What is a linguistic variety?

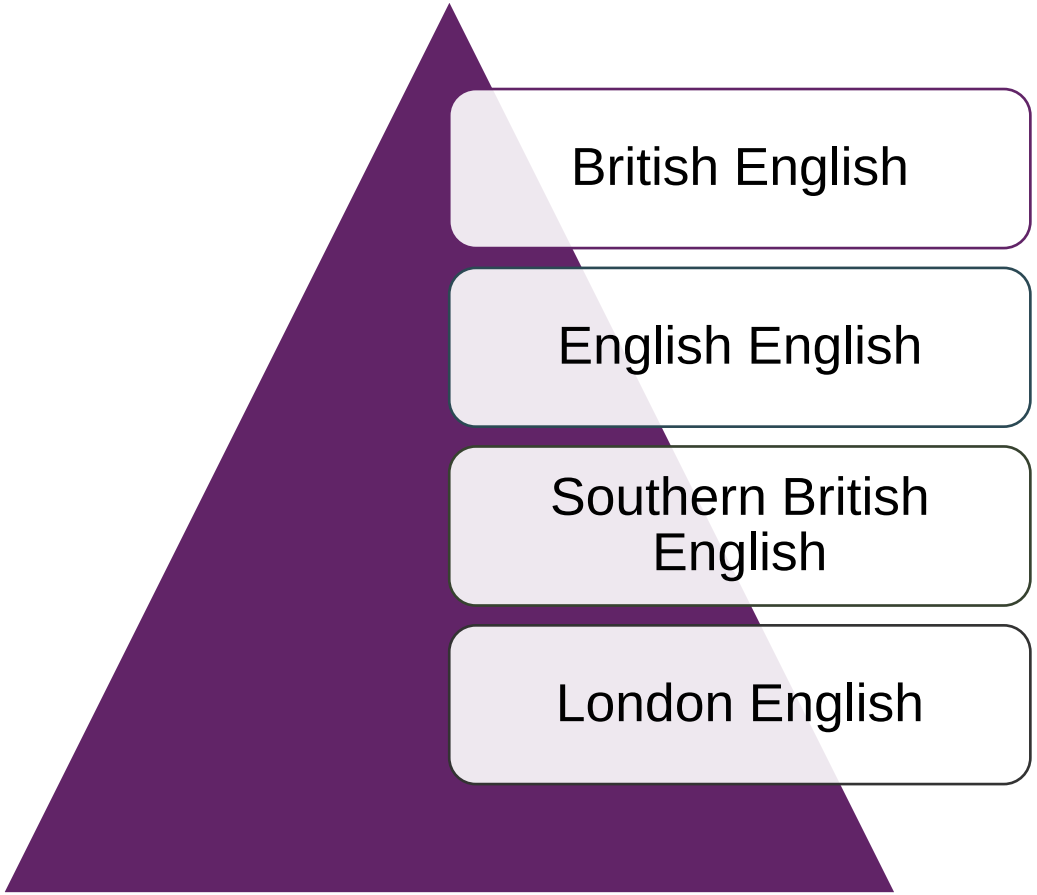
Linguistic coherence, variation and nominalisation processes in Cockney

Dr Amanda Cole

Amanda.cole@essex.ac.uk

Linguistic Data Consortium talk, 30th Nov 2022

What is a linguistic variety?



Top-down approach to documenting linguistic varieties

Top-down approach to documenting linguistic varieties

Issues with this approach

1. “Authentic” speakers from a community are selected the template for the variety which excludes certain groups of speakers.
2. The boundaries of the variety are not defined empirically but are based on ideological and boundary-marking processes.

Who is an authentic speaker?

- Selection criteria around the “authenticity” of a speaker measures them against expectations around linguistic and social belonging
- E.g., **N**on-mobile **O**lder **R**ural **M**ales (NORMs) in Survey of English Dialects
- Our models of linguistic varieties may be biased and exclude certain speakers

London dialect

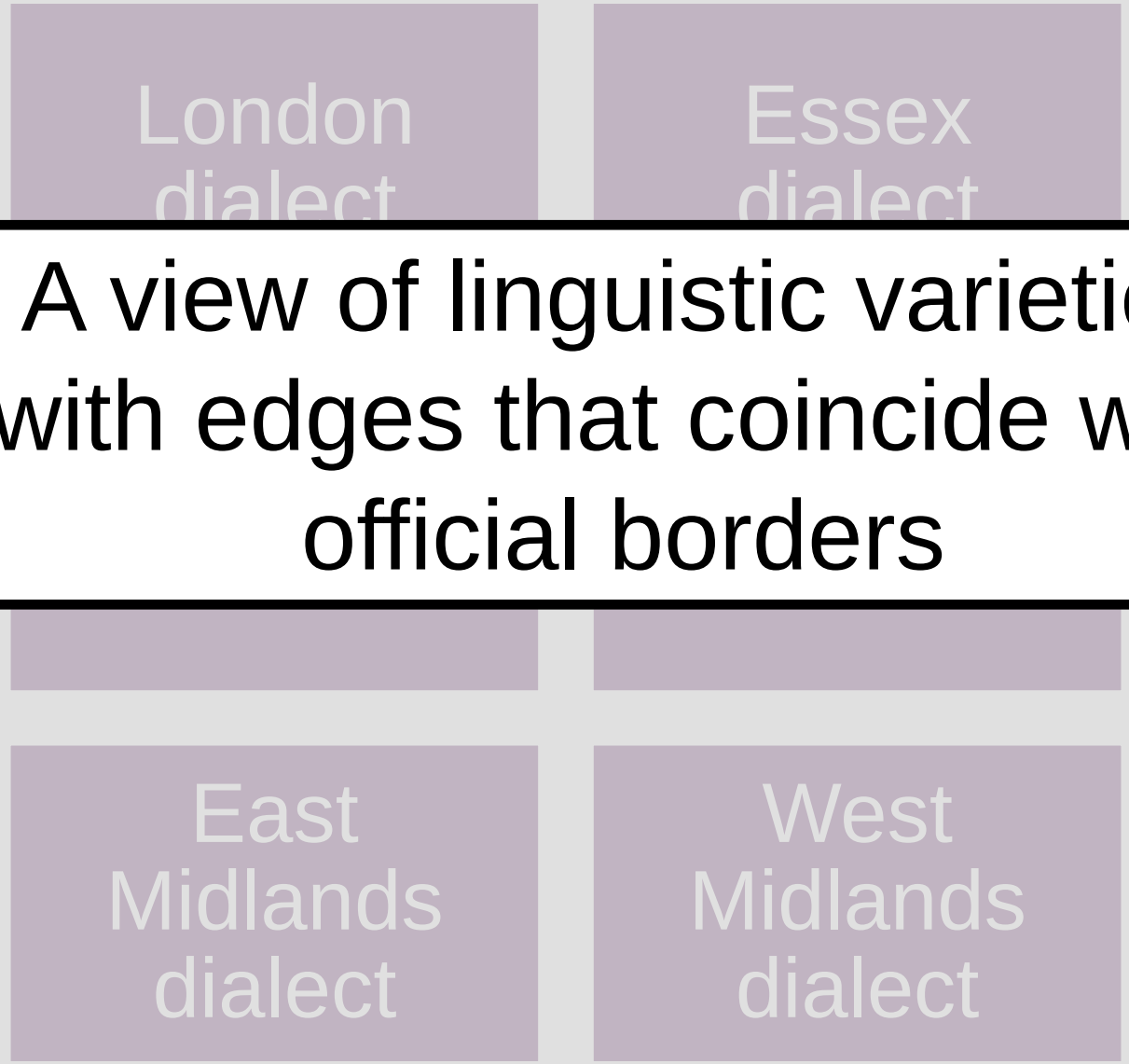
Essex dialect

Suffolk dialect

South West dialect

East Midlands dialect

West Midlands dialect



A view of linguistic varieties with edges that coincide with official borders

Where does one variety end and another begin?

linguistic variation can be categorized as either:

- **different variety**
- variation within the **same variety**

Language vs. dialect

- Danish, Norwegian and Swedish are mutually intelligible but considered **different languages**
- Gheg and Tosk are **dialects** of Albanian despite being mutually incomprehensible
- The Galician spoken in Spain is considered a different **language** to the Galician in Portugal
- There is a linguistic continuum between France and Italy but dialects are officially split according to the official border

Language vs. dialect

concepts of “language” and “dialect” are **nationalistically and ideologically** defined

- Danish, N… considered **different**
- Gheg an… lly incomprehensible
- The Galician spoken in Spain is considered a different **language** to the Galician in Portugal
- There is a linguistic continuum between France and Italy but dialects are officially split according to the official border

The issue

- All varieties occur on a continuum and involve internal variation.
- How do we determine where one continuum ends and another begins, constituting separatable varieties?

What is a linguistic variety?

- A variety is a set of **coherent** linguistic features spoken by a certain language communities with **shared** regional and/or **social** information.

Linguistic coherence

Variable X and variable Y may share a similar social distribution in a speech community

Speakers who have relatively high rates of the vernacular form of variable X should have relatively high rates of the vernacular variant of variable Y

Co-variation for multiple linguistic variables

- A variety should have **co-variation** between a "**cluster** of variables" which are **coherent** within a single accent (Guy 2013: 64)
- If there is not linguistic coherence "the cognitive and social reality of the 'sociolect' is problematic" (Guy 2013: 63)

Studies finding limited linguistic coherence

- Brazilian Portuguese (Oushiro & Guy, 2015)
- New York City English (Becker 2016)
- Copenhagen Danish (Gregersen & Pharao 2016)

Possible reasons for limited linguistic coherence

1. Linguistic coherence is not necessarily a requirement of a lect (Gregersen and Pharao 2016).
2. Other factors mediate rates of co-variation
3. The way we choose speech communities from a **top-down** approach =/ the most logical and verifiable split in terms of **linguistic content**.

Possible reasons for limited linguistic coherence?

1. Linguistic coherence is not necessarily a requirement (Gregersen and Pharao 2016).
2. **Other factors mediate rates of co-variation**
3. **The way we choose speech communities from a top-down approach =/ the most logical and verifiable split in terms of linguistic content.**

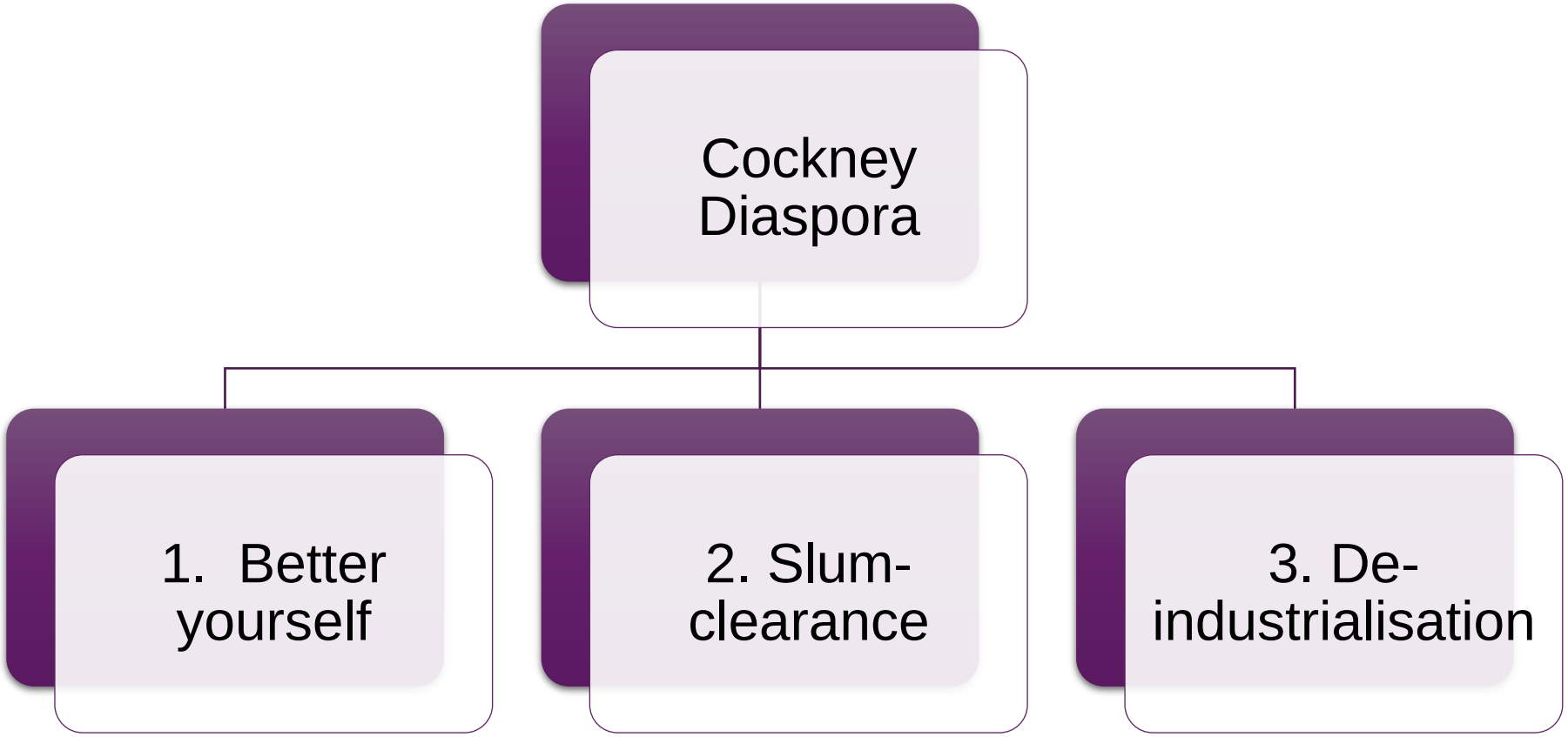
What are the varieties spoken in Southeast England?

Cockney	**working-class** dialect of East London which is the most vernacular in London
	Shifted diphthong system: /ʊ/- and /ɪ/-diphthongs are rotated clockwise and anti-clockwise respectively (see Cole and Strycharczuk 2022).
RP	“standard” accent used as neutral reference point in linguistic research
	Spoken by the **higher classes** and linked to the fee-paying school system
SSBE/SBE	Middle and upper-middle class accent evolved from RP (Lindsey 2019: 4)
	In most studies, speakers of SSBE are required to be southern and/or of a higher class
Estuary English	Somewhere between Cockney and RP
	Spoken across Southeast England by middle- and working-class speakers
MLE	Spoken in London, mainly by young people from an **ethnic minority** background (Cheshire et al. 2011)
	Has some features in common with other southern varieties but has an **innovative diphthong system**

Cockney	**working-class** dialect of East London which is the most vernacular in London
	Shifted diphthong system: /ʊ/- and /ɪ/-diphthongs are rotated clockwise and anti-clockwise respectively (see Cole and Strycharczuk 2022).
RP	"standard" accent used as neutral reference point in linguistic research
	system
SSBE/SB	019: 4)
	d/or of a
Estuary English	Somewhere between Cockney and RP
	Spoken across Southeast England by middle- and working-class speakers
MLE	Spoken in London, mainly by young people from an **ethnic minority** background (Cheshire et al. 2011)
	Has some features in common with other southern varieties but has an **innovative diphthong system**

- Inconsistency and imprecision in the demarcation of these accents based on linguistic and social factors.
- Are these varieties all contemporaneous, who speaks them, and what is the linguistic and social makeup?

The post-WWII relocation of traditional East London communities into the London peripheries, in particular, Essex (see Fox 2015; Watt, Millington & Huq 2014; Cohen 2013).



Milton Keynes
Bletchley
Ampthill
Bedfordshire
Letchworth Garden City
Hitchin
Leighton Buzzard
Dunstable
Luton
Stevenage
Hertfordshire
Bishops Stortford
Aylesbury
Harpenden
Welwyn Garden City
Ware
Hertford
Harlow
Hoddesdon
Berkhamsted
Hemel Hempstead
St Albans
Hatfield
Buckinghamshire
Chesham
Amersham
Potters Bar
Cheshunt
Watford
Borehamwood
Loughton
High Wycombe
Rickmansworth
Gerrards Cross
Maidenhead
Slough
Windsor
Berkshire
London
Woodley
Wokingham
Staines
Egham
Ashford
Walton-on-Thames
Esher
Weybridge
Ewell
New Addington
Sandhurst
Camberley
Woking
Epsom
Farnborough
Leatherhead
Fleet
Caterham
Bracknell
Redhill
Guildford
Reigate
Farnham
Surrey
Godalming
Horley
Crawley
East Grinstead
Horsham
Crowborough
Sudbury
Ipswich
Harwich
Colchester
Braintree
Essex
Witham
Clacton-on-Sea
Chelmsford
Maldon
Billericay
Wickford
Brentwood
Basildon
Rayleigh
Southend-on-Sea
Stanford le Hope
Canvey Island
South Ockendon
Grays
Northfleet
Medway
Swanscombe
Gravesend
Rochester
Gillingham
Chatham
Minster
Whitstable
Herne Bay
Sittingbourne
Faversham
Ditton
Canterbury
Sevenoaks
Maidstone
Tonbridge
Kent
Royal Tunbridge Wells
Ashford
Folkestone

University of Essex

Research site: The Debden Estate

University of Essex

Did Cockney move to Essex?

- First generation to grow up in Debden after their parents were relocated by the LCC in late 1940s or 1950s
- 15 participants (female = 9; male =6) born between 1944 and 1969 (mean = 1956; SD = 6.7 years).
- All lived in Debden for the remainder of their lives
- Either born in Debden (n= 10) or moved there with their families before the age of five (n=5).

Cole 2022

Shared between Debden and “Cockney”

1. shifted vowel system
2. [ɐ] for LETTER/COMMA
3. THOUGHT/NORTH/FORCE distinction
4. L-vocalisation
5. Roland-roller
6. [n] variant of (ING)
7. TH-fronting/stopping
8. H-dropping
9. glottal replacement/reinforcement
10. labiodental /ɹ/
11. verbal {-s} over-generalisation
12. distinction between CLOTH and LOT
13. past tense form for past participle (including *done* for *did*)
14. levelling to *was* in positive contexts but *weren’t* in negatives
15. *come* as past tense form
16. negative concord
17. prevalence of *ain’t*, *never* and *init*
18. object pronouns in standard subject pronoun positions
19. *what* or zero relative pronouns
20. reflexive pronouns *hisself* and *theirselves*
21. -n occasionally forming possessive pronouns
22. nonstandard production of *my*
23. adjectives in adverbial positions
24. *them* for *those*

University of Essex

Differences between Debden and "Cockney"

1. **"Cockney" features extended in Debden**
   - /l/-vocalisation between vowels
2. **Previously reported as relics of Cockney**
   - schwa epenthesis,
   - -n possessive pronouns
   - CLOTH/LOT split
3. **"Cockney" features Not found in Debden**
   - yod deletion not as extensive
   - glottal replacement is not observed for /p/
   - short front vowels lower than anticipated
4. **Non-standard features in Debden but not previously documented in Cockney.**
   - *bring*/*buy* levelling in past tense and the past participle forms
   - *init* as an invariant tag
5. **What is the benchmark of "Cockney" for comparison?**
   - Wells (1982); Tollfree (1999); Mott (2012); Hughes, Trudgill & Watt (2012) Wright (1981); Sivertsen (1960)
   - Differences in these accounts due to regional and social variation; different methods, language change

Differences between Debden and “Cockney”

1. “**Cockney” features extended in Debden**
   - /l/-vocalisat
2. **Previousl**
   - schwa epe
   - -n possess
   - CLOTH/LOT
3. “**Cockney**
   - CLOTH/LO
   - yod deletion not as extensive
   - glottal replacement is not observed for /p/
   - short front vowels lower than anticipated

4. **Non-standard features in Debden but not**

ey.

e past

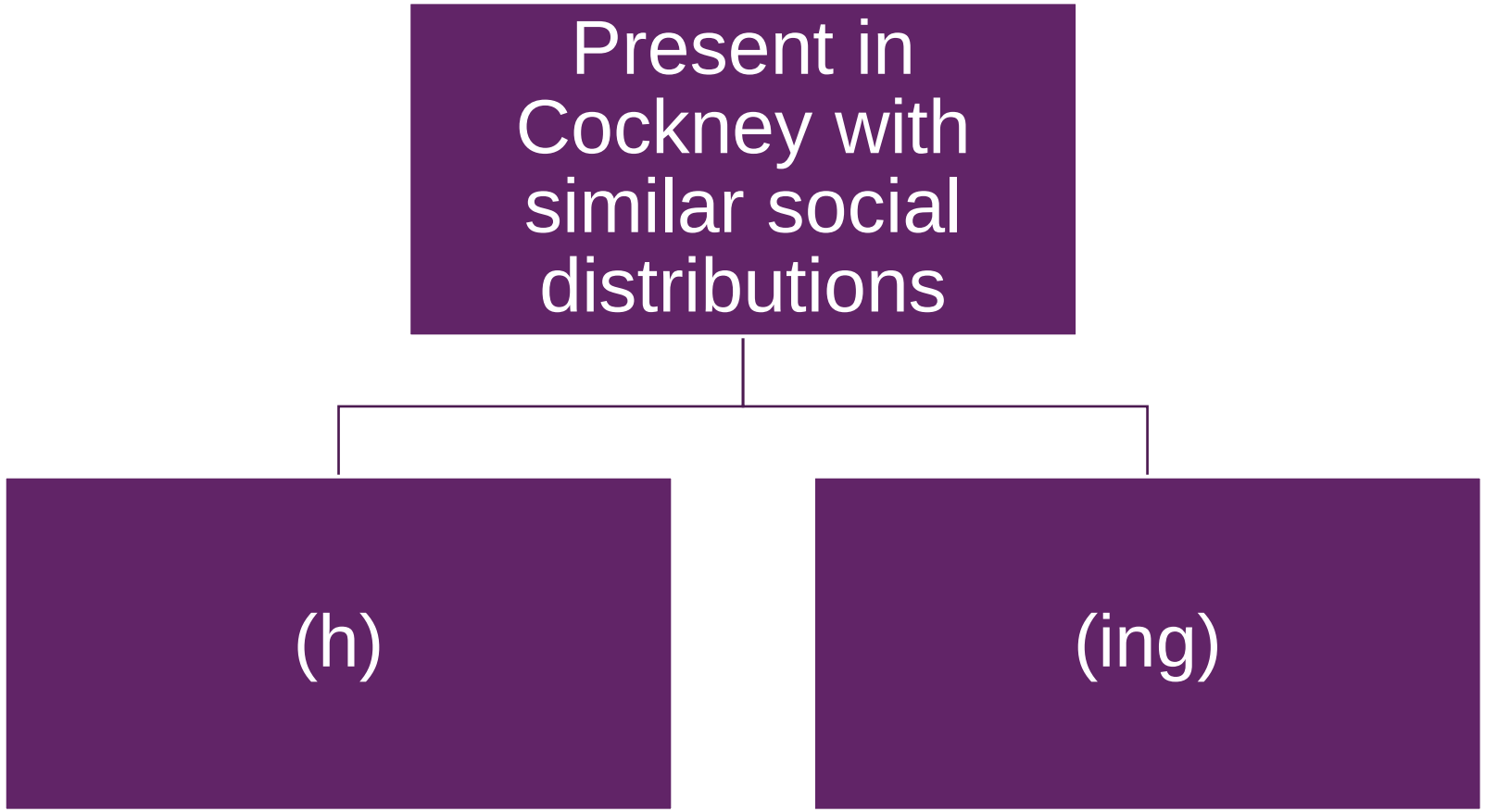
ney” for

(2012);
Hughes, Trudgill & Watt (2012) Wright (1981); Sivertsen (1960)

- Differences in these accounts due to regional and social variation; different methods, languag change

There seems to be sufficient overlap between Debden and Cockney to conclude that Cockney has moved to Debden.

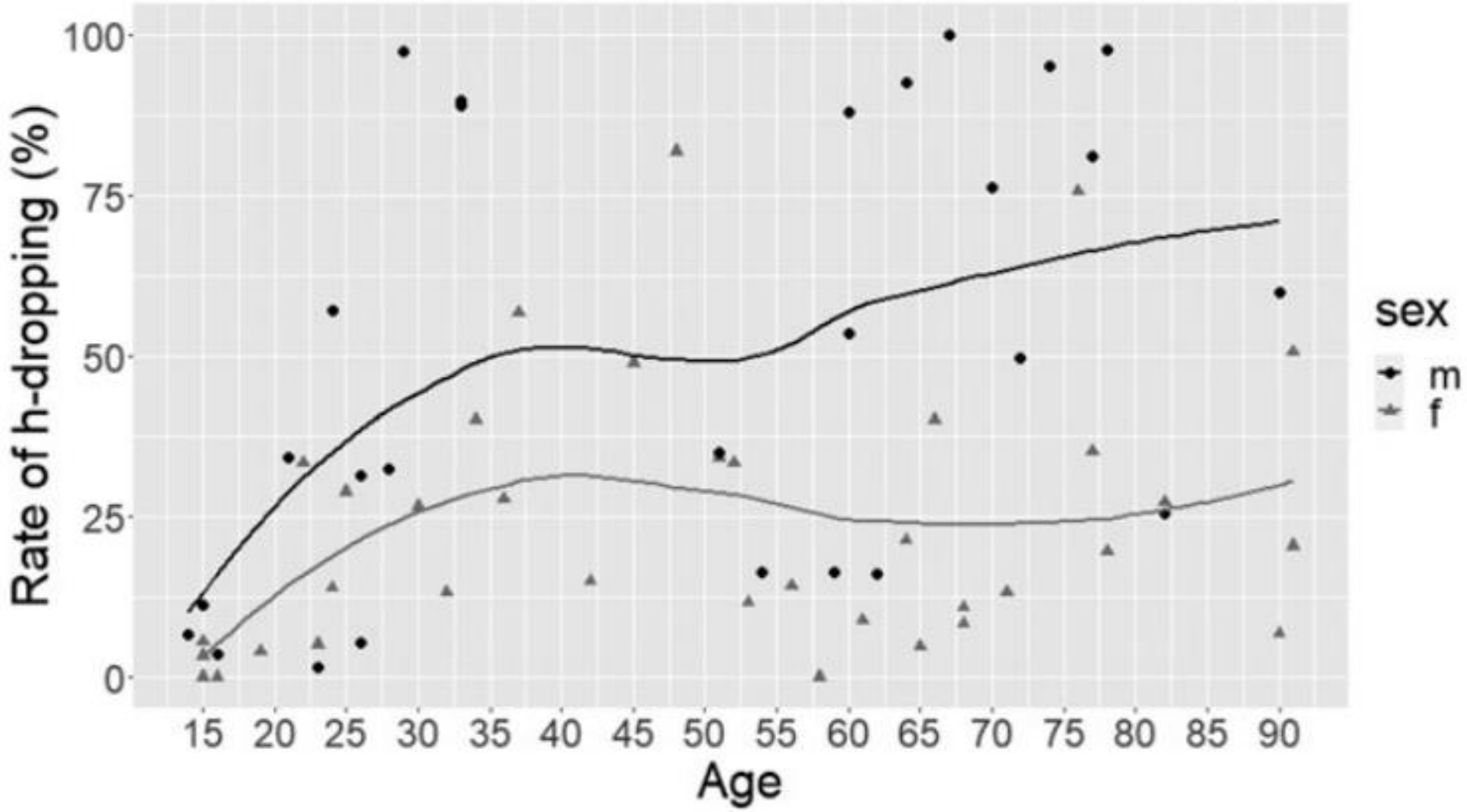
Is the Debden variety linguistically coherent?



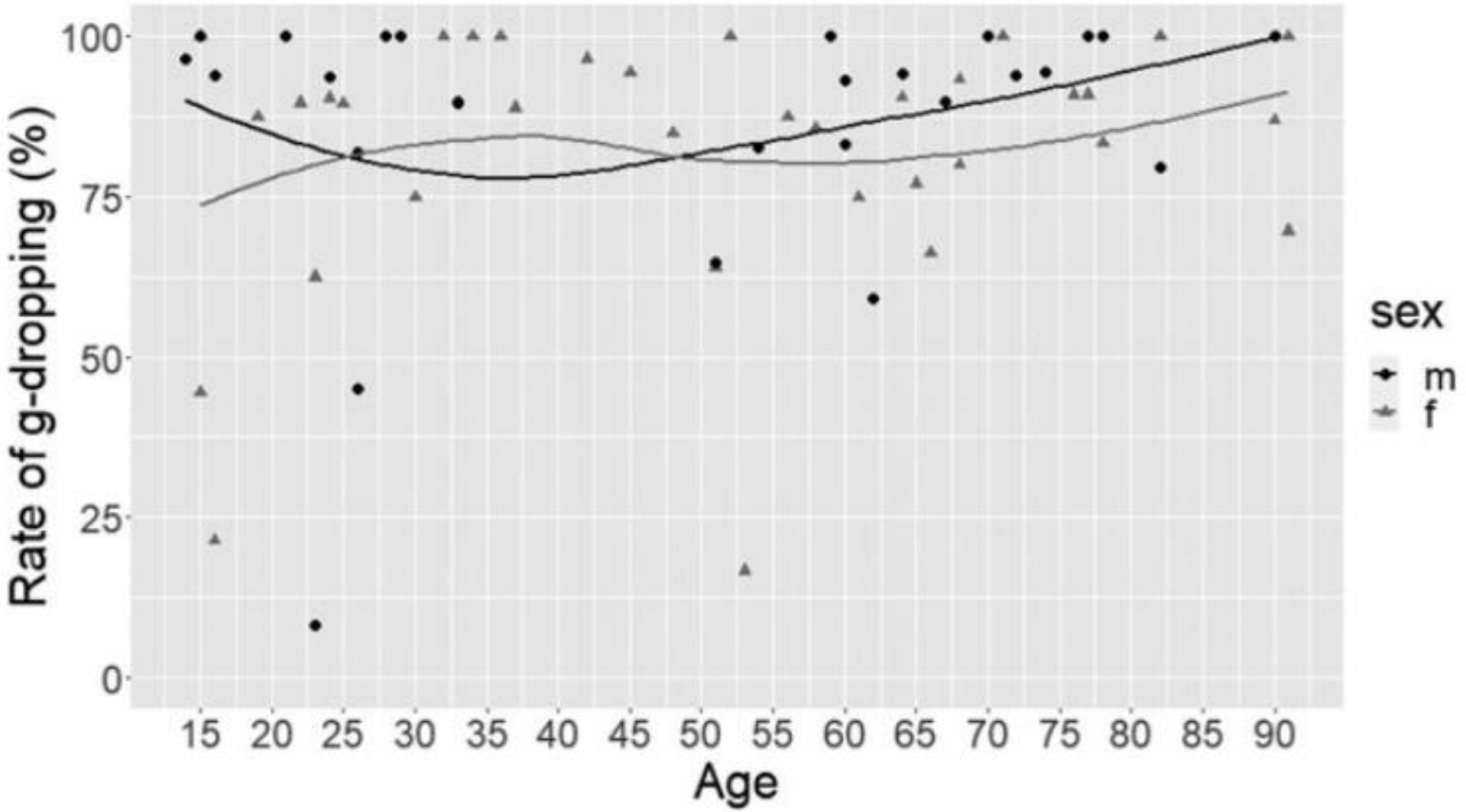
Speakers and tokens

- 63 speakers (36 female) from Debden
- 14 to 91 years of age (M= 49.3yrs, SD = 23.8)
- Rates of (h) and (ing) coded from casual speech (range 20 – 50 mins)
- 2,183 tokens of (ing) and 4,058 tokens of (h)

Rates of (h) in Debden

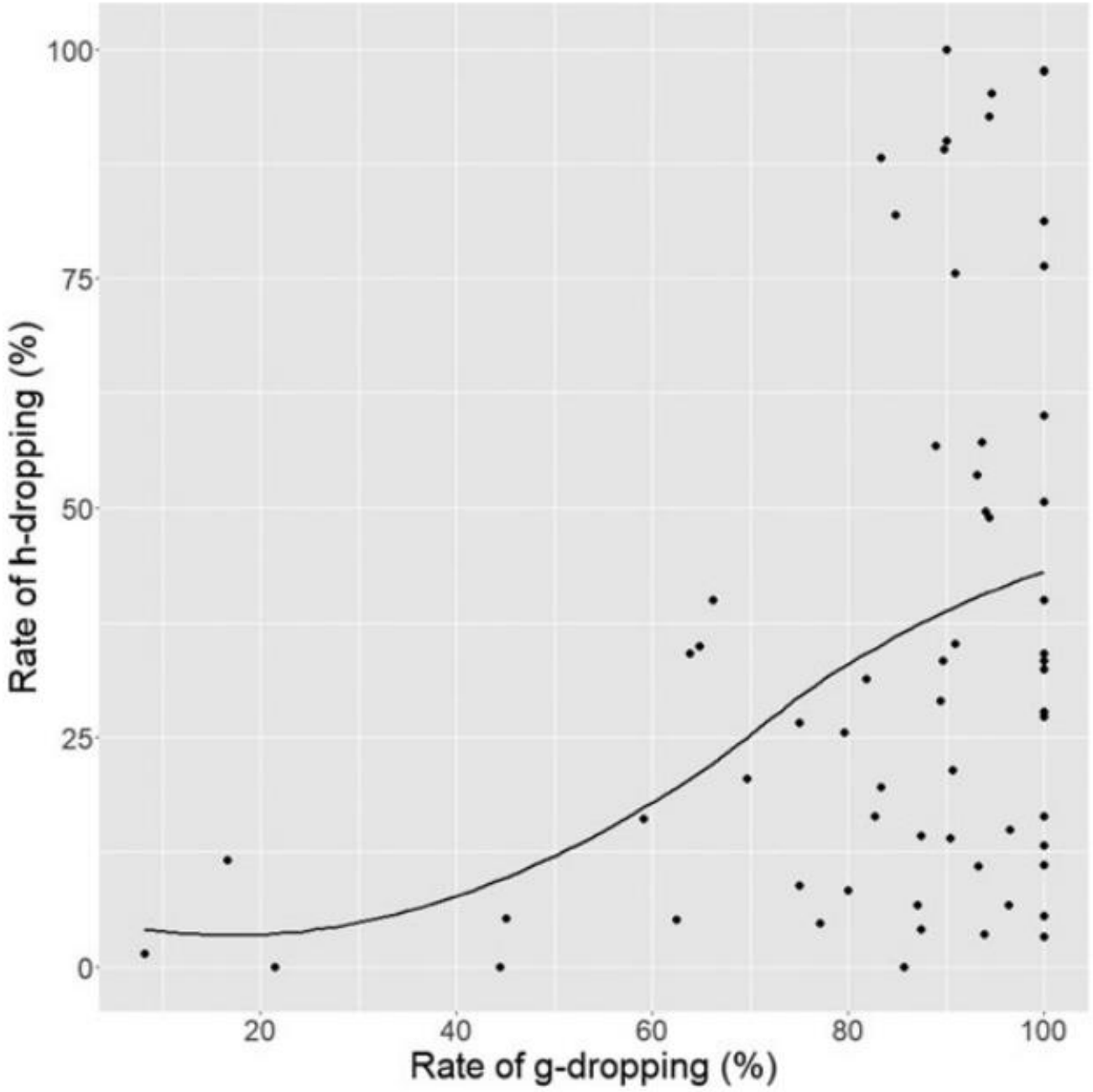


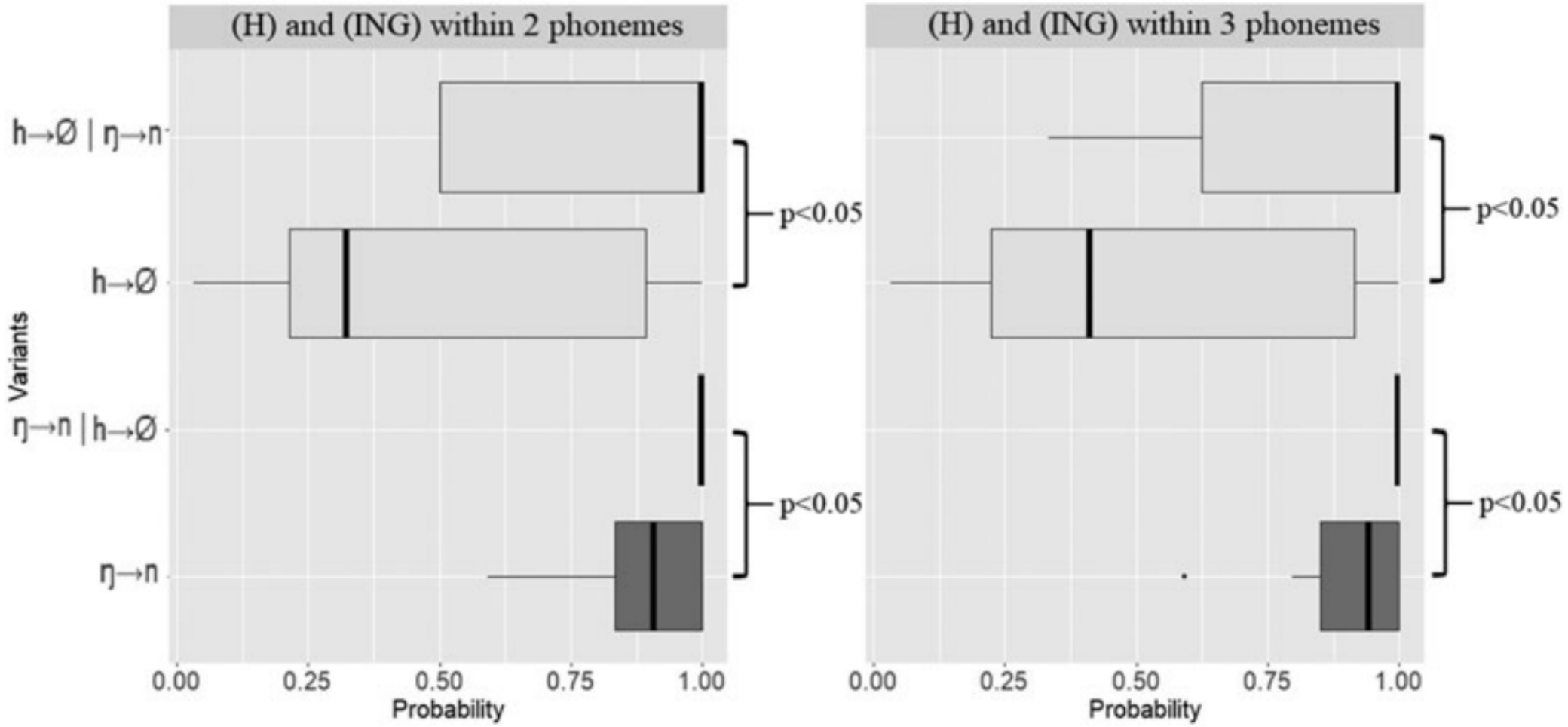
Rates of (ing) in Debden



There is a weak correlation (r = 0.36) between rates of (ING).

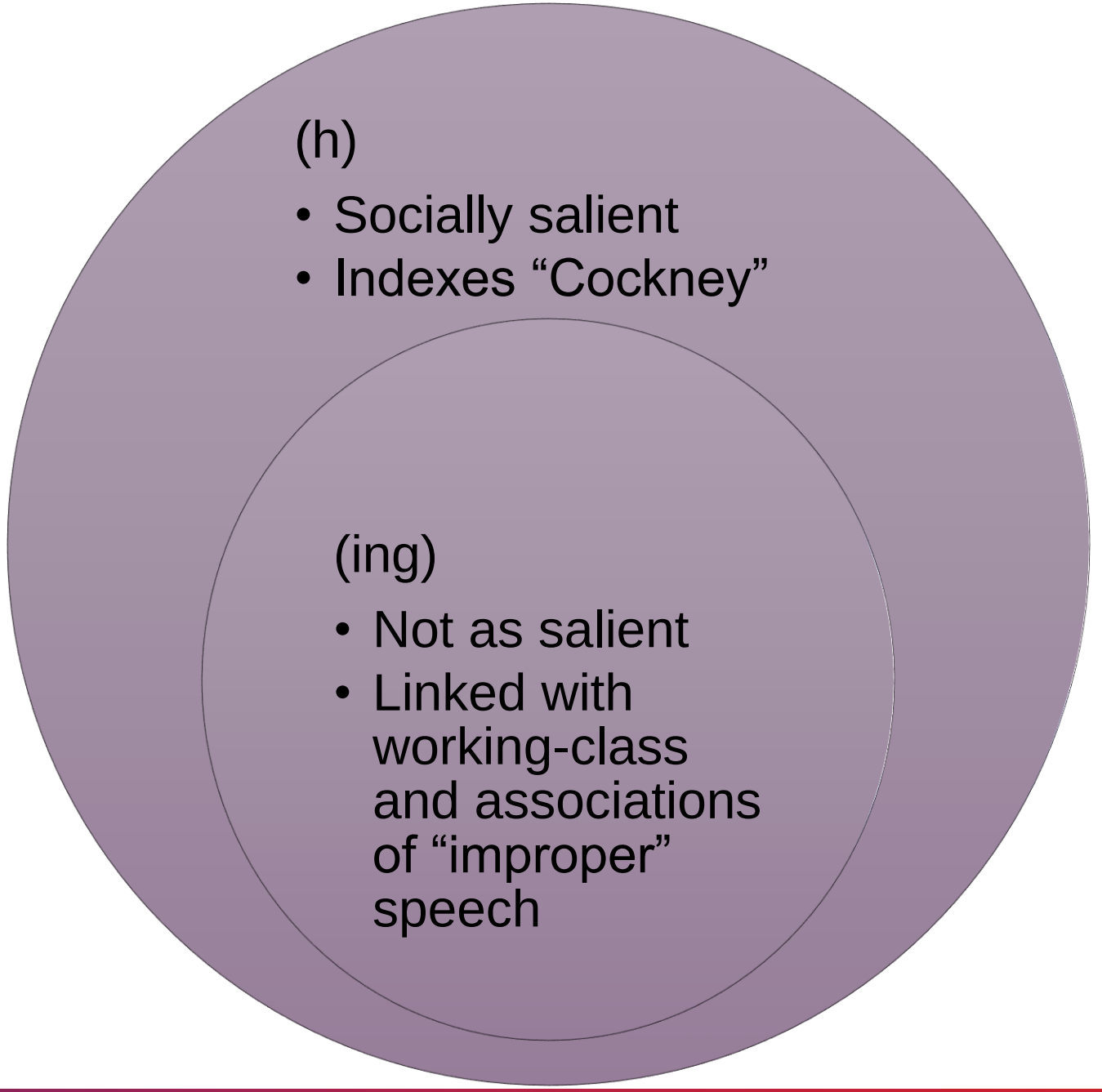
There is an implicational relationship: h-dropping implies g-dropping, but the reverse is not true





Speakers are significantly more likely to produce h-dropping within two (left panel) or three (right panel) phonemes of g-dropping compared to the probability of h-dropping occurring independently (and vice-versa).

Implicational relationship between (h) and (ing) due to overlapping social meanings



Metalinguistic commentary on (ing)

I was saying, 'I'm going to speak much better today, I'm going to speak and I'm going to say all my words properly and all my letters properly.' And they were laughing at me 'cause I suppose I'll say 'laughin'' and 'jokin'' and we don't put a 'g' on the end and–but I know–it was far too much effort 'cause it's not me, is it?

51-year-old woman

Metalinguistic commentary on (h)

Well, it seems - it seems to me that if people can't pronounce their words properly, they seem to – they assume you come from London, init. If they're not saying their t's or h's or anything like that, there's–they'll say, "Oh, you come from London then, don't you?"

54-year-old man

Possible reasons for limited linguistic coherence

1. Linguistic coherence is not necessarily a requirement (Gregersen and Pharao 2016).
2. **Other factors mediate rates of co-variation**
3. **The way we choose speech communities from a top-down approach =/ the most logical and verifiable split in terms of linguistic content.**

Bottom-up approach to splitting linguistic varieties

- Trudgill (1990:32) draws isoglosses on a single composite map of England for eight linguistic features
- Splits England into thirteen dialect areas based on where the most “abrupt transitions” or overlapping isoglosses occur (Trudgill 1990:32)
- “we draw boundaries between dialect areas at places where we find a situation most closely resembling an abrupt transition.” (Trudgill 1990: 6)

University of Essex

The search for linguistically coherent accents: Unsupervised clustering of diphthong variation in Southeast England

Amanda Cole[1] & Patrycja Strycharczuk[2]

University of Essex[1] University of Manchester[2]

The premise

- A variety includes cluster of variables which are coherent
- A single variety has a centre of gravity
- But internal variation within a single variety exists
- A variety is a set of linguistic features spoken by a certain language communities.

Speakers

- 193 speakers from across Southeast England
- Lived there for at least half of the years between the ages of 3 and 18
- Born 1986 – 2001

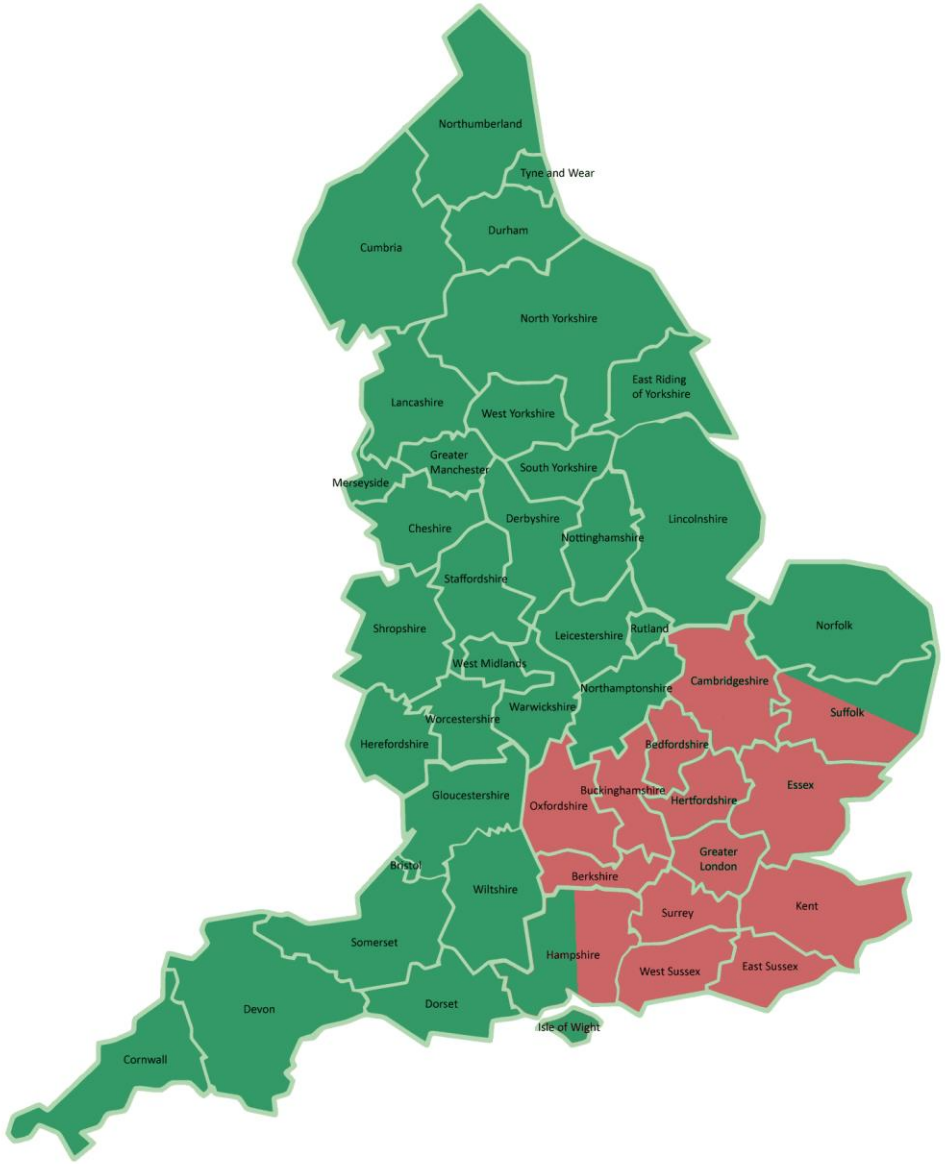


Image adapted from Visit North West

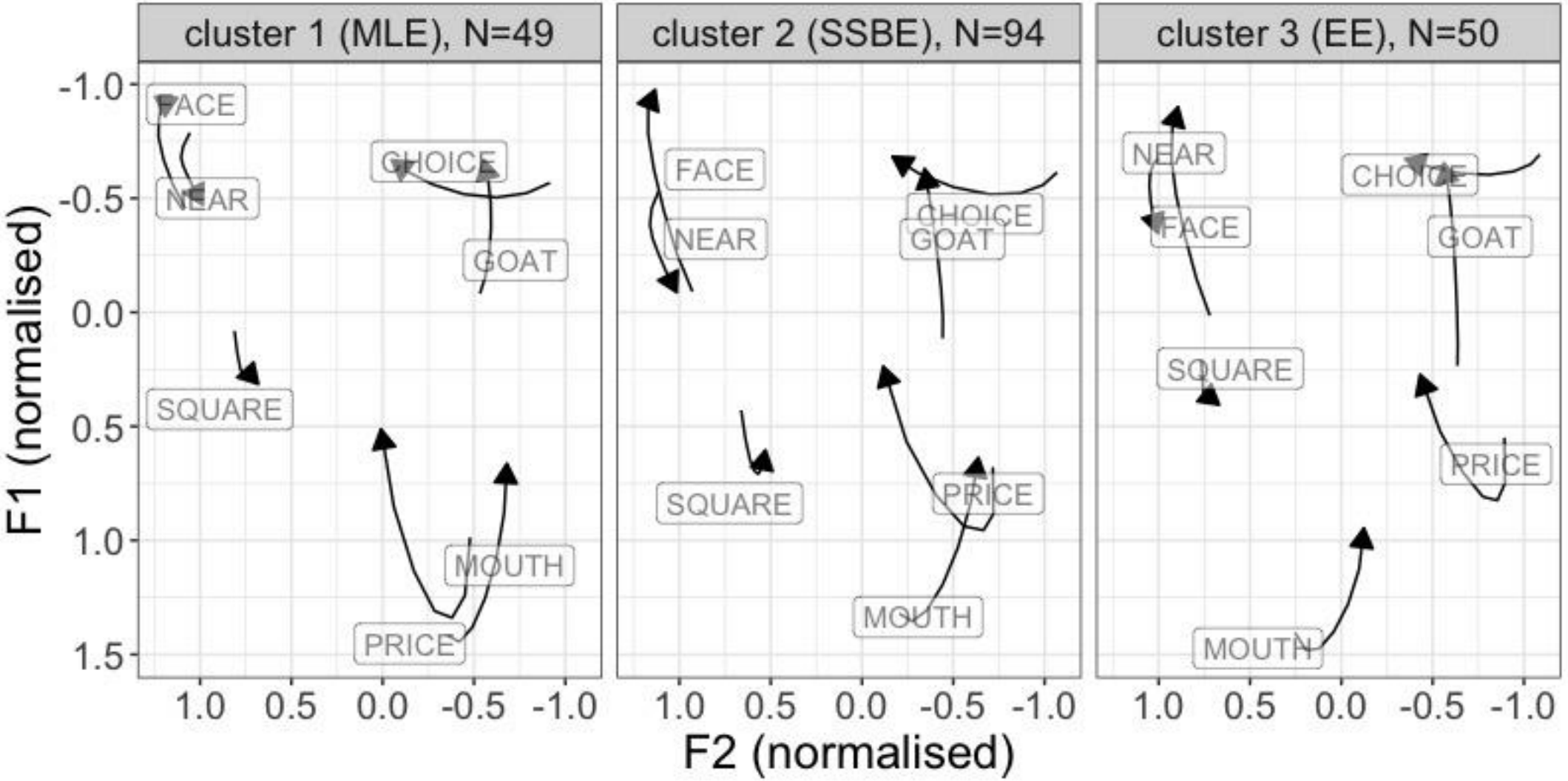
Data

- Passage and wordlist data
- Productions were transcribed in ELAN
- Automatic segmentation and extraction with FAVE (Rosenfelder et al., 2014).
- F1 and F2 measured at 10% intervals
- Outliers removed and z-score normalised the F1 and F2 values within speaker (Lobanov, 1971)
- Vowels in analysis: FACE, PRICE, CHOICE, MOUTH, GOAT, NEAR, SQUARE

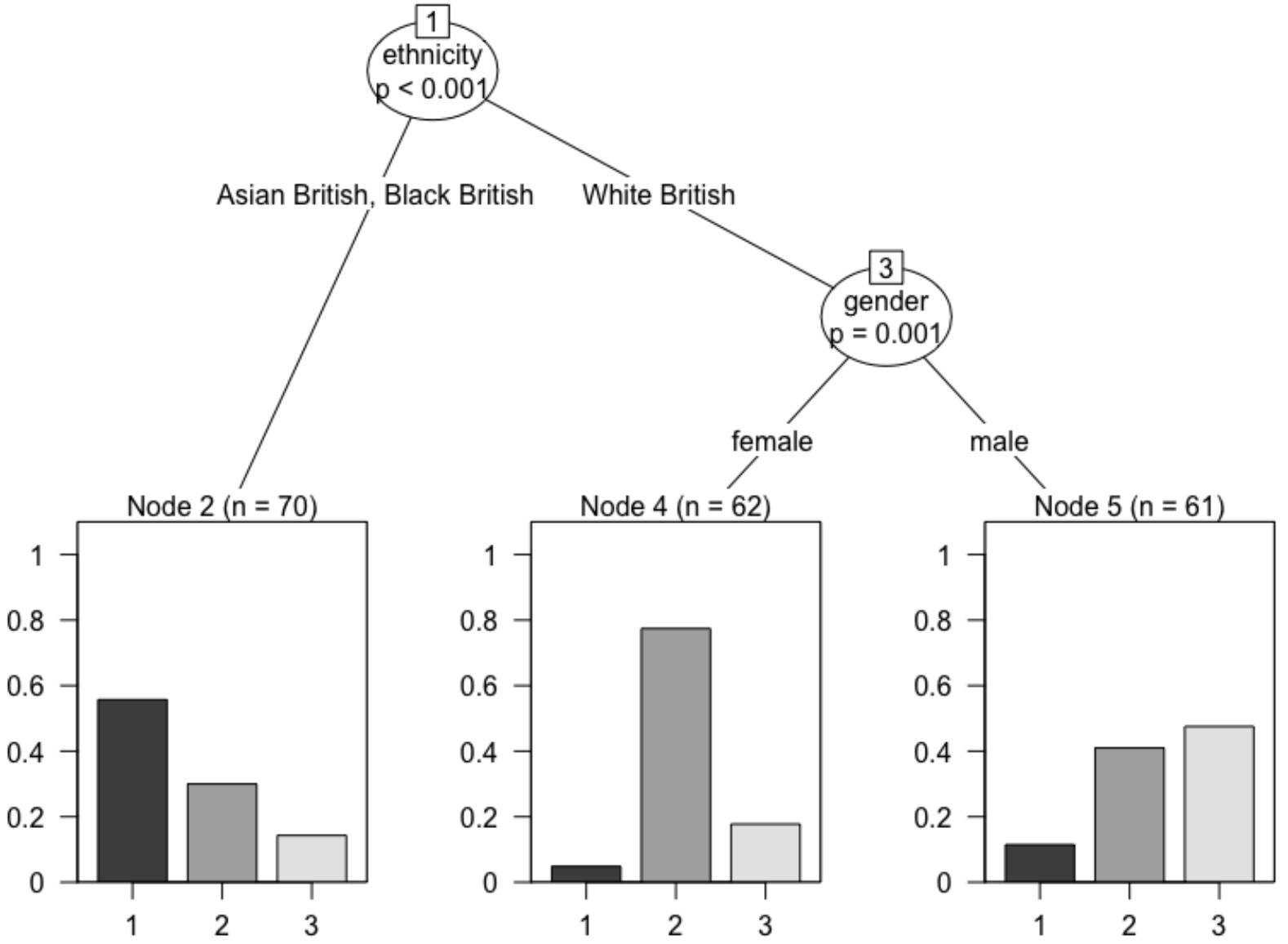
Analysis

- functional Principal Component Analysis (fPCA; Gubian, Torreira & Boves, 2015) applied to the formant values
- Principal Components entered into a clustering analysis, exploring patterns of co-occurrence
- conditional inference tree: how does speakers' **social information** relate to their **cluster membership**
  - age, gender, ethnicity, social class and region (London vs. other south-eastern counties)

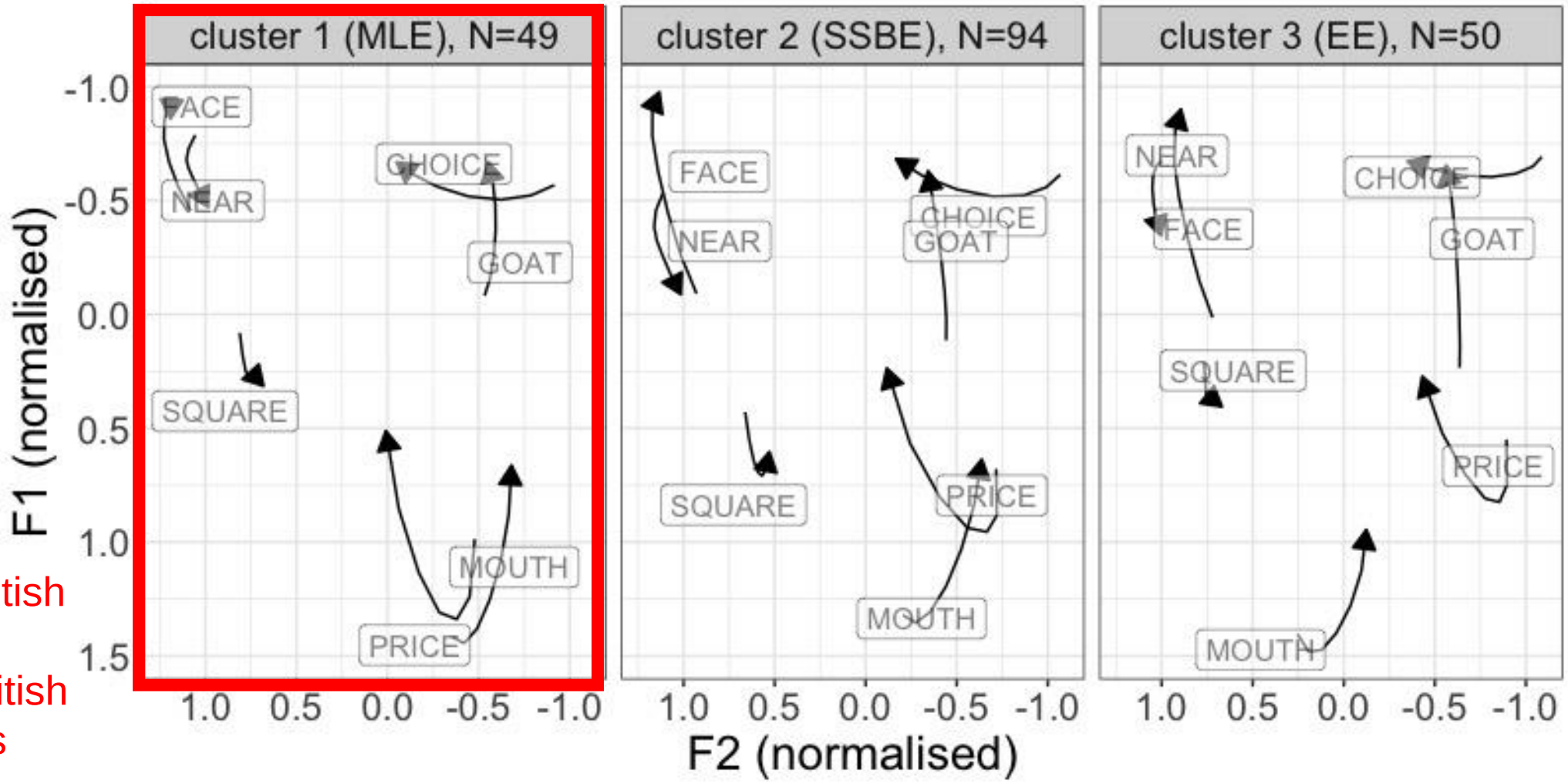
Results: cluster analysis



mapping cluster membership onto social predictors



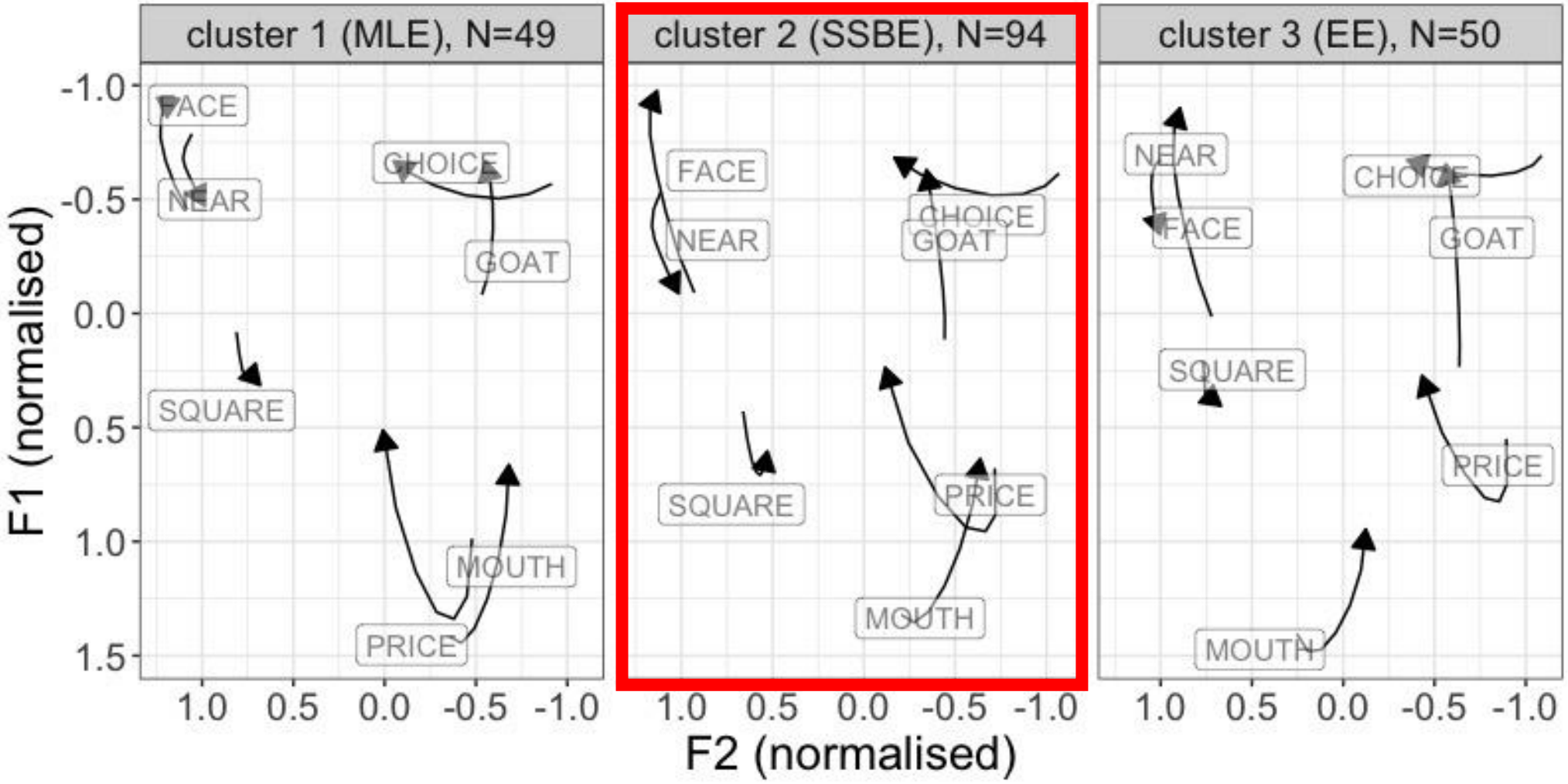
mapping cluster membership onto social predictors



mainly Black British and Asian British speakers

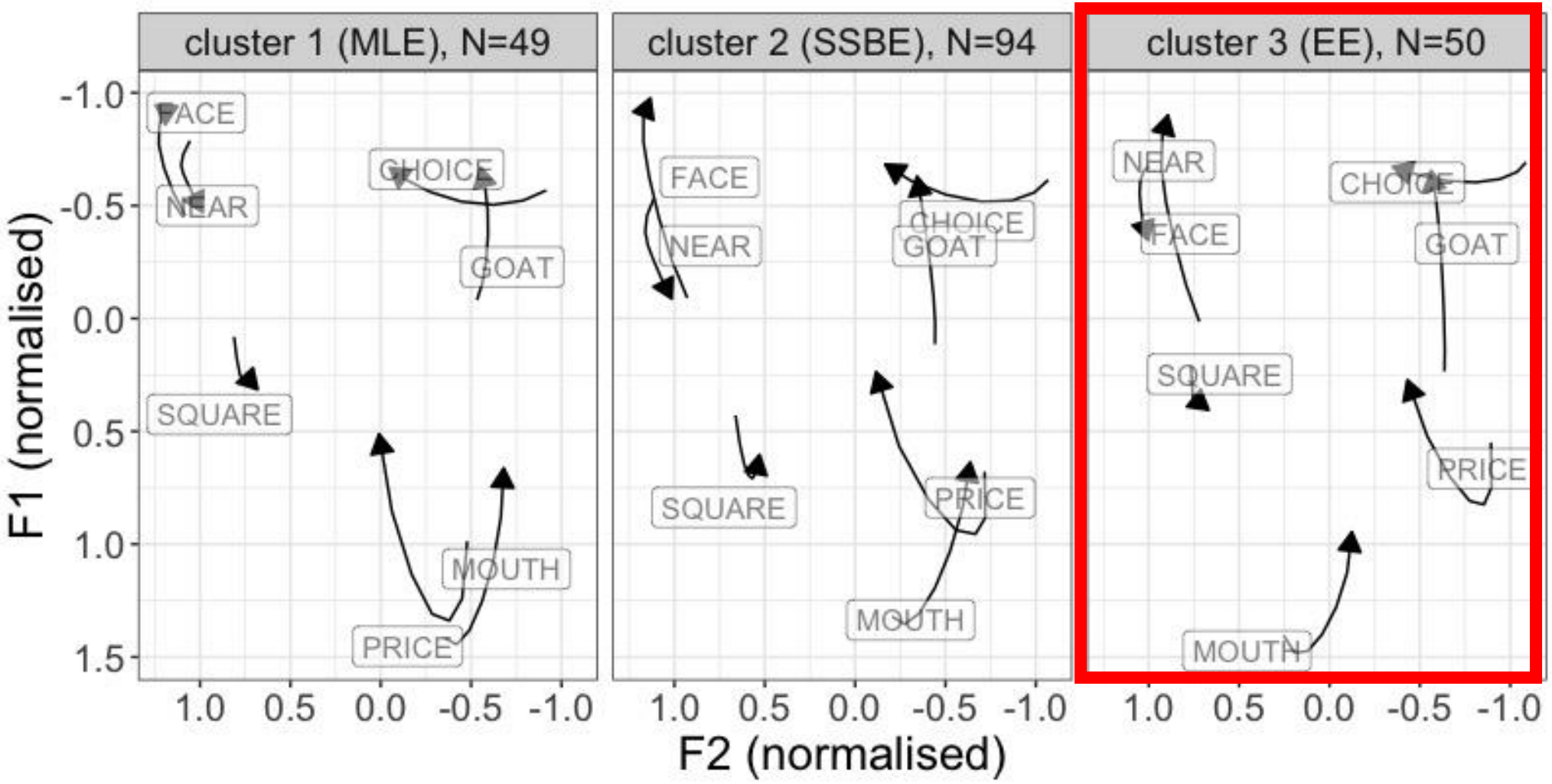
mapping cluster membership onto social predictors

mainly White British females



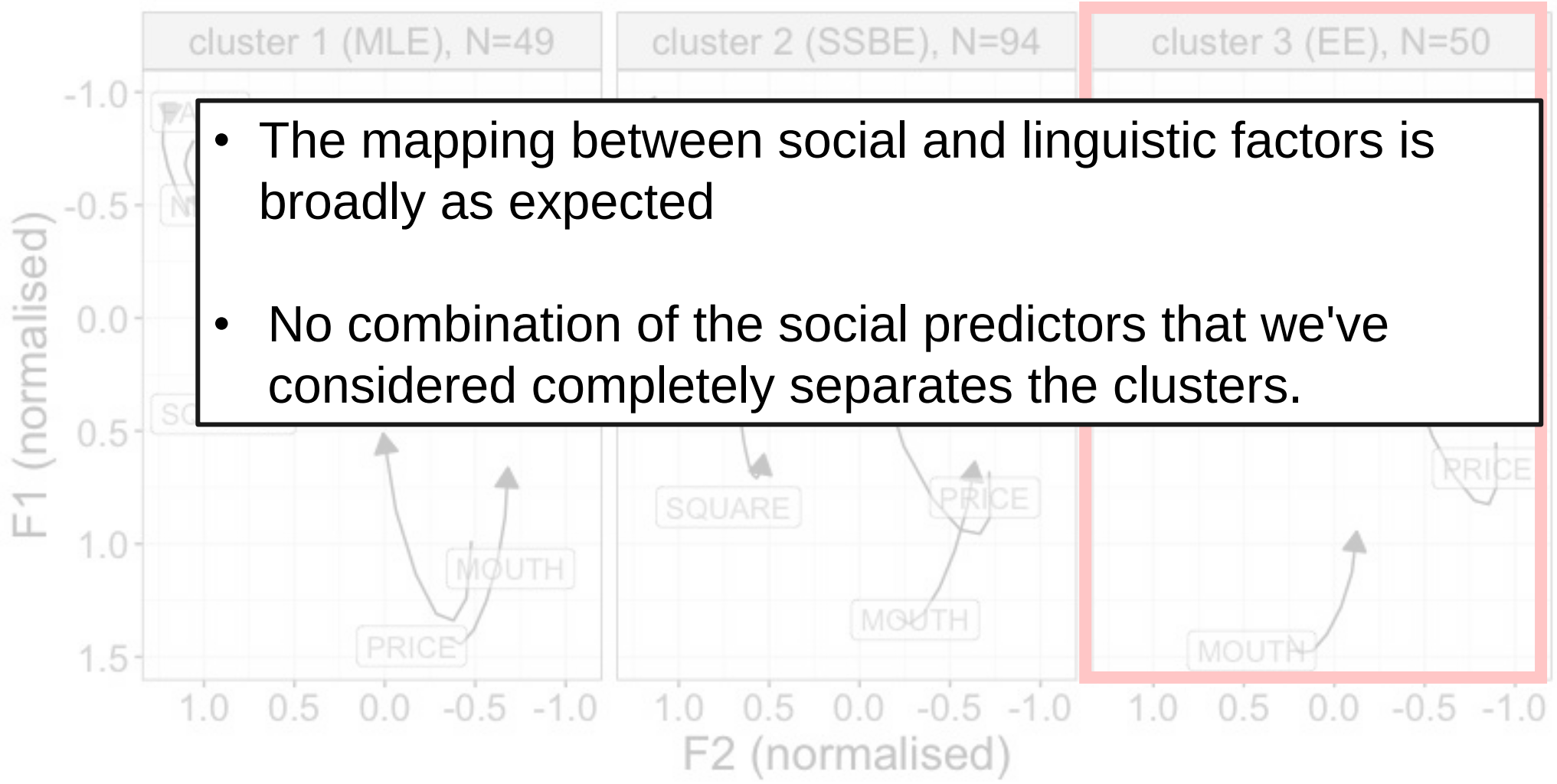
mapping cluster membership onto social predictors

mainly White British males



mapping cluster membership onto social predictors

mainly White British males



- The mapping between social and linguistic factors is broadly as expected
- No combination of the social predictors that we've considered completely separates the clusters.

Contributions of the study

- A method of uncovering variation in the data without pre-determining of **social groups** or **linguistic factors**.
- We show structured variation in vowel production in line with three previously described accents: **MLE, SSBE and EE**
- The linguistic centre of gravity for these three diphthong systems can be used as reference points in future research

Possible reasons for limited linguistic coherence

1. Linguistic coherence is not necessarily a requirement (Gregersen and Pharao 2016).
2. **Other factors mediate rates of co-variation**
3. **The way we choose speech communities from a top-down approach =/ the most logical and verifiable split in terms of linguistic content.**

What is a linguistic variety?

- A variety is a set of **coherent** linguistic features spoken by a certain language communities with **shared** regional and/or **social** information.

Thank you ☺