

What is 'typicality'? Salience and standardness in UK accents

CARMEN LLAMAS, VINCE HUGHES (UNIVERSITY OF YORK)
THOMAS KETTIG (YORK UNIVERSITY)

METHODS IN DIALECTOLOGY XIX

VANCOUVER, BC

AUGUST 7, 2026

Humans & Machines: Novel methods for assessing speaker recognition performance

- Forensic interest in performance of lay listeners at speaker recognition
- Limited amount of work on recognition of unfamiliar voices. We aim:
 1. To compare and combine human and automatic speaker recognition judgements and evaluate them on the same scales
 2. To explore how human judgements are affected by cognitive biases related to criminal trials



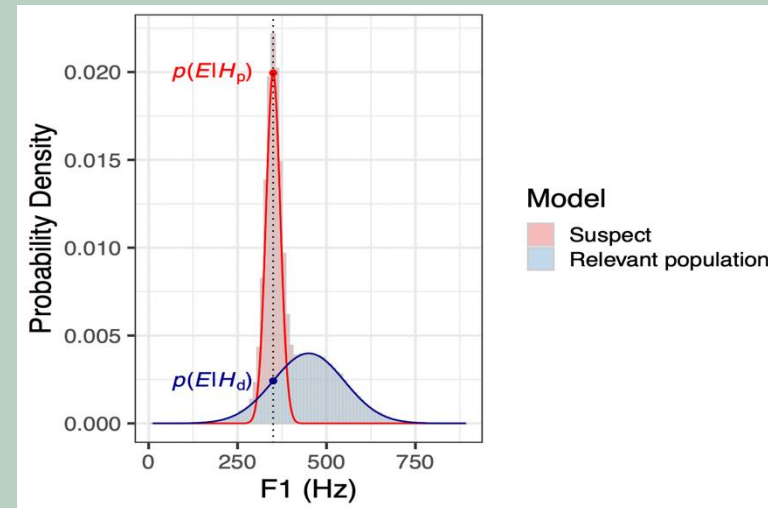
Humans and Machines project

- Previous work with binary/Likert scale responses – not comparable with automatic systems
- Forensic phoneticians express the likelihood that two samples come from the same person via a Likelihood Ratio (LR):

Similarity – “Prosecution hypothesis”

Typicality – “Defense hypothesis”

ASR systems compare these measures using available data



$$LR = \frac{0.02}{0.0025} = 8$$

- But what does it mean to be “typical”?

What is typicality?

- In forensics: would include features like voice quality, clinical idiosyncrasies
 - Purely frequency: Typical = what happens most in the group/context relevant to that case
 - e.g. if someone has a lisp, they are less “typical”, and therefore we can be more confident in our comparison of two samples
- In sociolinguistics/dialectology?
 - Fischer’s (1958) pre-Labovian study of use of [ɪn] or [ɪŋ] by New England schoolchildren – differentiated between ‘model’ boy and ‘typical’ boy: ‘typical’ boy had higher usage of [ɪn]
- So... is a typical speaker one who uses more **non-standard** features?
- But then what does it mean to be a “typical” speaker of a **standard** accent?

Method

Immersive jury game

Participants encounter pairs of sound samples (24 pairs per person), judge:

- **Typicality** of the first voice
 - “This is a Newcastle speaker. How typical is this voice relative to other speakers of the same accent?”
- **Similarity** of the two voices
- Whether they thought the speakers were the **same** or not

0–100 scale for each response

Two conditions

- Immersive jury game (n = 1,505)
- Qualtrics (n = 301)

Stimuli (for first sample)

- **Standard Southern British English**
 - 30 samples from 30 speakers from DyViS (Nolan et al. 2009)
- **Newcastle**
 - 45 samples from 15 speakers from TUULS (Watt et al. 2018)
- **Middlesbrough**
 - 45 samples from 15 speakers from TUULS (Watt et al. 2018)
- All voices from male speakers
- Listeners never rated typicality of same voice twice
- Forensically-realistic quality
 - First sample = landline phone quality
 - Second sample = high quality, from mock police interviews
 - Short (10-11 seconds)

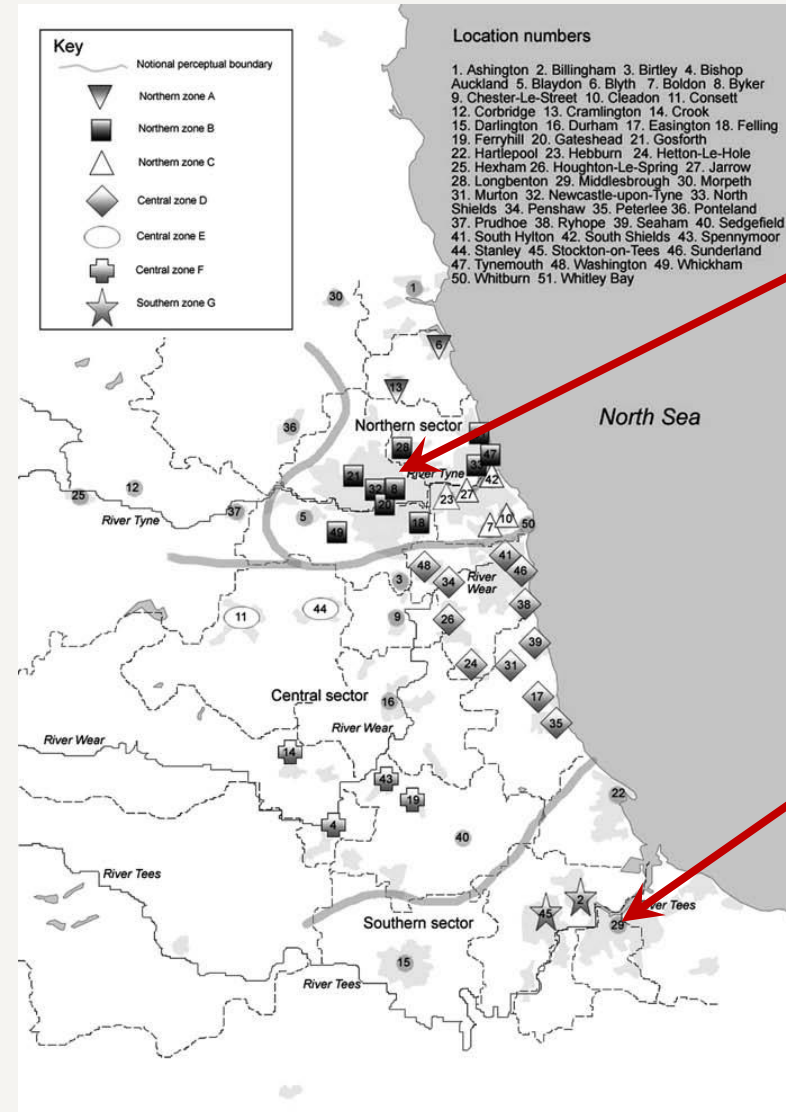


Map 1. The North East of England (from Buchstaller et al. 2011:3, based on two outline images: UK and Ireland)

North East of England

‘No one from Middlesbrough would mistake a Tynesider [Newcastle] for someone from Middlesbrough – but the accents are sufficiently similar to be grouped together, and sufficiently different from those of other areas. Londoners, for instance, might mistakenly think that Middlesbrough speakers were from Newcastle, but they would be much less likely to think that they were from, say, Sheffield.’

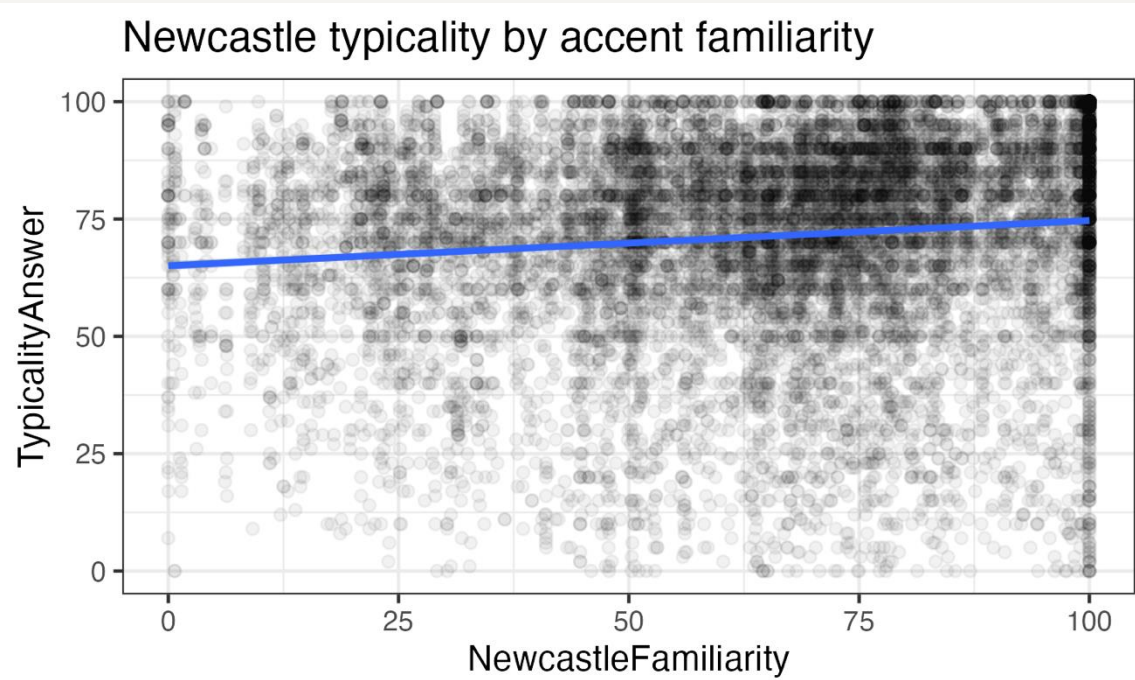
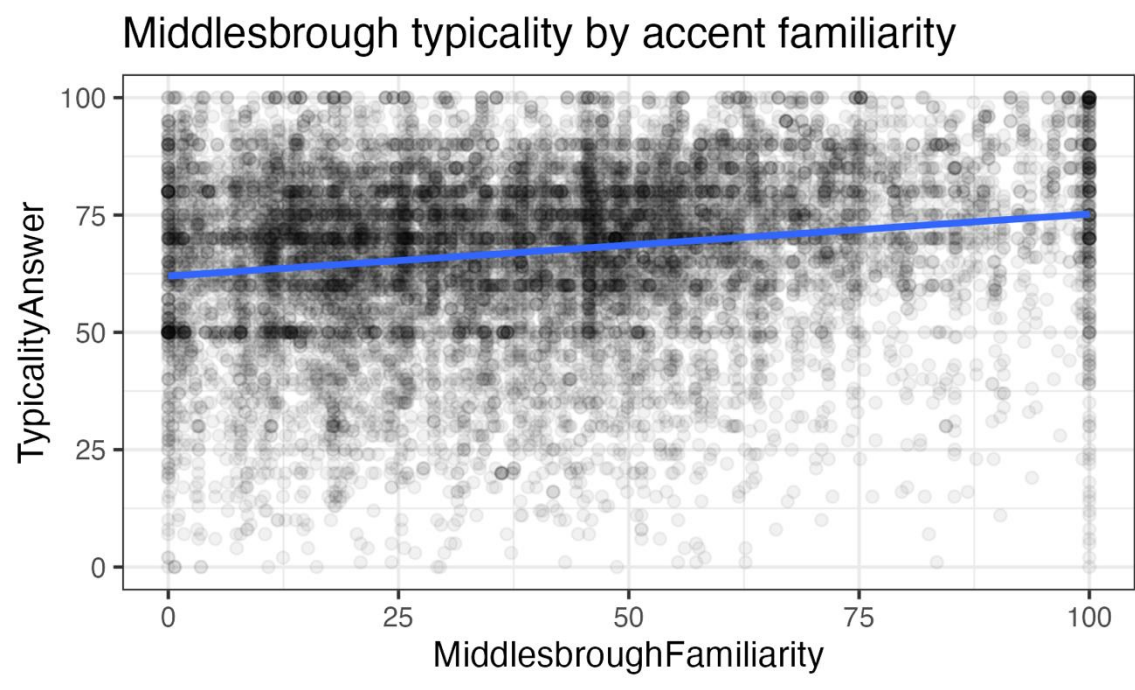
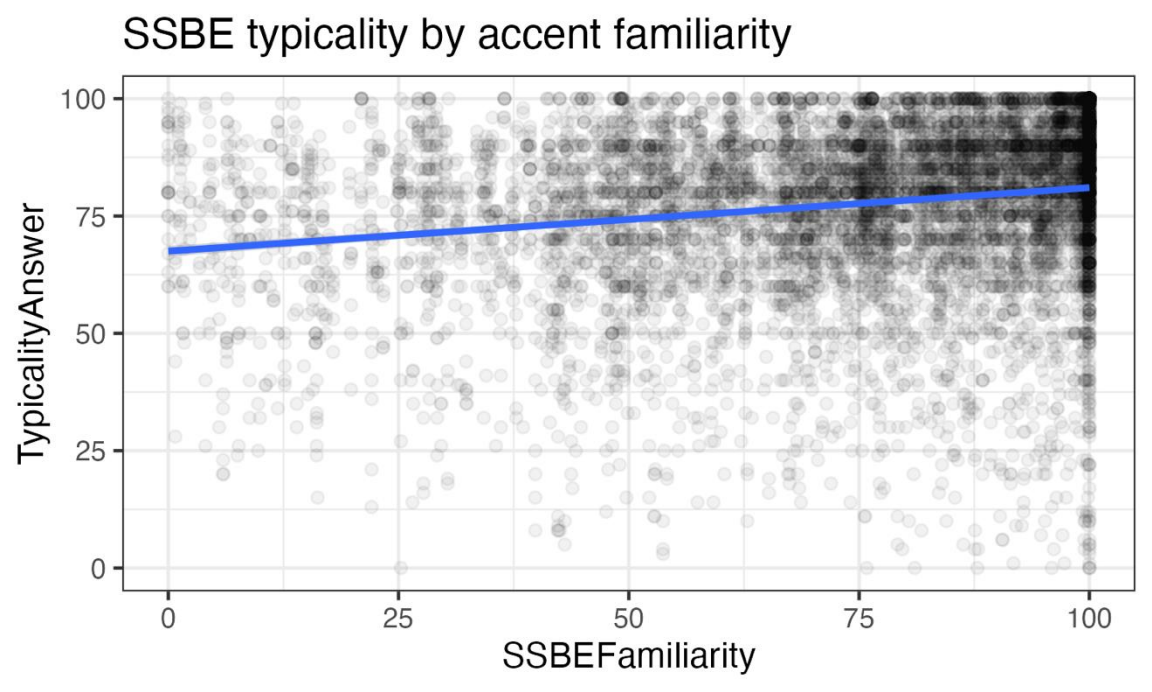
(Trudgill 1990: 77)



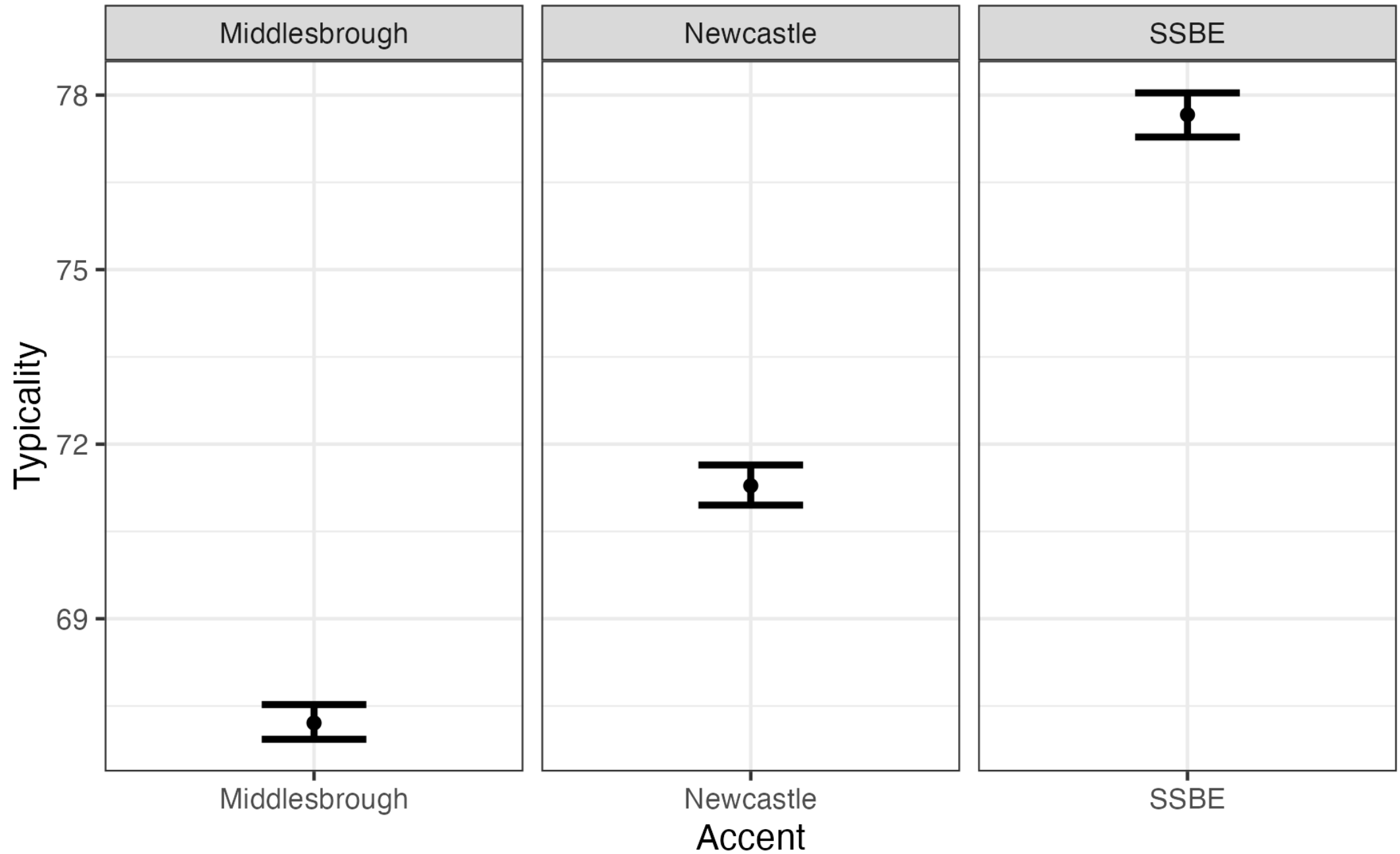
Research questions

- 1) What are the distributions of typicality responses for Newcastle and Middlesbrough accents compared with SSBE?
- 2) To what extent are typicality responses affected by listener familiarity with the accent?
- 3) To what extent are typicality responses affected by the presence/absence of salient phonetic features in the non-standard accents?

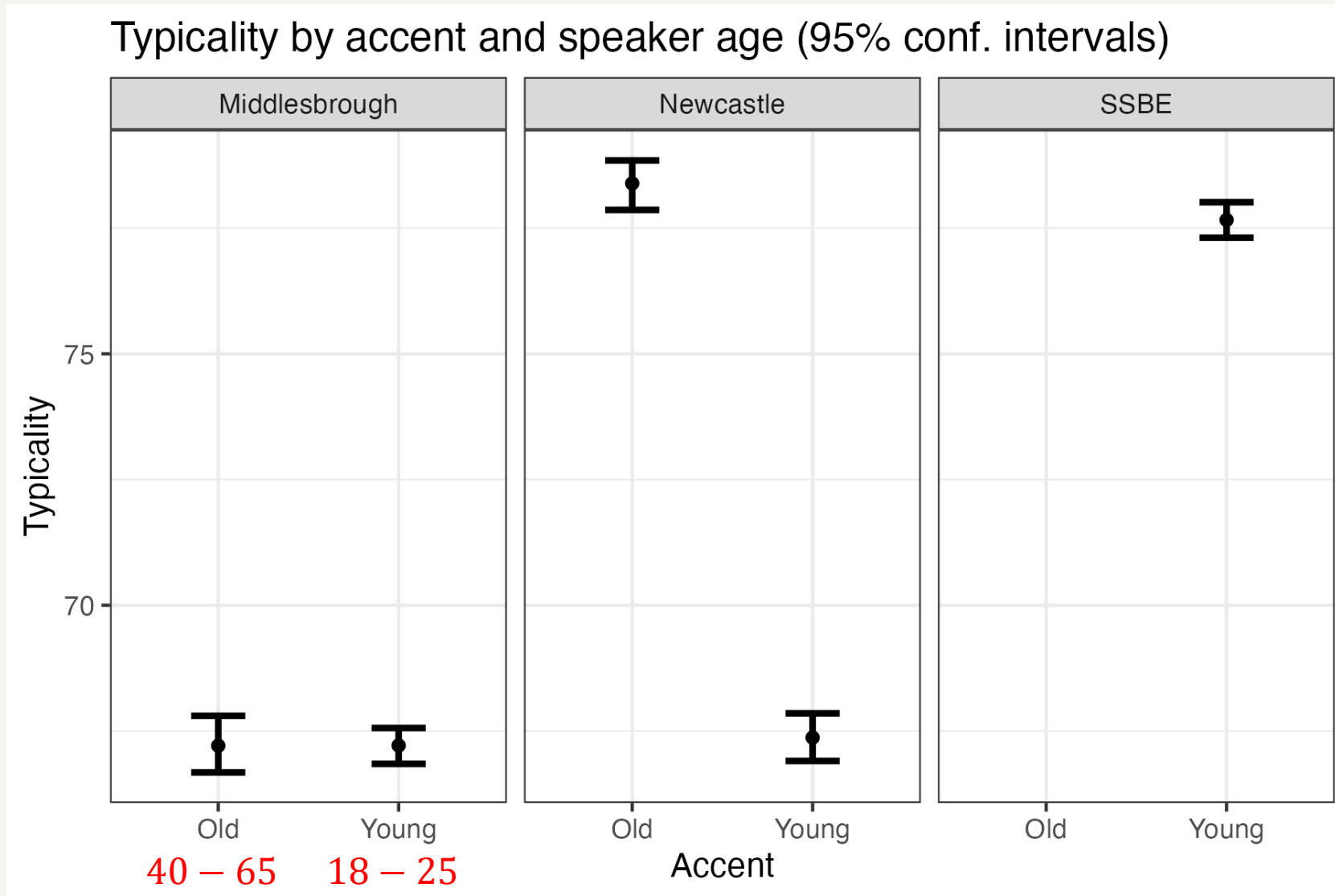
Typicality vs. familiarity



Typicality by accent (95% conf. intervals)



Typicality by Speaker Accent and Age group



Rating accentedness of non-standard stimuli

Carmen Llamas (native of Middlesbrough) rated the North East samples on 1 -4 scale for different types of salient *accentedness*

- **Regional** (features found throughout North East)
- **Localized** (features differentiating Middlesbrough & Newcastle)
- **“Overall”**
 - Holistic impression that includes both NE features and the localised features, and in some cases other non-standard features (usually grammatical) that may be found throughout BrE
 - Takes into account things like position and stress of the accent feature in the utterance that may contribute to the perception of the overall accentedness

Example linguistic features

Regional features – Common to all of North East

- glottalised intervocalic /p,t,k/
 - e.g. [p^he:p^ʔə] *paper*, [æp^ʔi] *happy* 
- t-to-r
 - e.g. [bʊ^ɹafil] *but I feel* 
- pre-aspirated non-intervocalic /p,t,k/
 - e.g. [bæ^ht] *bat*, [ʃɔ^hps] 
- *aye* (yes), *hoy* (throw)

Example linguistic features

Localized features: Found in NC or M'b but not the other

NURSE

Newcastle = $\emptyset$:  or $\text{ɔ}:$ 

Middlesbrough = ɛ 

START

Newcastle = $\text{ɑ}:$ 

Middlesbrough = $\text{a}:$ 

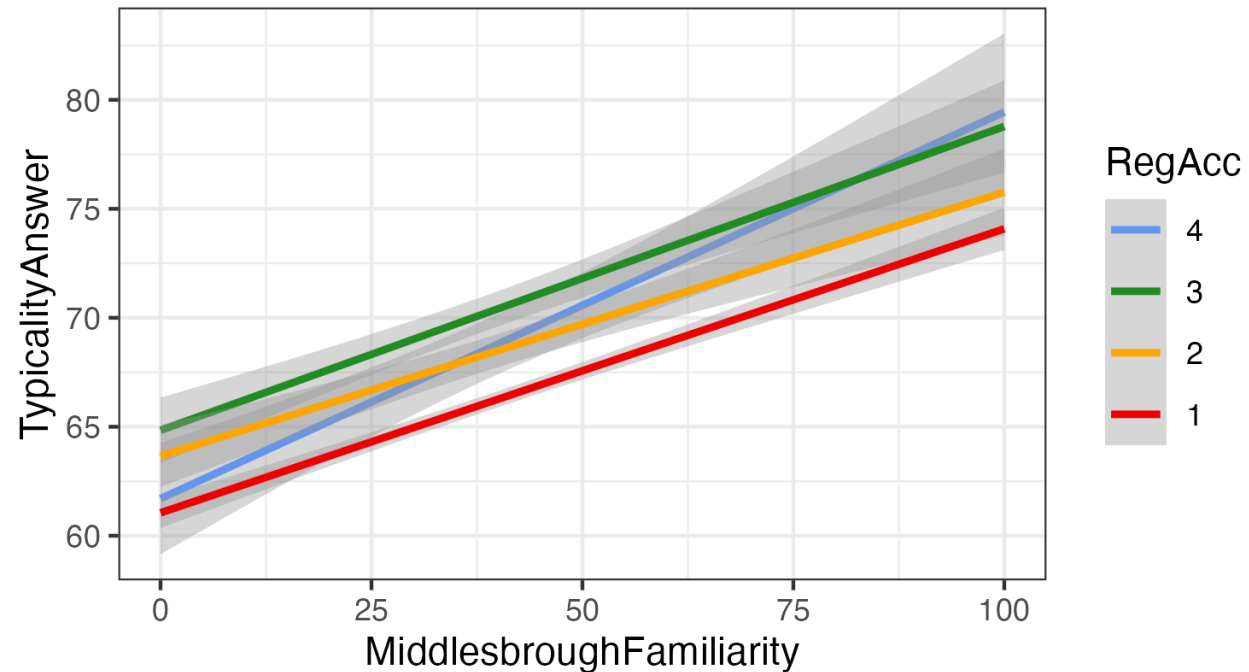
lettER

Newcastle = ɐ 

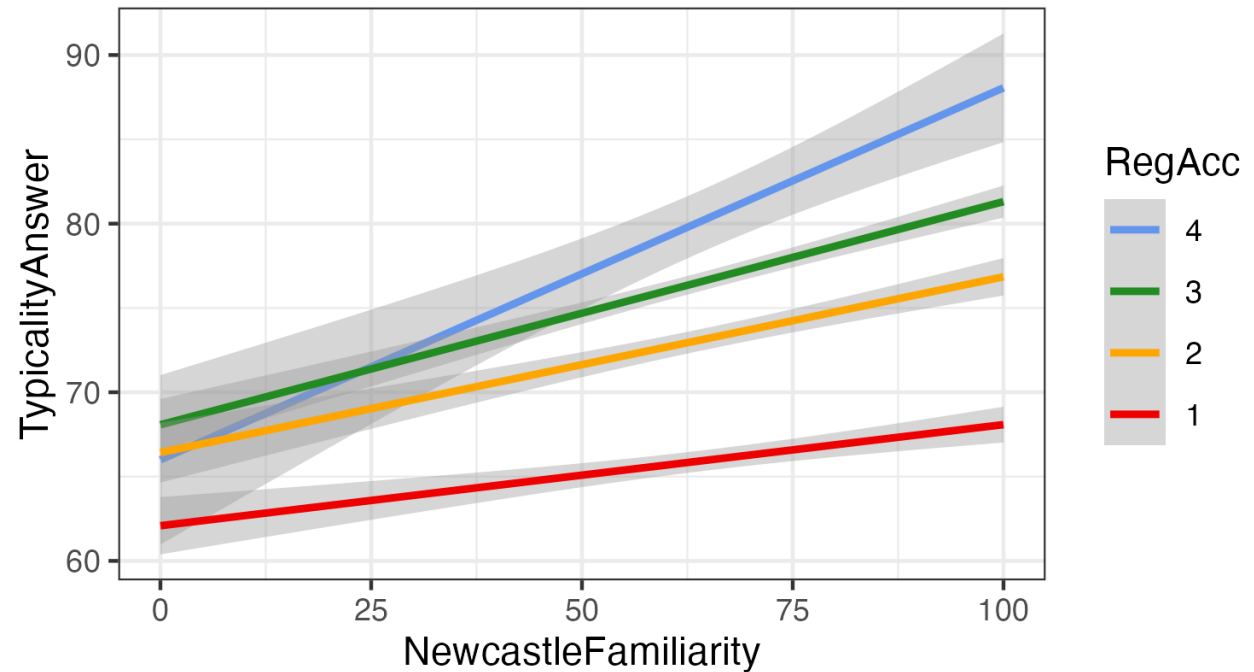
Middlesbrough = ɛ 

Regional accentedness vs. familiarity

M'b typicality by accent familiarity

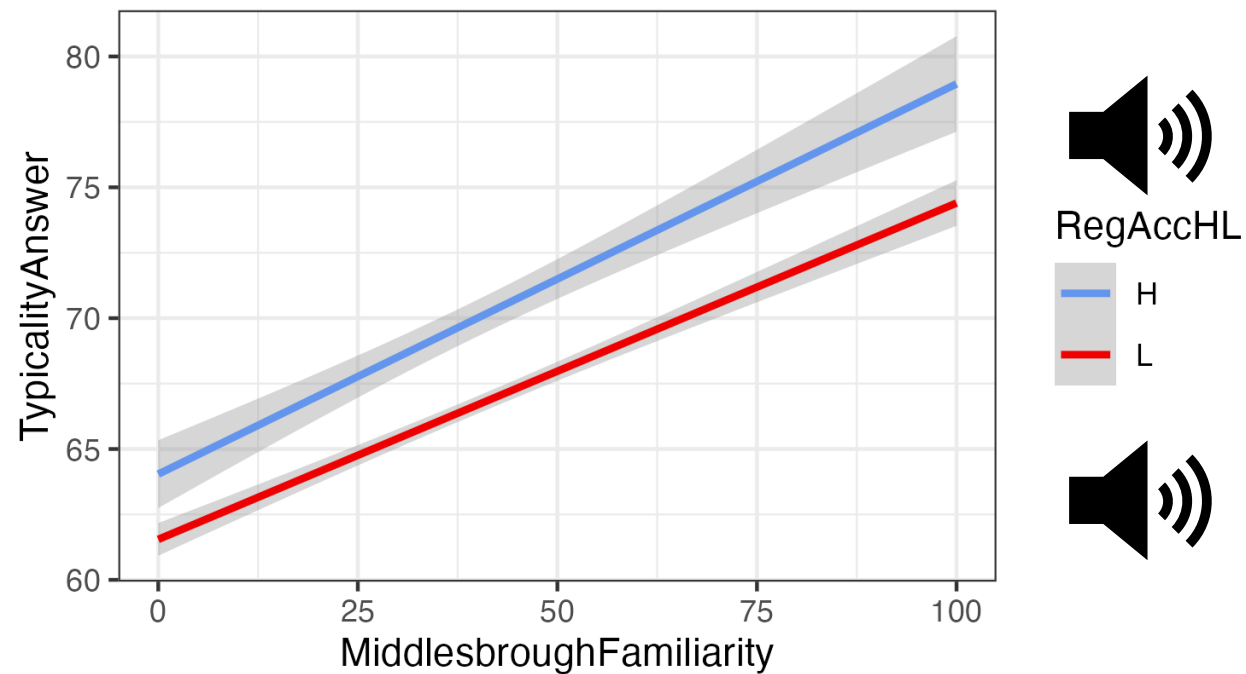


N'c typicality by accent familiarity



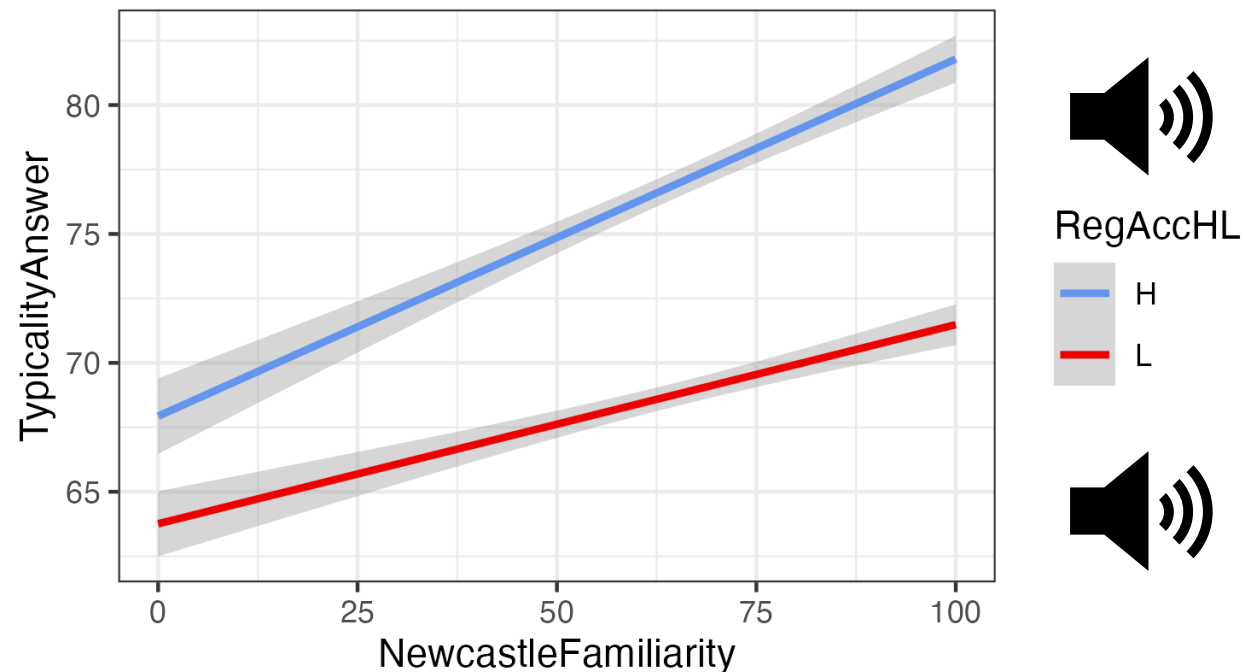
Regional accentedness vs. familiarity

M'b typicality by accent familiarity



Familiarity effect: $p < 0.0001$
RegAccHL effect: $p = 0.005$
Familiarity * OverallAccHL: $p = 0.02$

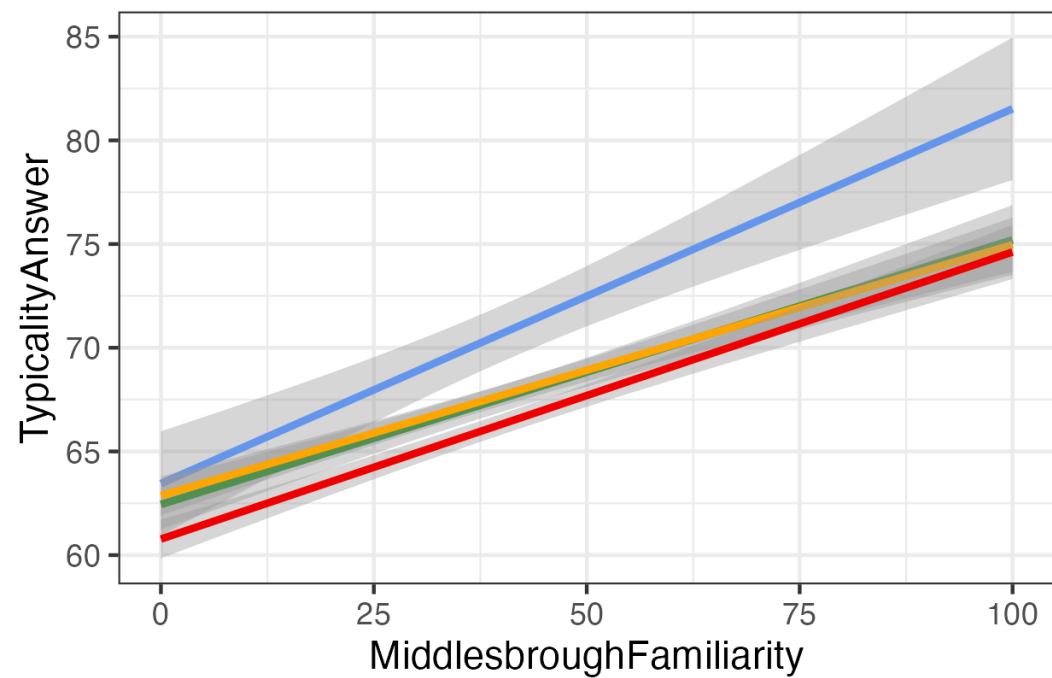
N'c typicality by accent familiarity



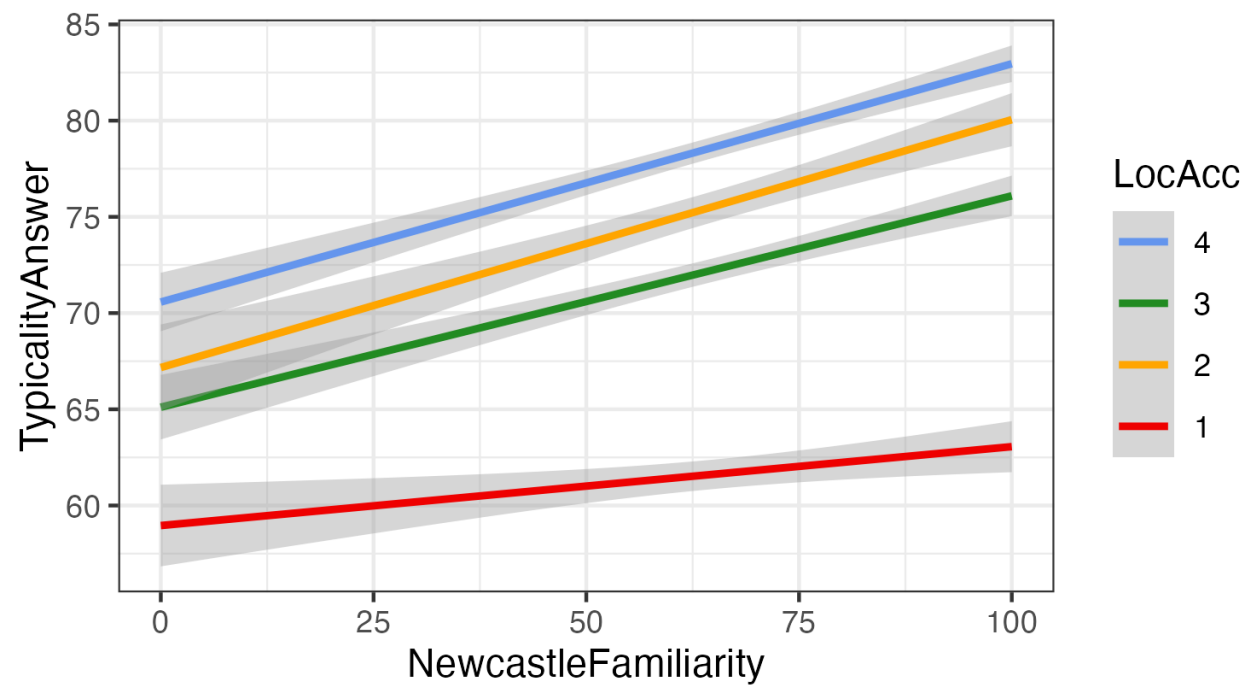
Familiarity effect: $p < 0.0001$
OverallAccHL effect: $p = 0.046$
Familiarity * OverallAccHL: $p < 0.0001$

Local accentedness vs. familiarity

M'b typicality by accent familiarity

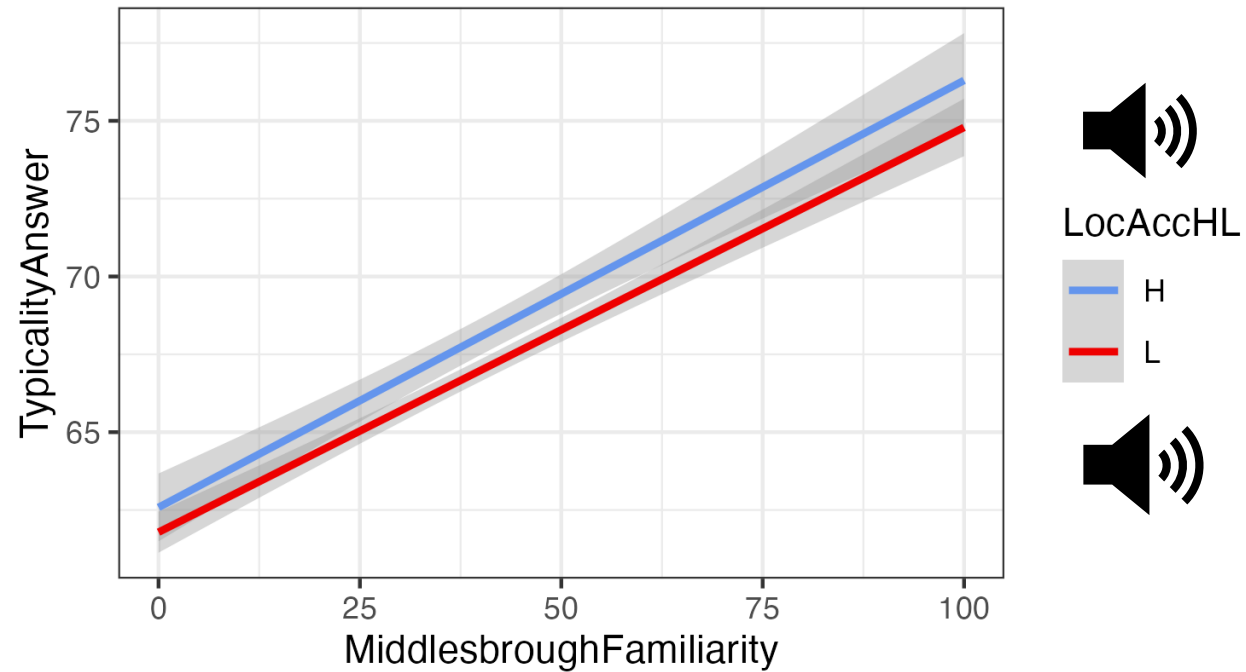


N'c typicality by accent familiarity



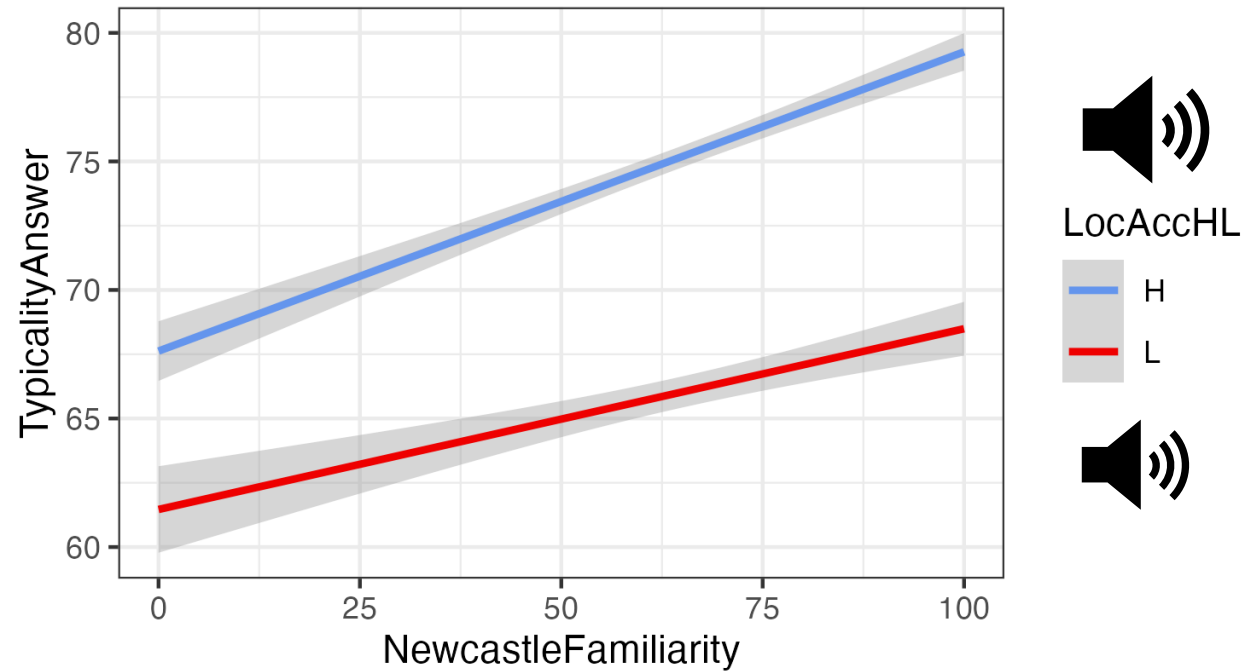
Local accentedness vs. familiarity

M'b typicality by accent familiarity



Familiarity effect: $p < 0.0001$
LocAccHL effect: $p = 0.09$ / $p = 0.51$ (no overall HL effect)
Familiarity * OverallAccHL: $p = 0.41$ (no slope difference)

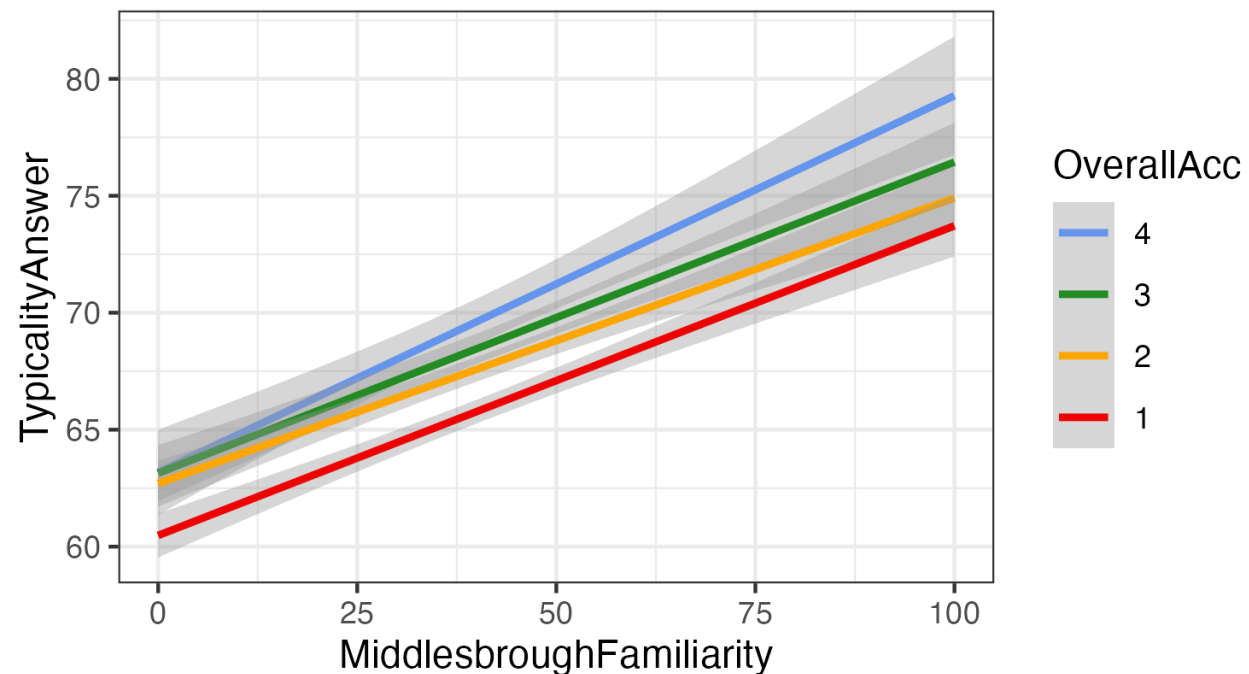
N'c typicality by accent familiarity



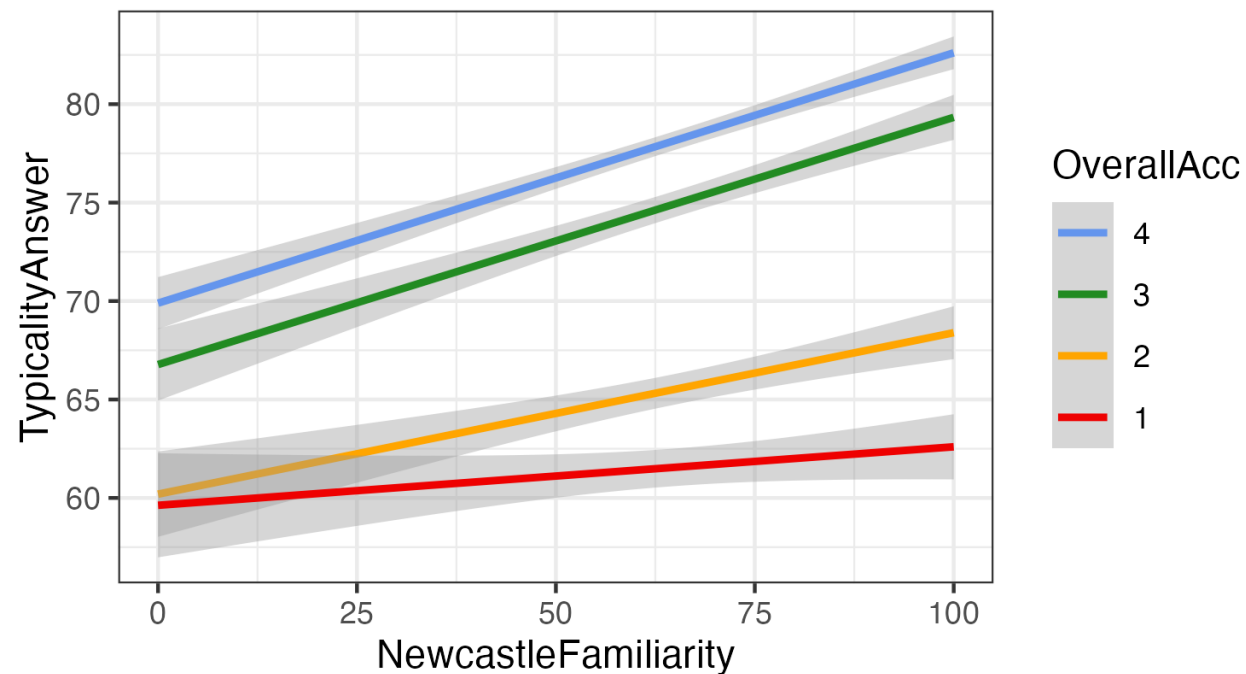
Familiarity effect: $p < 0.0001$
OverallAccHL effect: $p = 0.009$
Familiarity * OverallAccHL: $p < 0.0001$

Overall accentedness vs. familiarity

M'b typicality by accent familiarity

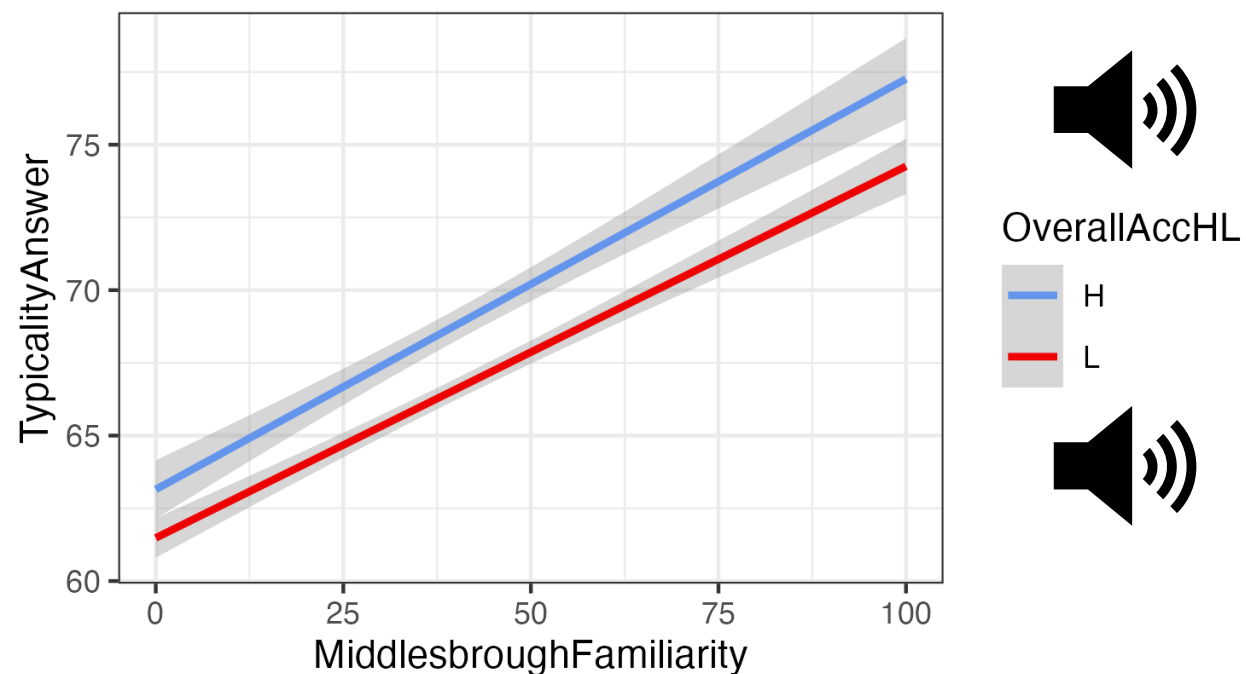


N'c typicality by accent familiarity



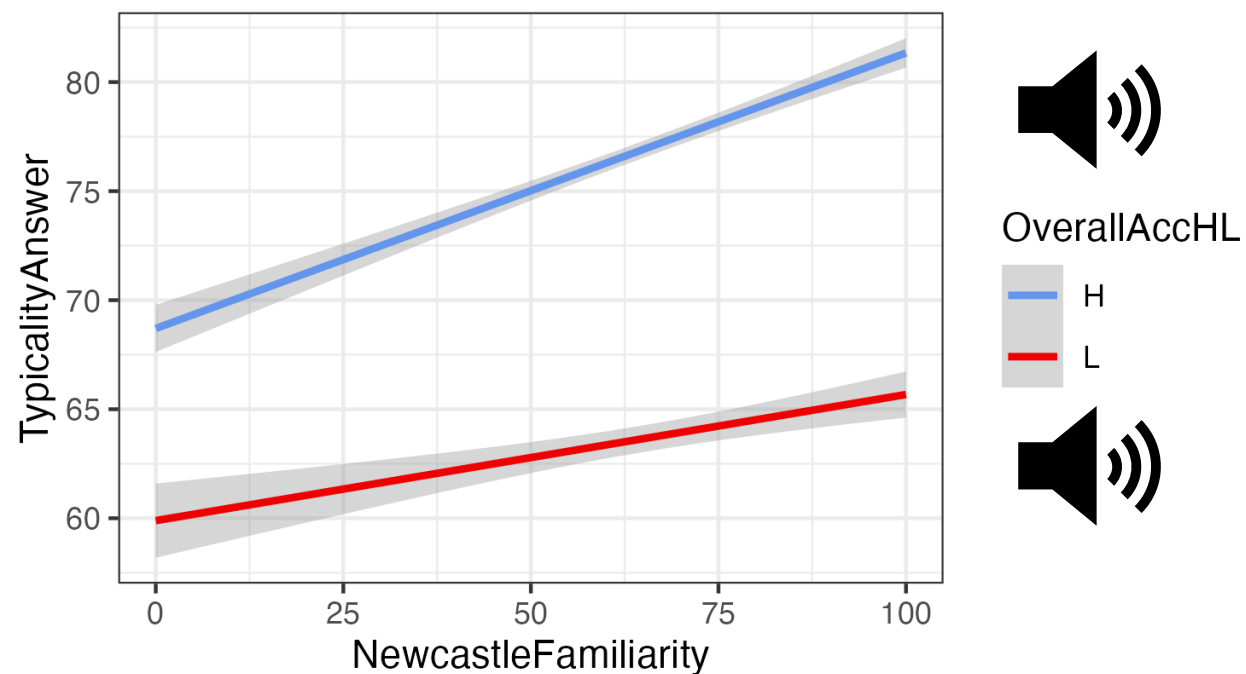
Overall accentedness vs. familiarity

M'b typicality by accent familiarity

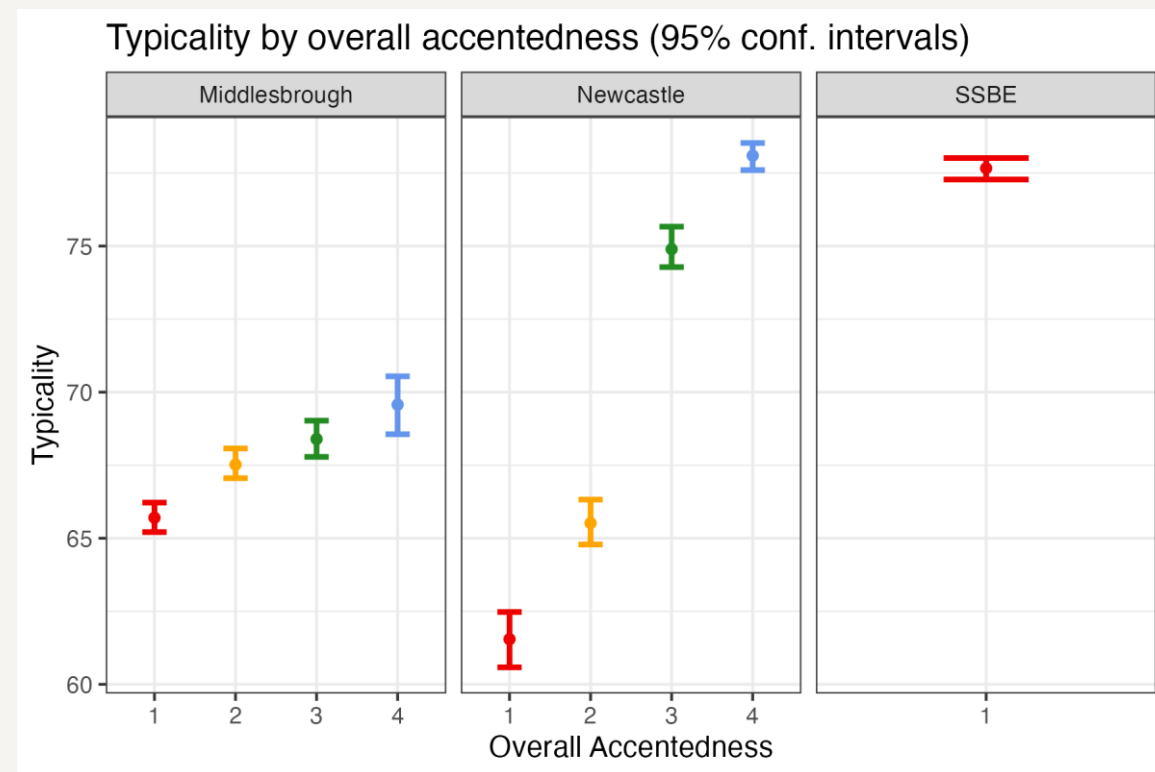
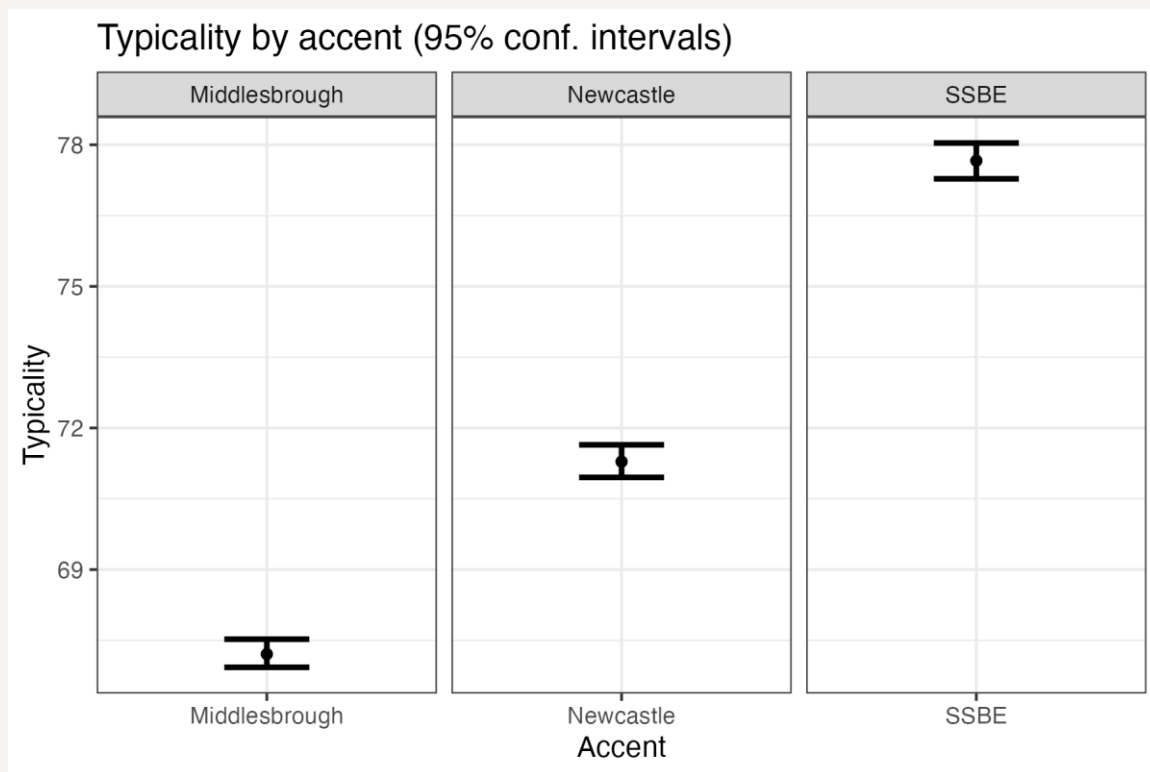


Familiarity effect: $p < 0.0001$
OverallAccHL effect: $p = 0.06$ (no H/L difference)
Familiarity * OverallAccHL: $p = 0.07$ (no slope difference)

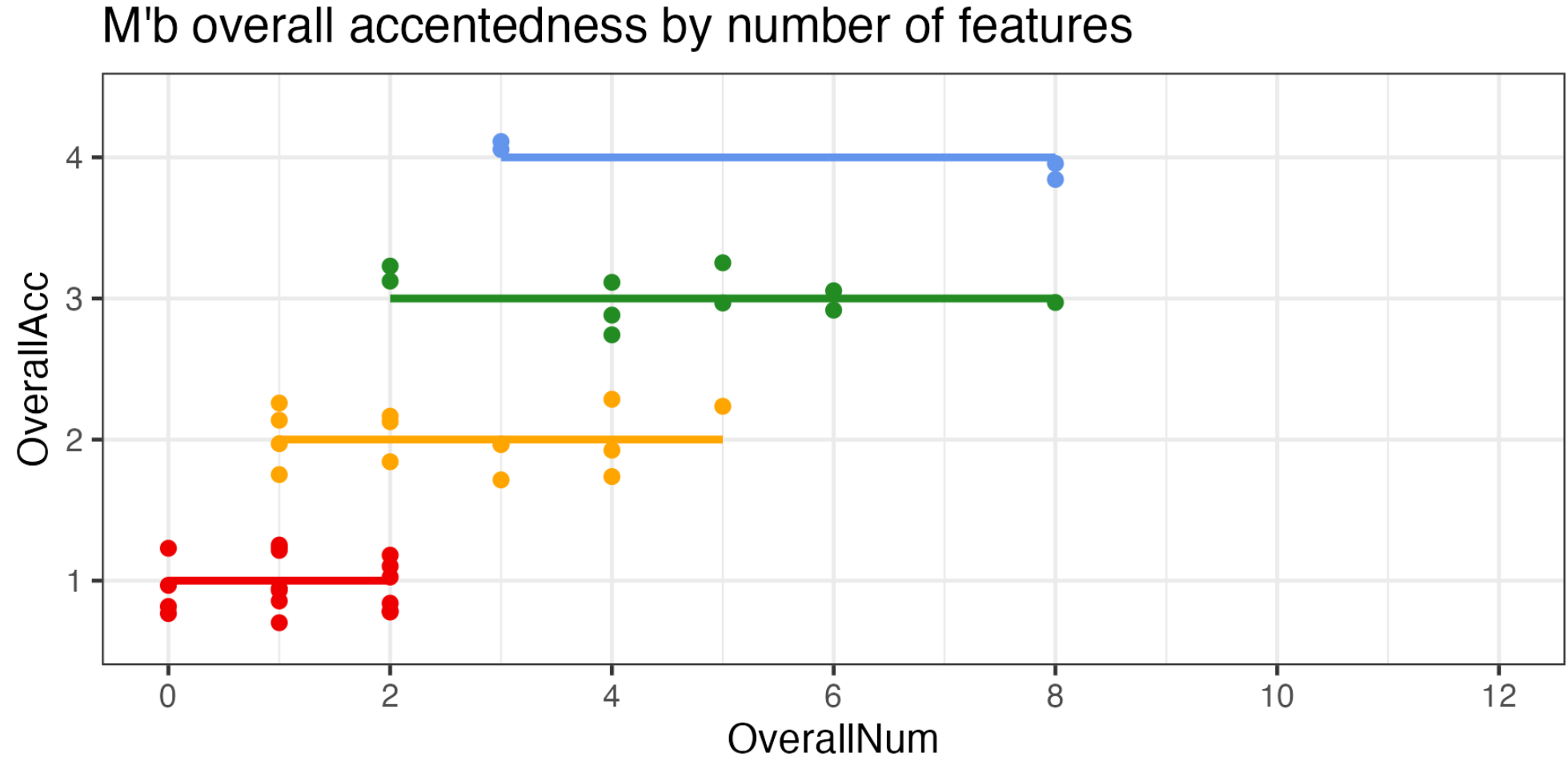
N'c typicality by accent familiarity



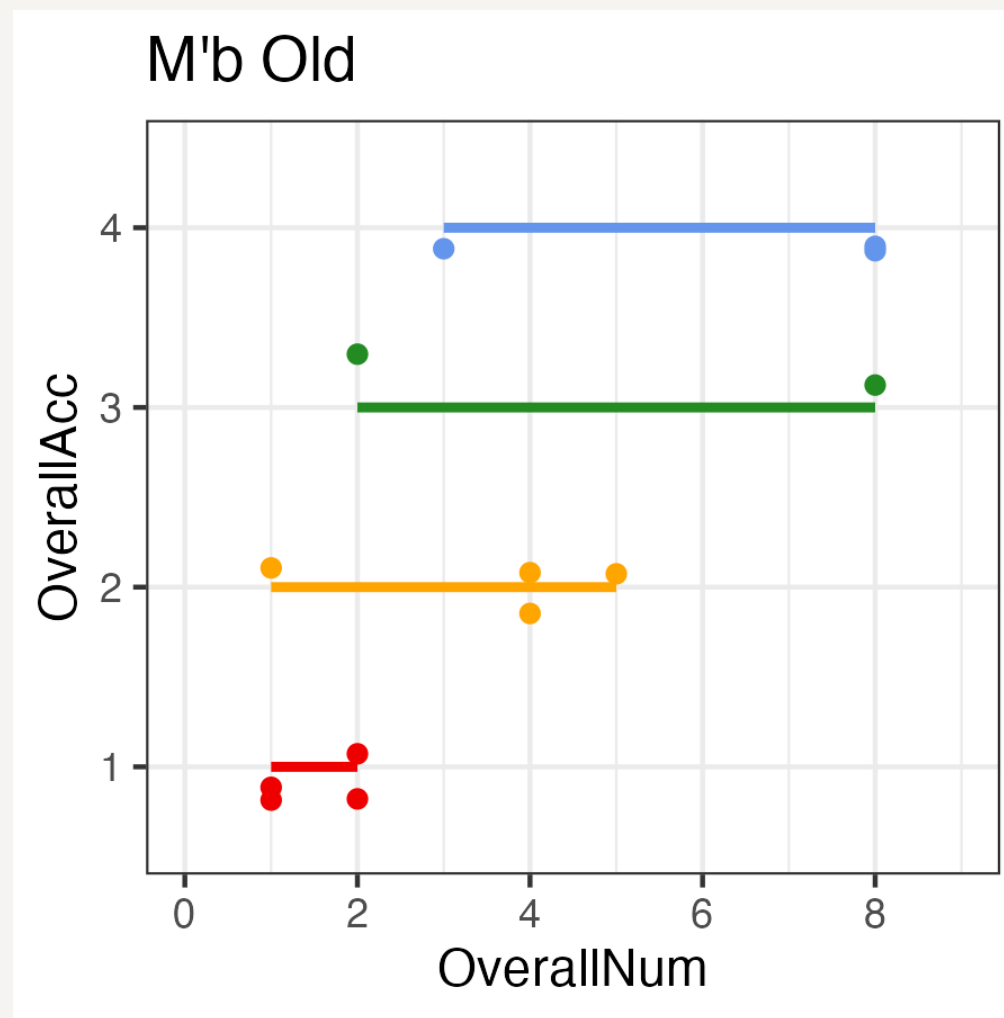
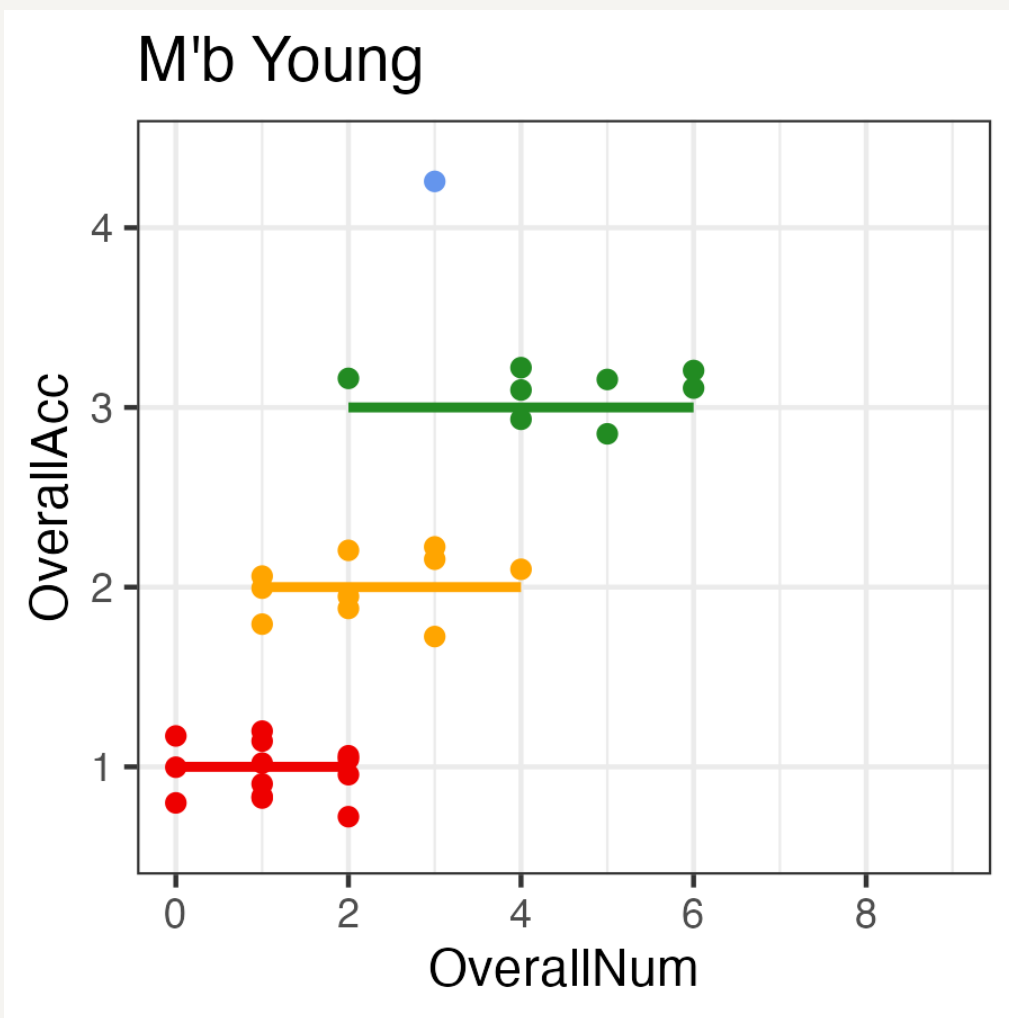
Familiarity effect: $p < 0.0001$
OverallAccHL effect: $p < 0.0001$
Familiarity * OverallAccHL: $p < 0.0001$



Number of features – Middlesbrough

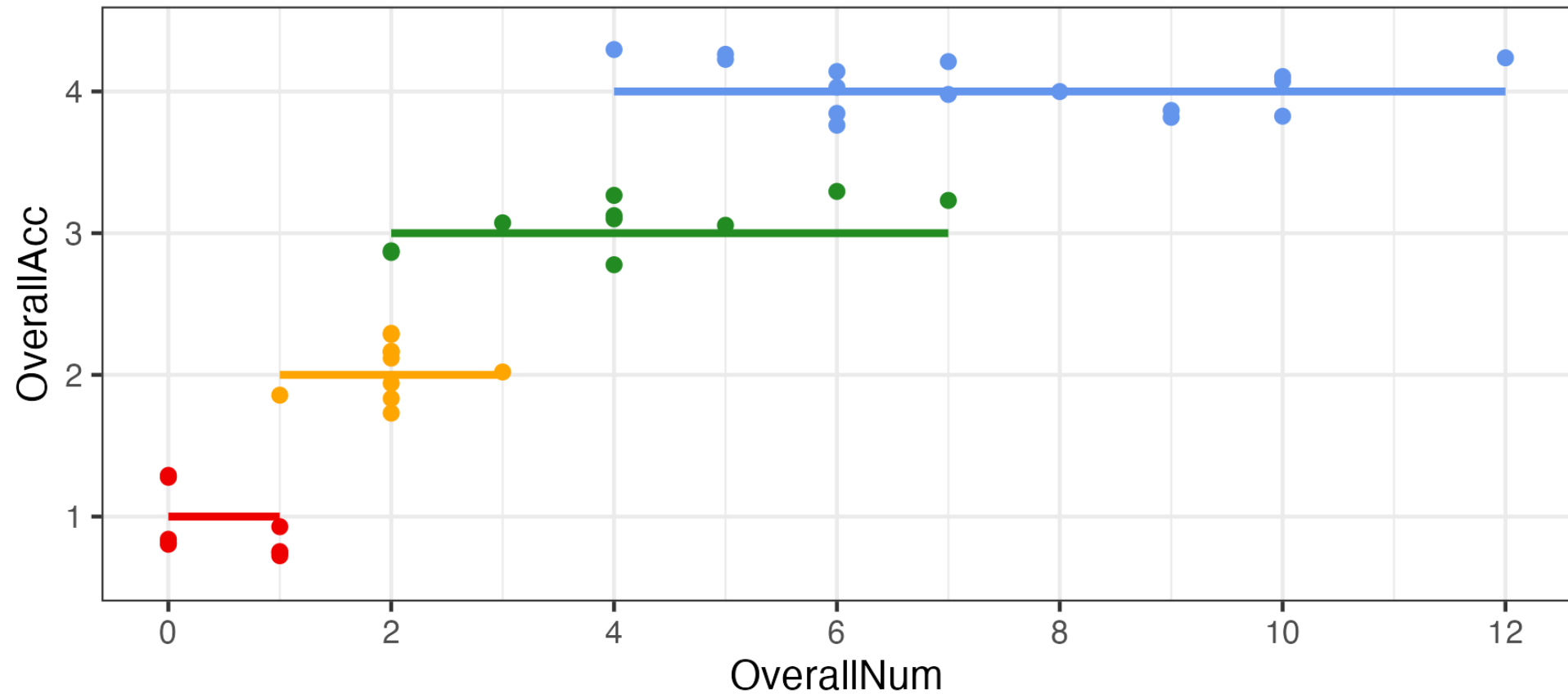


Number of features – Middlesbrough

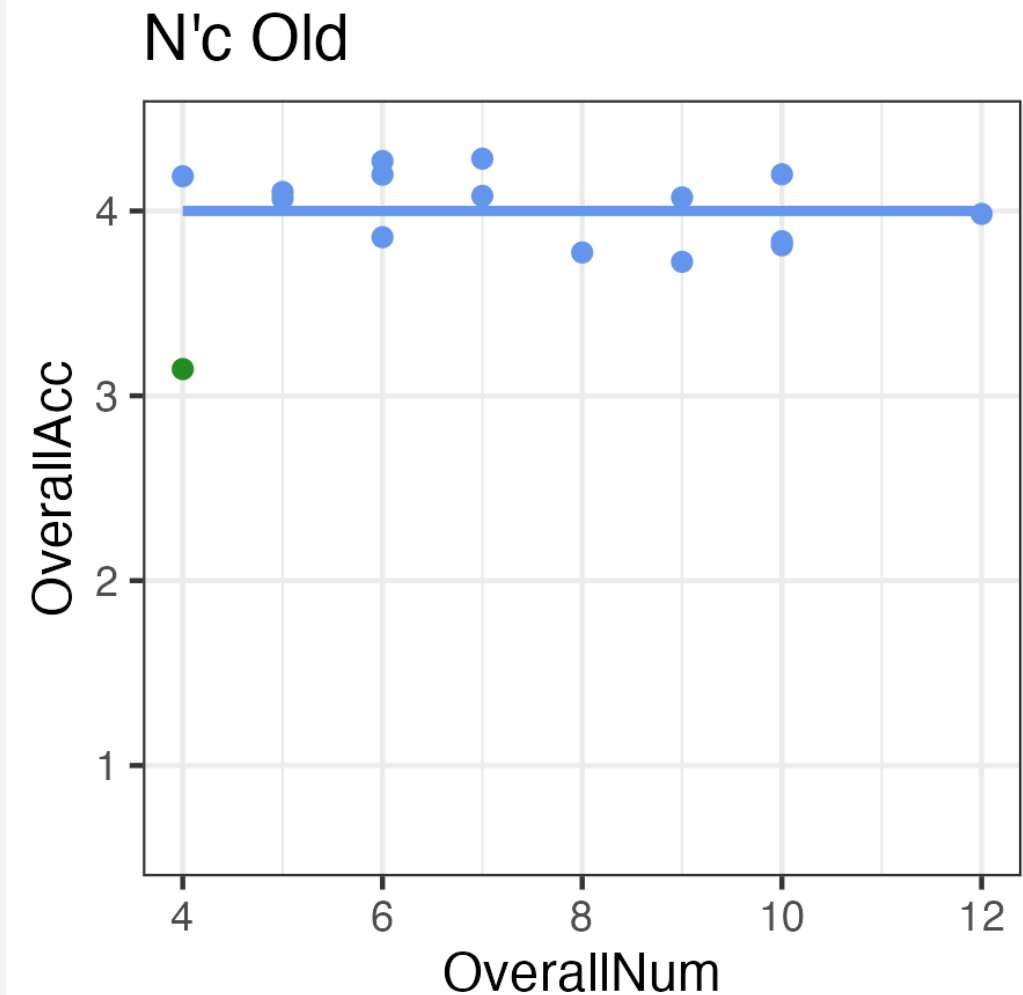
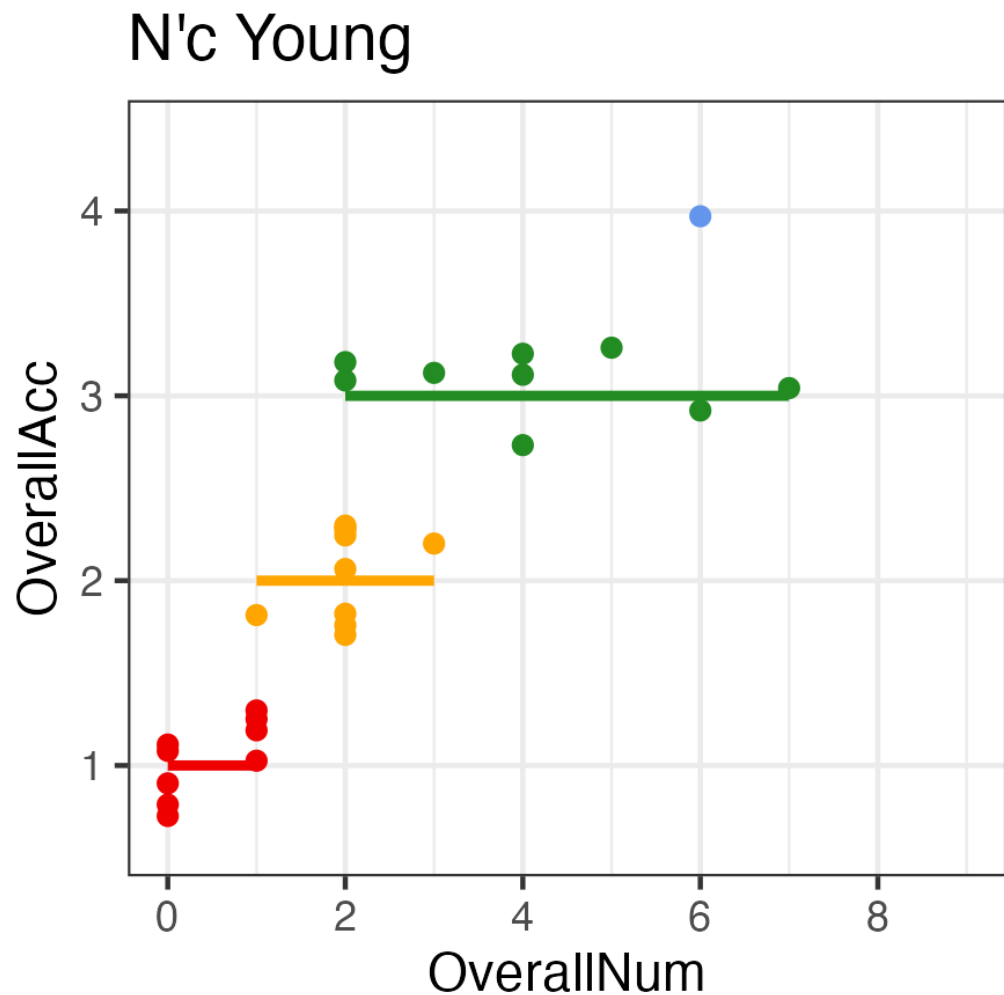


Number of features – Newcastle

N'c overall accentedness by number of features



Number of features – Newcastle



Summary: North East

- Number of features has clear relation with accentedness rating, but not a one-to-one mapping
- Clear difference in the number of features in N'c samples compared with M'b
- Helps account for differences in the perception of typicality (N'c higher rates than M'b)
- For non-standard accents, the **more dialect features** that are noticed, the **more typical** they are perceived to be
- **Familiarity also a factor**
 - Cultural dominance of N'c in the North-East contributes to general familiarity of accent and dialect nationally
 - Greater general awareness of features associated with N'c than M'b, hence different responses to high- vs. low-accentedness samples for N'c, but not for M'b

Thoughts on SSBE

- Very homogeneous group of speakers (Cambridge undergrads)
- Only one sample judged to have a noticeable feature – and it has the *lowest* typicality rating!
- What we're thinking about: For a standard accent, the **fewer features** that are noticed, the **more typical** they are perceived to be

Overall conclusions

- We have to notice features in speech in order to make an assessment of typicality
- The *more features* noticed in *regional accents*, the *more typical* they are perceived
- The *fewer features* noticed in the *standard accent*, the *more typical* it is perceived
- But – we're still figuring out exact links with theory, so please send our way any thoughts or citations we should be aware of!

Thank you!