' formant trajectory (for each vowel)?

Data

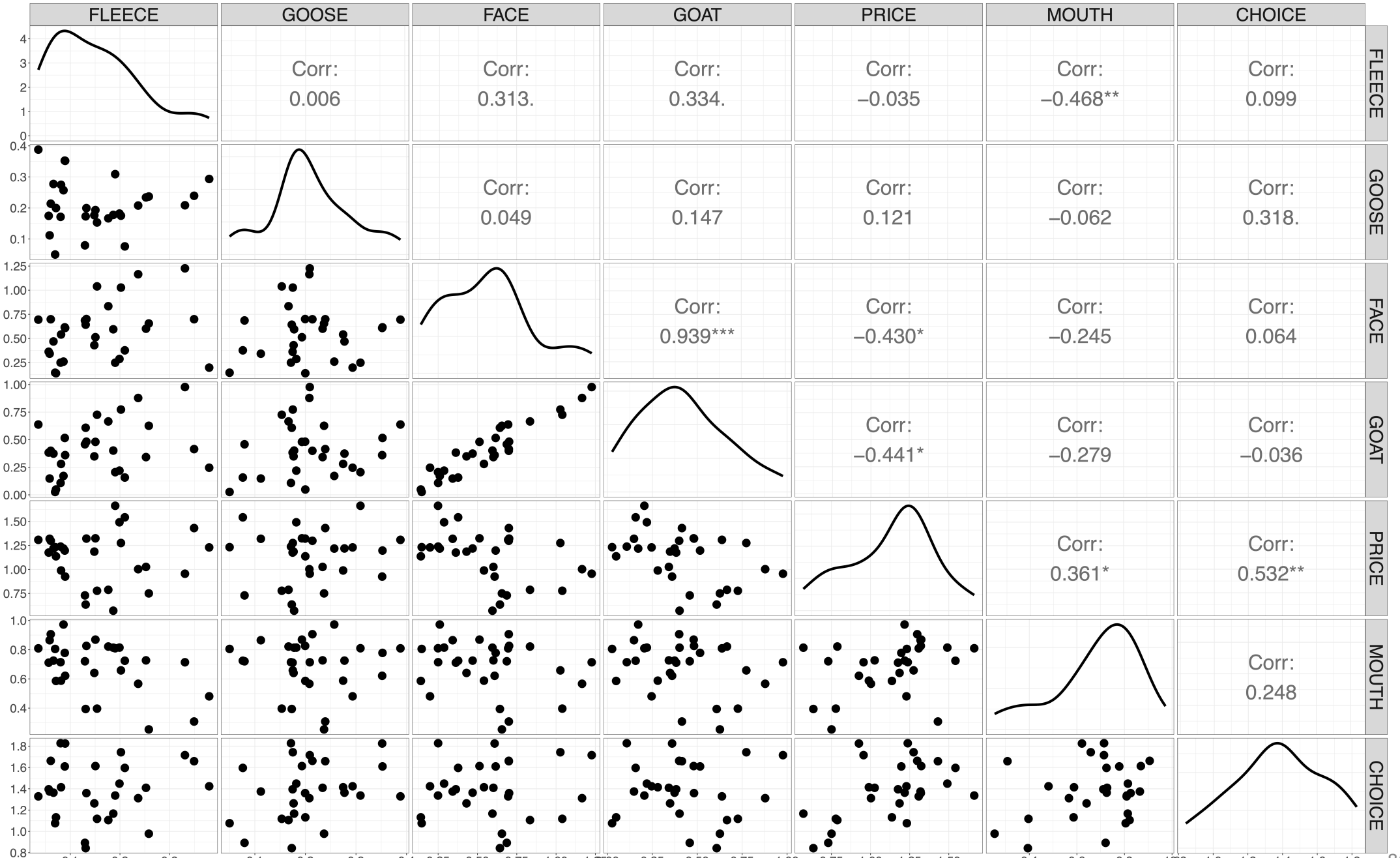
- Each dialect: X tokens from Y speakers
- How do we represent each dialect's 'average' formant trajectory (for each vowel)?
  - One option: average over all (relevant) tokens
  - Another option: average over each speaker's average ('mean of means')
  - …

Ferragne & Pellegrino (2010a, 2010b), Brand et al. (2021)

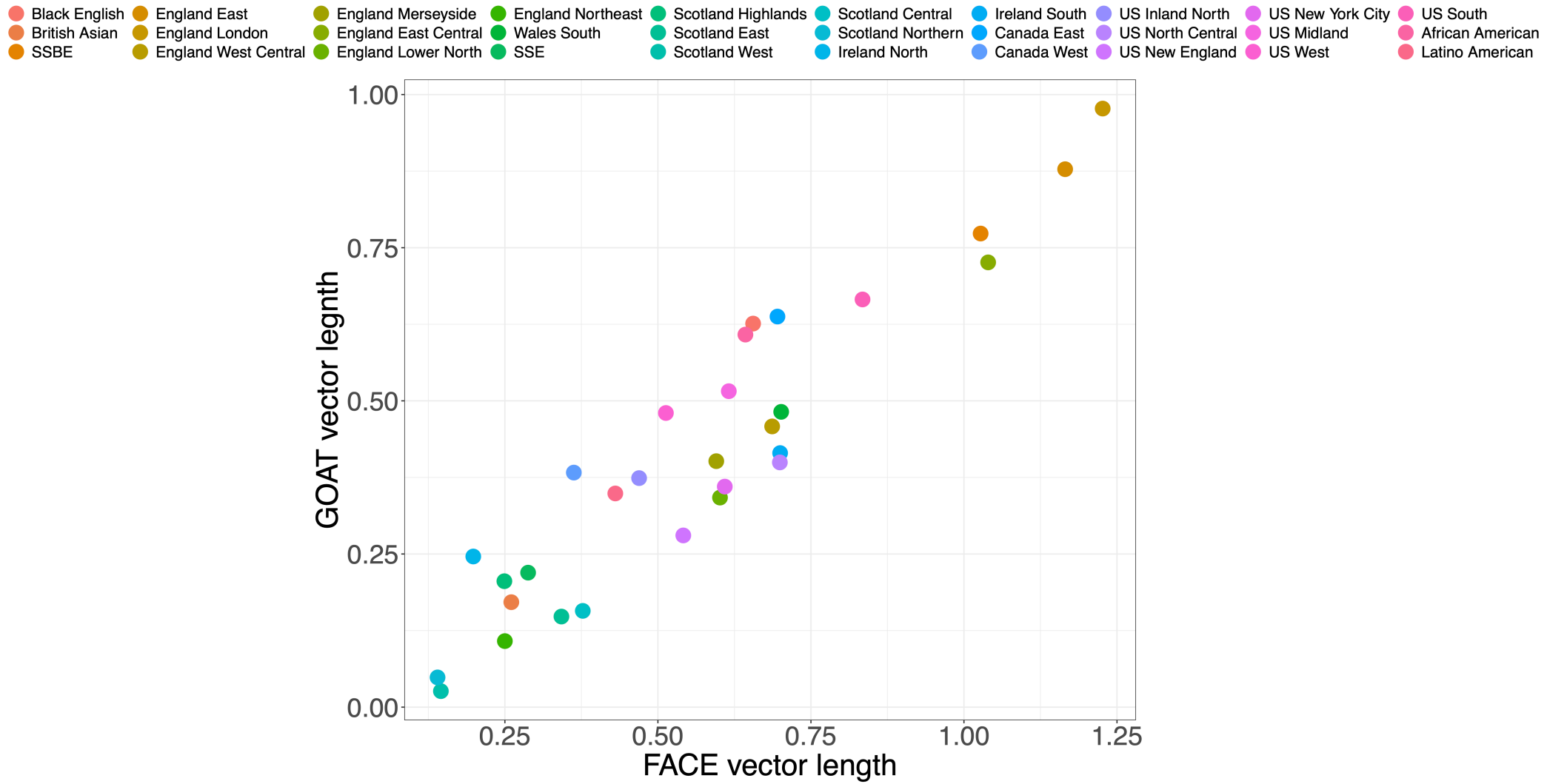
Data

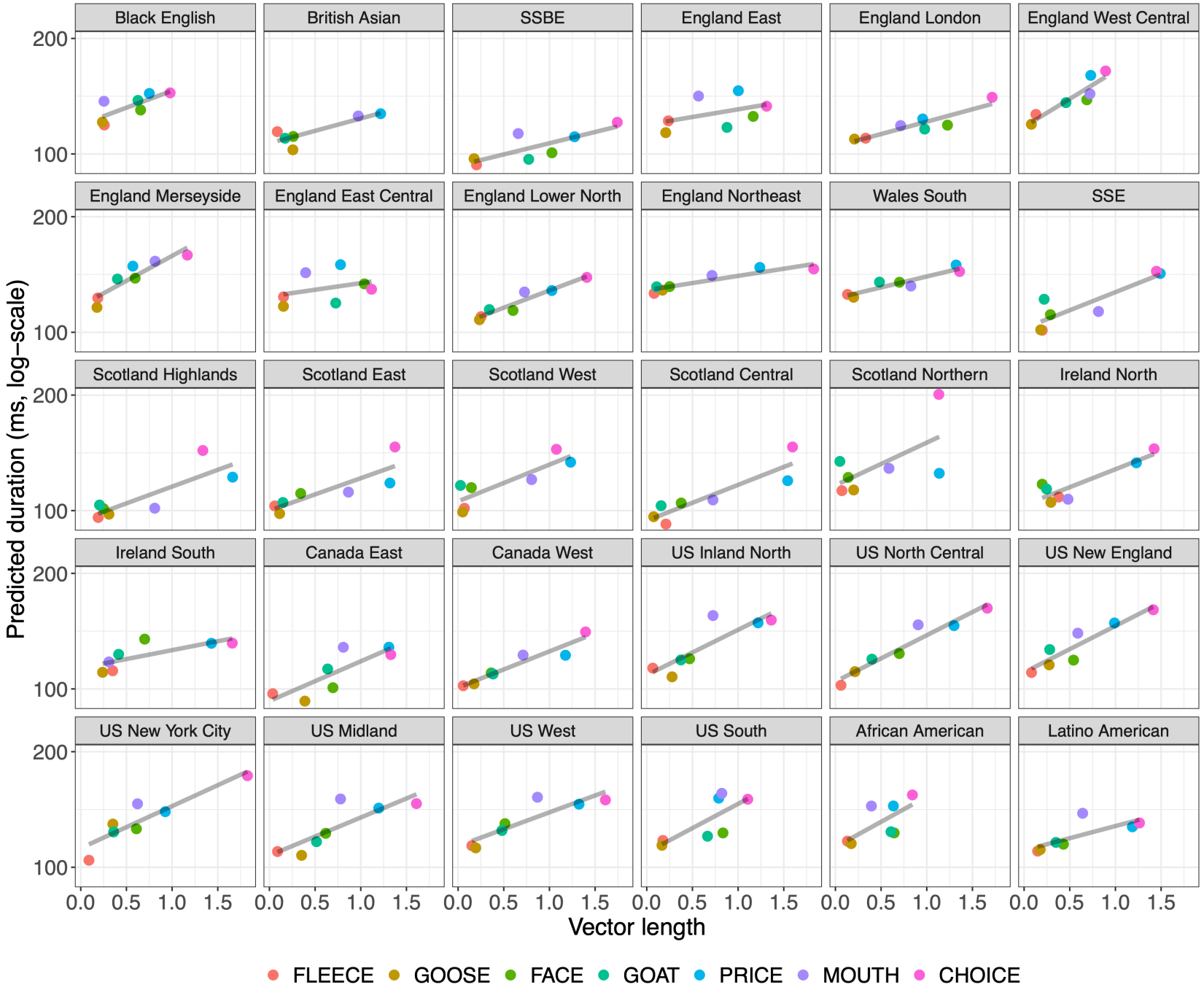
- Each dialect: X tokens from Y speakers
- How do we represent each dialect's 'average' formant trajectory (for each vowel)?
  - One option: average over all (relevant) tokens
  - Another option: average over each speaker's average ('mean of means')
  - …
- Here: directly *model* [F1, F2] trajectories

Ferragne & Pellegrino (2010a, 2010b), Brand et al. (2021)


FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE
Corr: 0.006
Corr: 0.313.
Corr: 0.334.
Corr: −0.035
Corr: −0.468**
Corr: 0.099
Corr: 0.049
Corr: 0.147
Corr: 0.121
Corr: −0.062
Corr: 0.318.
Corr: 0.939***
Corr: −0.430*
Corr: −0.245
Corr: 0.064
Corr: −0.441*
Corr: −0.279
Corr: −0.036
Corr: 0.361*
Corr: 0.532**
Corr: 0.248

Results: Relationships between measures




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
Predicted duration (ms, log−scale)
200
100
0.0
0.5
1.0
1.5
Vector length
FLEECE
GOOSE
FACE
GOAT
PRICE
MOUTH
CHOICE

This study

- explores how English vowels vary in different acoustic *dimensions*
  - position in F1 x F2 space
  - formant dynamics
  - duration
- looks across *many dialects*
  - allows us to draw 'big picture' observations across different versions of the 'same' language
  - made possible through data sharing and advances in data processing tools

Dialect classification: DCTs

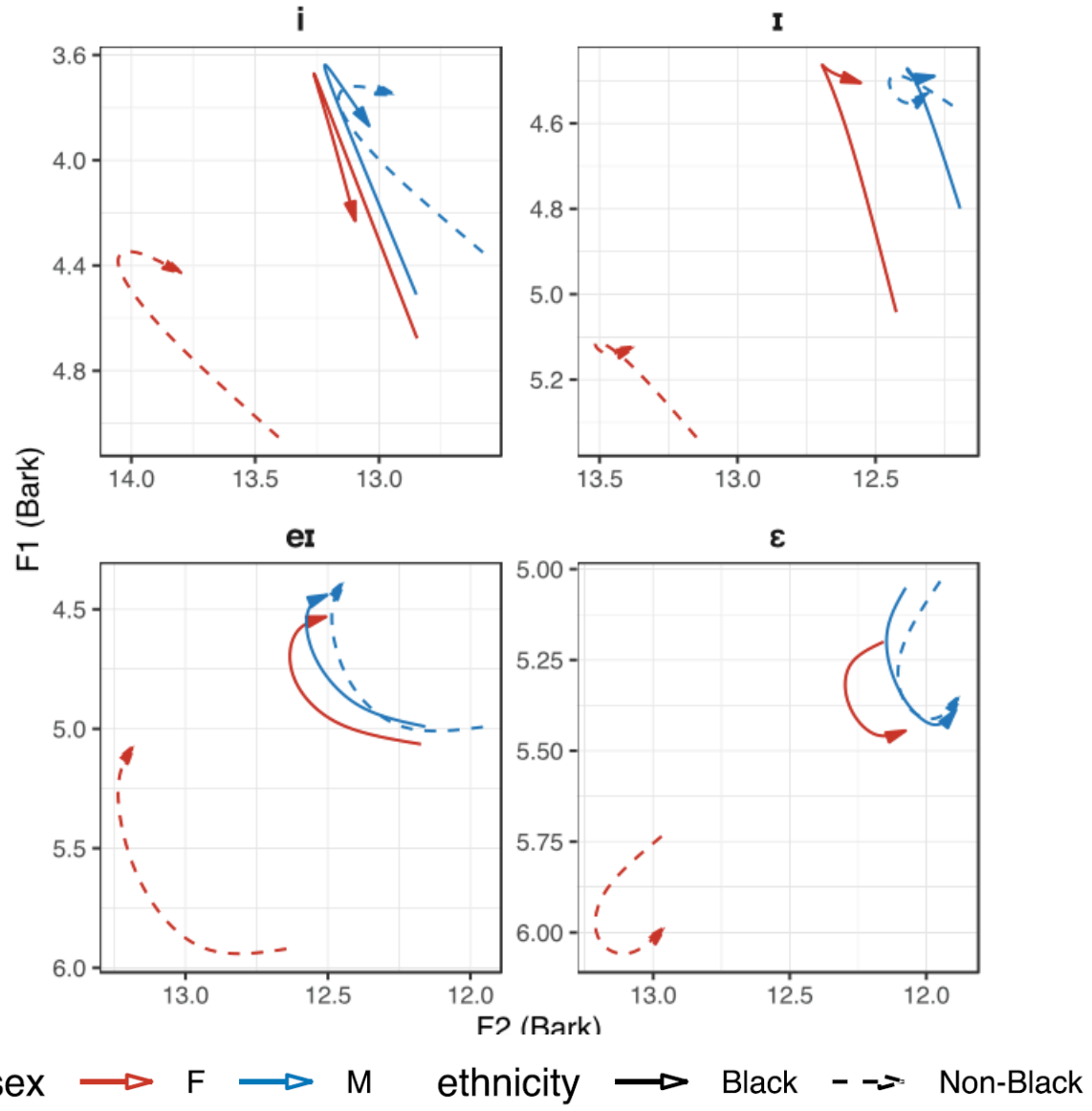
- Previously used in range of studies examining dynamic properties of vowels
- C_0 and C_1 shown to be informative for classifying different vowels within a dialect
- C_2 less informative/marginal increase in performance

Watson & Harrington (1999), Williams & Escudero (2014), Williams et al. (2019)

Vowels in English dialects

- differences within + across dialects in *dynamic quality*

- clear indications of dialect variation in vowels beyond static quality



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)