



UNIVERSITY OF LEEDS

Using eye-tracking to investigate the role of regional variation in spoken word recognition

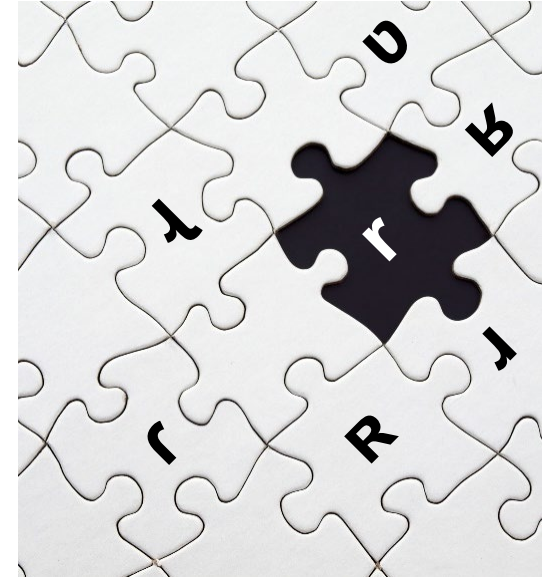
Gisela Tomé Lourido (University of Leeds)

with Robert Lennon (Lancaster University) & Bronwen G. Evans (UCL)

Lancaster University, 9 February 2021

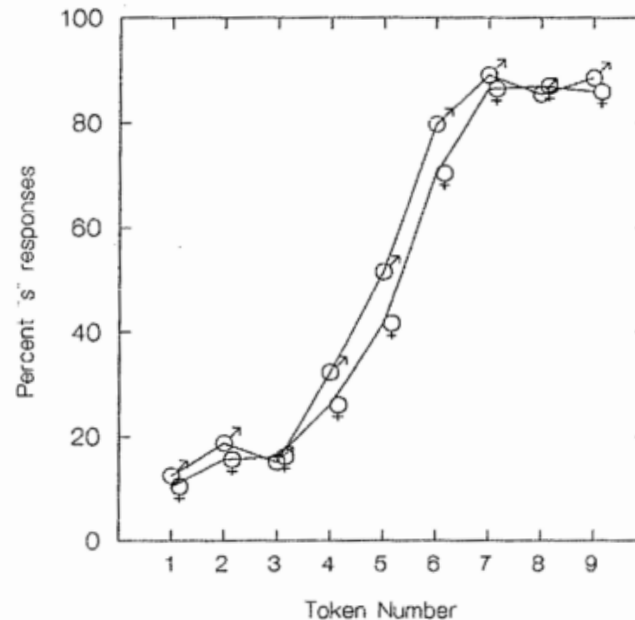
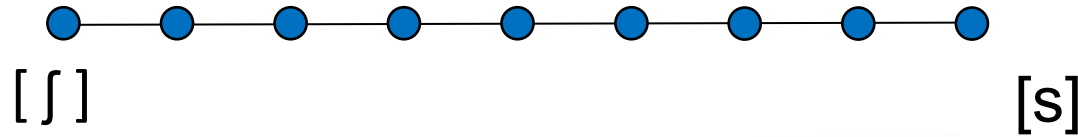
Background: Processing of variation

- Early models of speech processing considered variability a “problem” listeners have to solve.
- Hearing speech produced in an unfamiliar accent has a processing cost, Adank et al. 2009; Floccia, et al. 2006.
- Listeners can rapidly adapt to novel talkers and accents, e.g., Bradlow & Bent, 2007.
- When listening to a familiar accent, **perceived information about the speaker** has been shown to affect low-level speech perception (e.g., Strand, 1999) and lexical access (e.g., Koops et al., 2008), arguably facilitating processing.



Background: Strand (1999)

- Visual stimuli: male and female faces
- Auditory stimuli: words, e.g., *ship* and *sip*



Background: Explicit vs implicit cues

These experiments often use pictures or words to cue a specific social category (e.g., gender, age, region) explicitly.

It is unclear whether **brief exposure to accent-specific phonetic features in the speaker's speech alone** would also influence speech processing.

Does information about the speaker's accent embedded in the speech signal affect the time course of spoken word recognition?

Background: Phonetic variables

TRAP-BATH



Word class	<u>North</u>	<u>South</u>
<i>pack, maps, fact, trap</i>	æ	æ
<i>pass, laugh, bath, ask</i>		ɑː
<i>park, heart, marks, calm</i>	ɑː	

FOOT-STRUT



Word class	<u>North</u>	<u>South</u>
<i>book, cook, bush, foot</i>	ʊ	ʊ
<i>buck, hug, cut, strut</i>		ʌ



Leeds.
Thanks to
Robert Lennon.

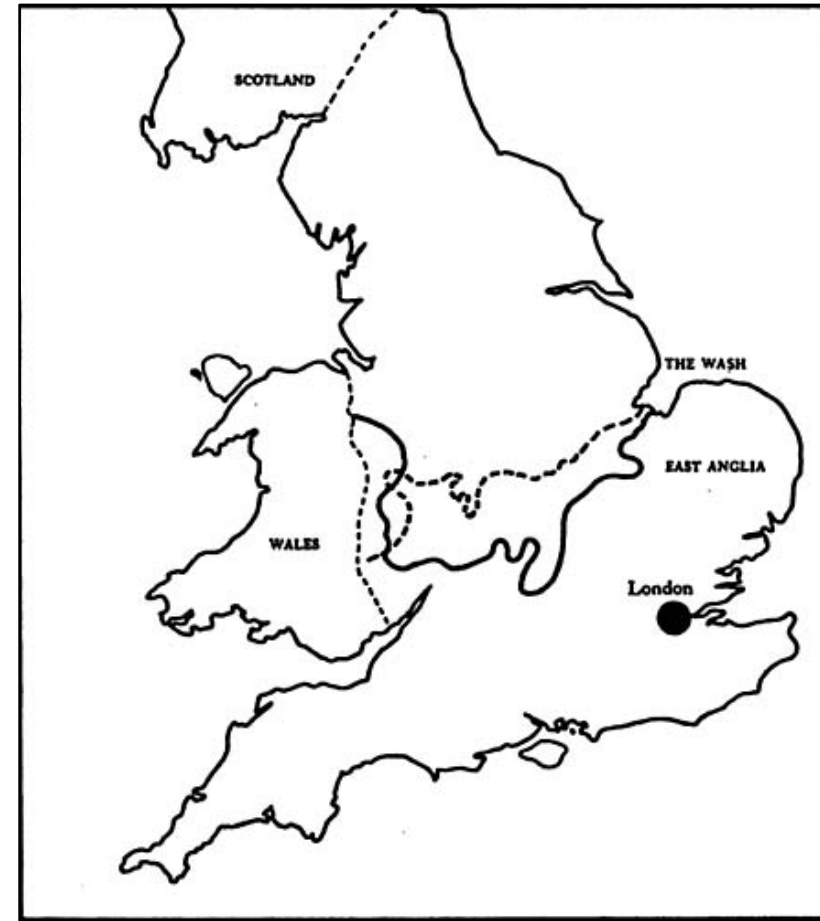
Background: Phonetic variables

TRAP-BATH

Word class	<u>North</u>	<u>South</u>
<i>pack, maps, fact, trap</i>	æ	æ
<i>pass, laugh, bath, ask</i>		ɑː
<i>park, heart, marks, calm</i>	ɑː	

FOOT-STRUT

Word class	<u>North</u>	<u>South</u>
<i>book, cook, bush, foot</i>	ʊ	ʊ
<i>buck, hug, cut, strut</i>		ʌ



Isoglosses for FOOT-STRUT (solid line), and TRAP-BATH (broken line). Adapted from Wells.
swphonetics.com/articulation/accents/sbe/

Background: Phonetic variables

TRAP-BATH

Word class	<u>North</u>	<u>South</u>
<i>pack, maps, fact, trap</i>	æ	æ
<i>pass, laugh, bath, ask</i>		ɑː
<i>park, heart, marks, calm</i>	ɑː	

Both categories exist,
different lexical distribution

FOOT-STRUT

Word class	<u>North</u>	<u>South</u>
<i>book, cook, bush, foot</i>	ʊ	ʊ
<i>buck, hug, cut, strut</i>		ʌ

STRUT does not exist in
Northern listeners'
phonological inventory

Background: L2 speech perception

When we hear a new category that doesn't exist in our phonological repertoire, the new category is 'assimilated' to the nearest native category, Best 1995

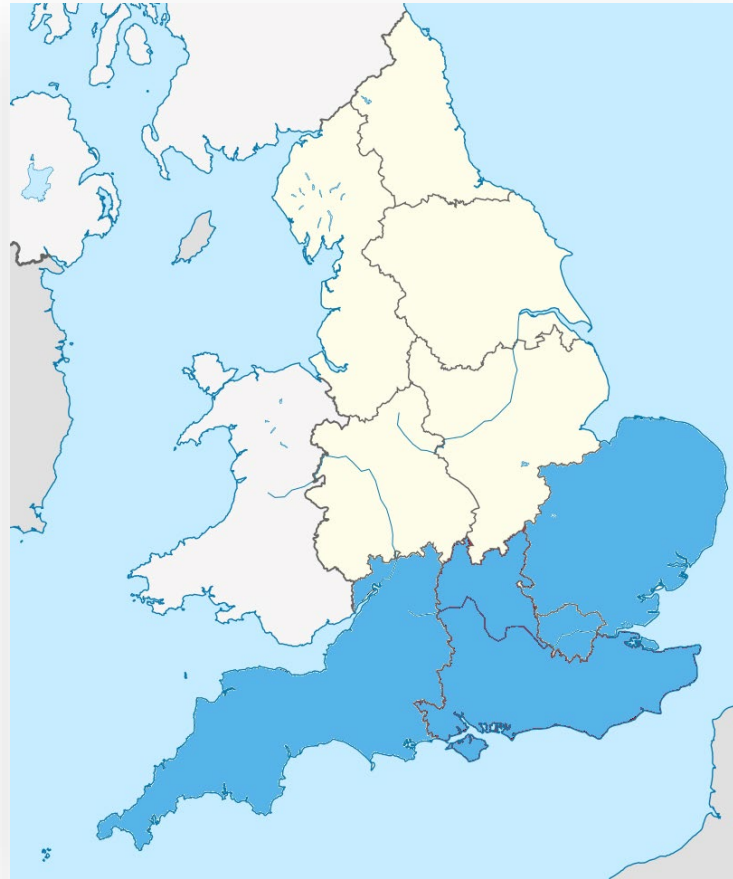
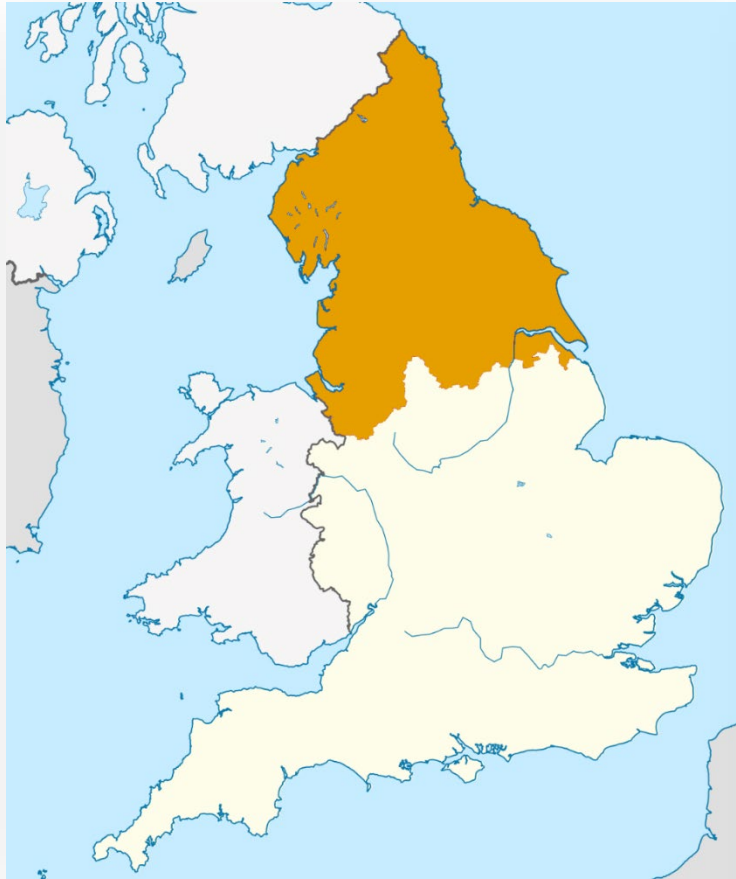
Words containing phonological contrasts which do not exist in the native phonetic repertoire create increased lexical competition for L2 learners, Weber & Cutler, 2004

FOOT-STRUT

Word class	<u>North</u>	<u>South</u>
<i>book, cook, bush, foot</i>	ʊ	ʊ
<i>buck, hug, cut, strut</i>		ʌ

STRUT does not exist in Northern listeners' phonological inventory

The current study: Design



2x2 design

- 2 speaker accents: Leeds, SSBE
- 2 listener groups: northern, southern English listeners

Northern listeners as experimental group

- SSBE familiar, standard variety

Eye-tracking task

The current study: Predictions

Do northern listeners differ from southern listeners when processing SSBE and Leeds accents?

TRAP-BATH

Northern listeners will be able to use the TRAP-BATH distinction in SSBE to disambiguate the pairs of words earlier

→ they will look at the target earlier in the SSBE condition.

FOOT-STRUT

The FOOT-STRUT distinction in SSBE will not necessarily help northern listeners disambiguate the words earlier, as this is not a native contrast

→ both accent conditions will be similar.

Method

Method: Participants

61 listeners were tested

11 were excluded due to language background

9 were excluded due to technical problems

The remaining **41 participants**

- were monolingual in English
- classified in two groups

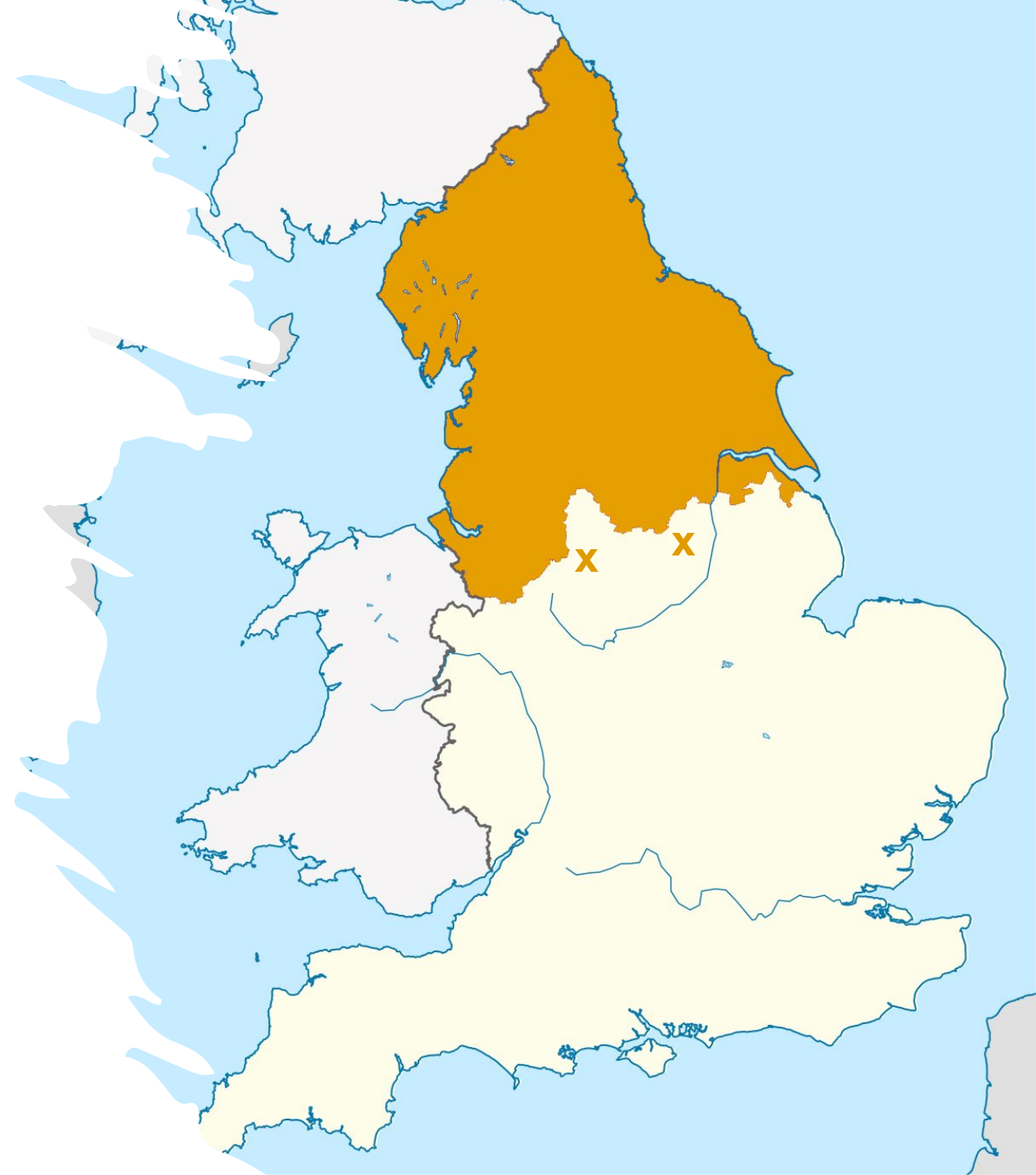
Northern listeners
(N = 24)

Southern listeners
(N = 17)

Method: Participants

Northern listeners

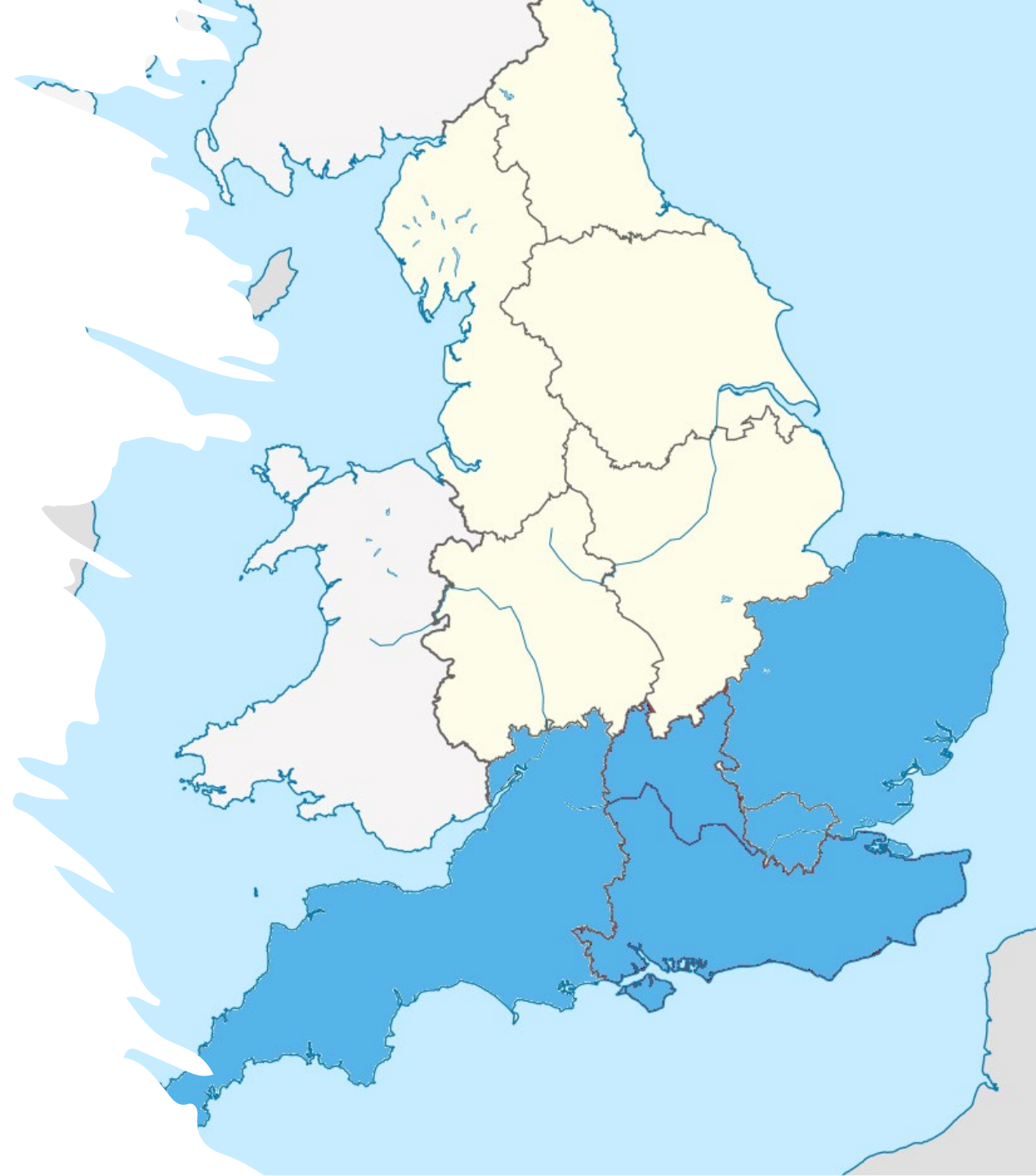
- were 18-44 years old, median= 24
- 17 f, 7m
- were raised in the North of England
- had not lived elsewhere for more than 8 months
- had parents who were monolingual and raised in the North of England



Method: Participants

Southern listeners

- were 18-25 years old, median= 19
- 11 f, 6m
- were raised in the South of England
- not lived elsewhere for more than 8 months
- had been in Leeds for less than 3 months
- had parents who were monolingual and raised in the South of England



Method: Audio stimuli

Naturally-produced words
recorded by 2 Leeds & 2
SSBE speakers

Embedded in carrier

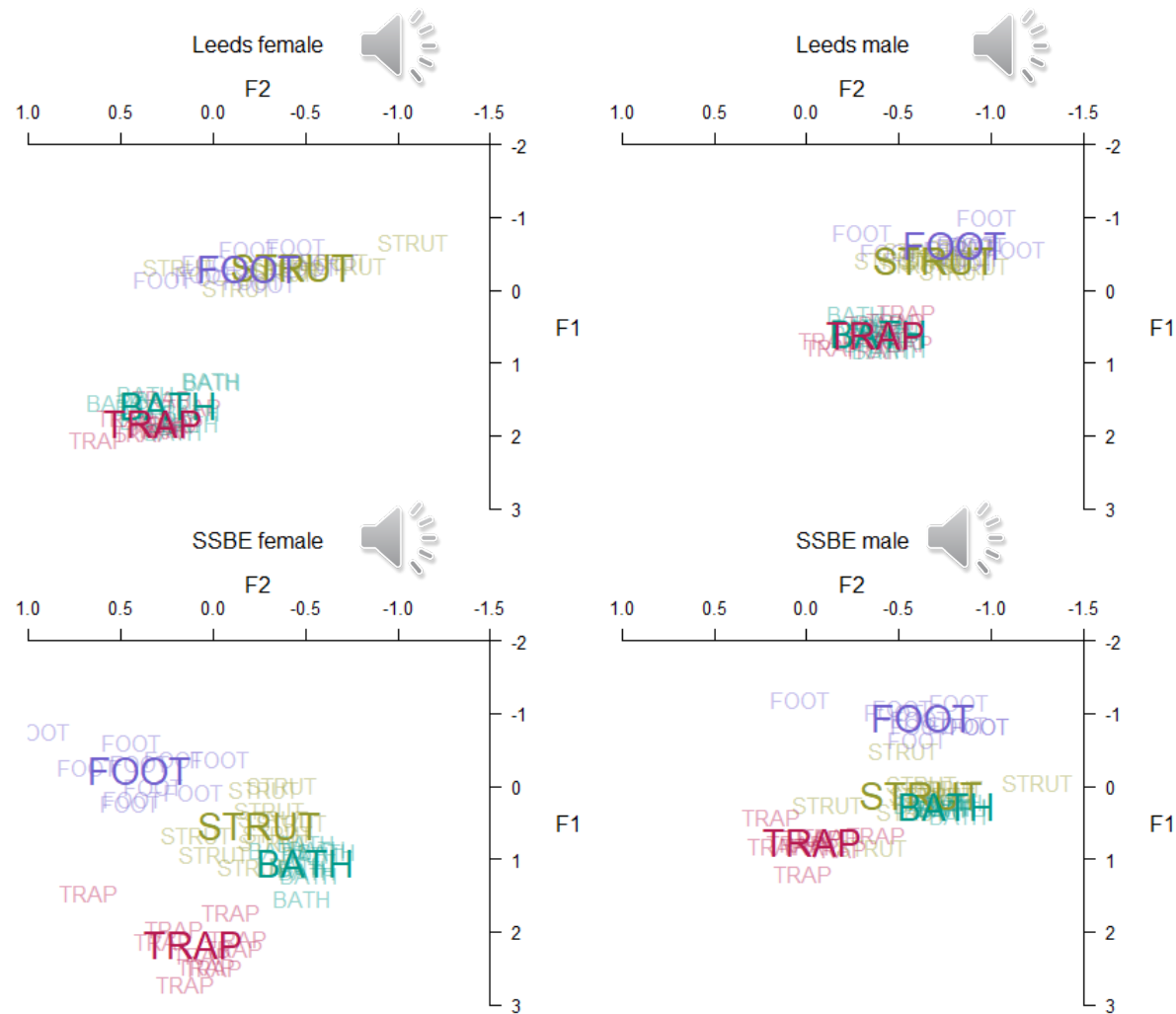
Evans & Iverson, 2004

Leeds accent:

/aɪm 'æskɪŋ ju tə 'ækses/

SSBE accent:

/aɪm 'ɑːskɪŋ ju tə 'ækses/



Method: Visual stimuli (following Best et al, 2013)

- Visual World Paradigm, Tanenhaus et al., 1995
- 2 printed words per trial (target-competitor)
- Words were CVC, CVCC, CVCV, CVCVC, CVCCVC
- Not semantically related

20 Test sets

10 TRAP-BATH contrasts

10 FOOT-STRUT contrasts

Controlled for frequency

20 Filler sets

Acoustically similar contrasts,
e.g., DRESS-KIT, LOT-THOUGHT

BUS + BOOK

PATH + PACK

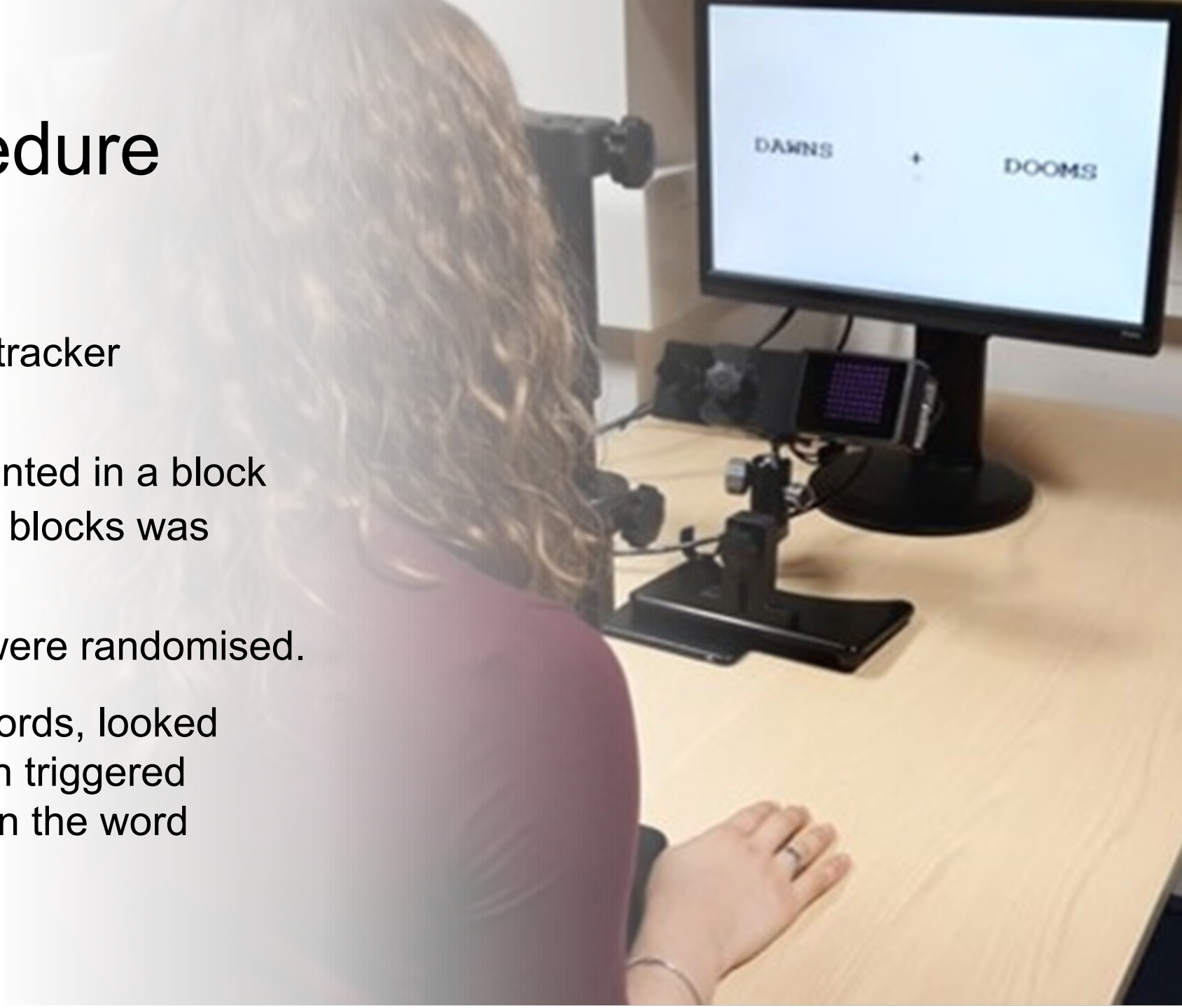
LEG + LID

POP + PORT

Method: Procedure

Eye-tracking task

- Eyelink 1000 Plus eye-tracker (500 Hz sampling rate).
- Each accent was presented in a block and the presentation of blocks was counterbalanced.
- Trials within the block were randomised.
- Participants read the words, looked at a fixation cross which triggered the audio and clicked on the word they heard.



Method: Procedure

Language Background Questionnaire

Do the following pairs of words rhyme when you say them?

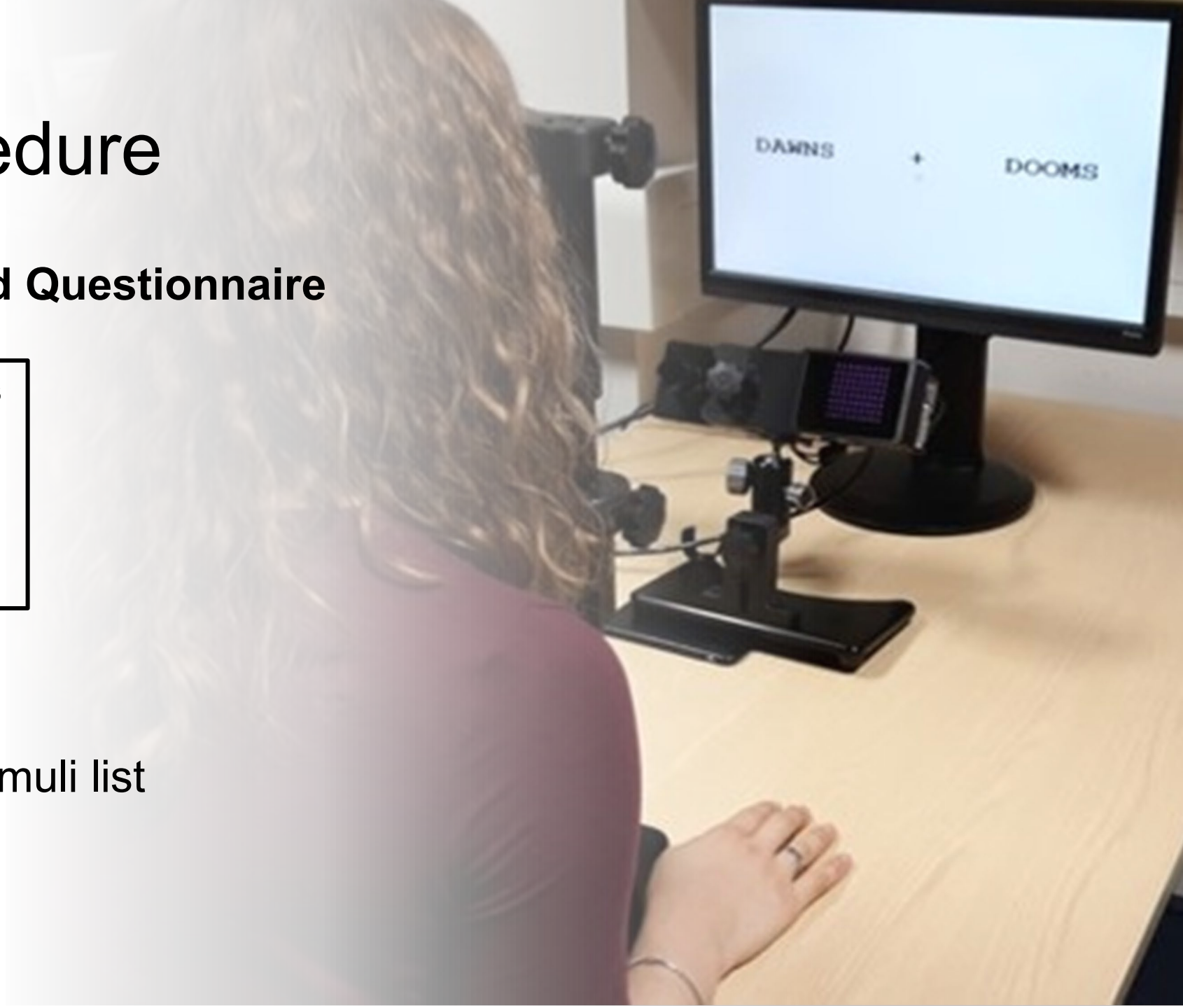
- *cut* *put*
- *bath* *math*
- *cart* *cat*

Production data

Participants read the stimuli list

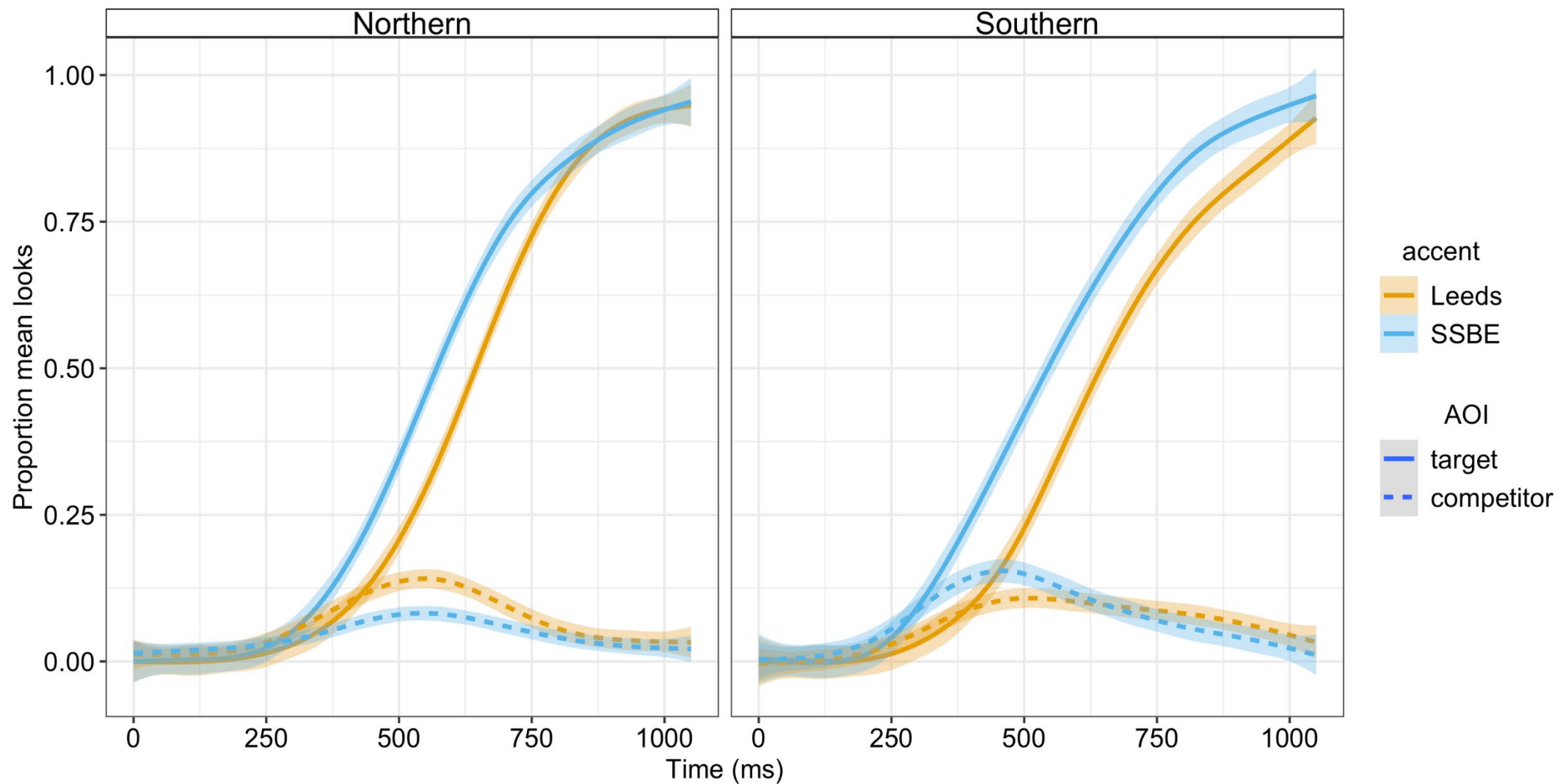
Northern listeners: 14

Southern listeners: 17



Results

TRAP-BATH raw data



TRAP-BATH statistical analysis

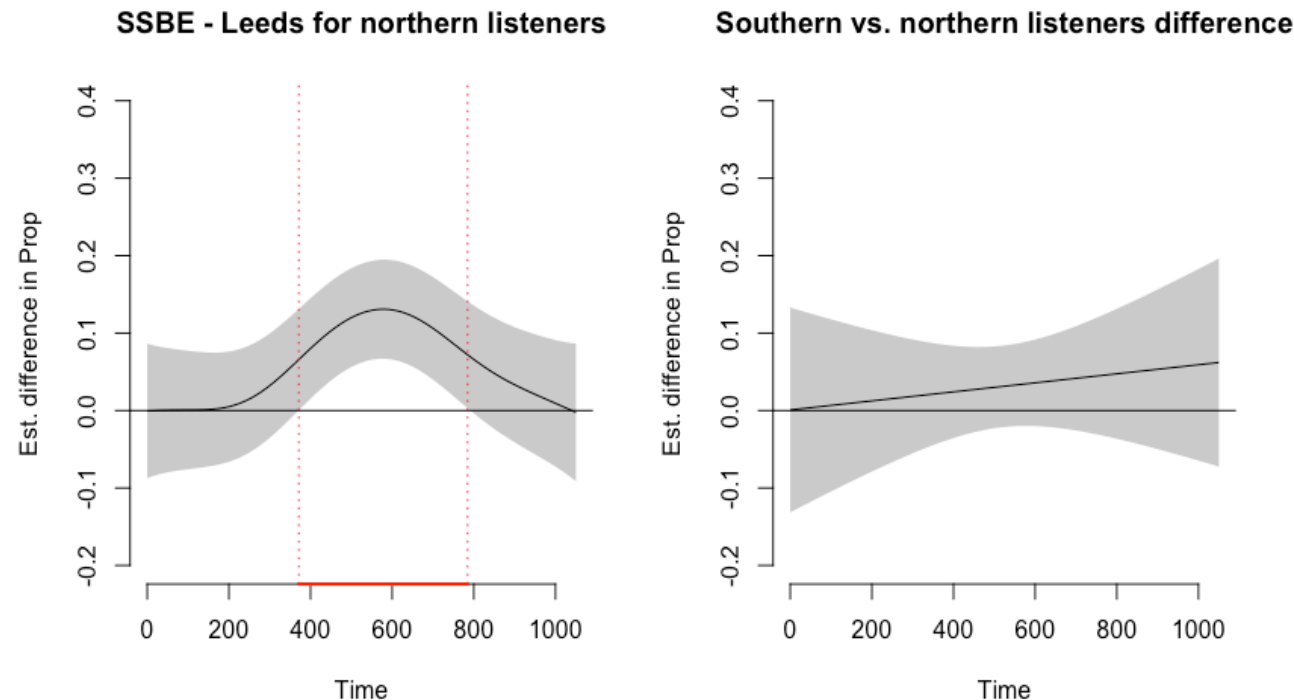
Generalised additive mixed models (GAMMs), Wieling, 2018.

- Separate models for TRAP-BATH and FOOT-STRUT sets.
- Model comparison was used to determine random-effects structure: 2 random smooths (participant by speaker's accent, item by listener's group)
- Model comparison indicated that speaker's accent improved the model, but listener's group did not.

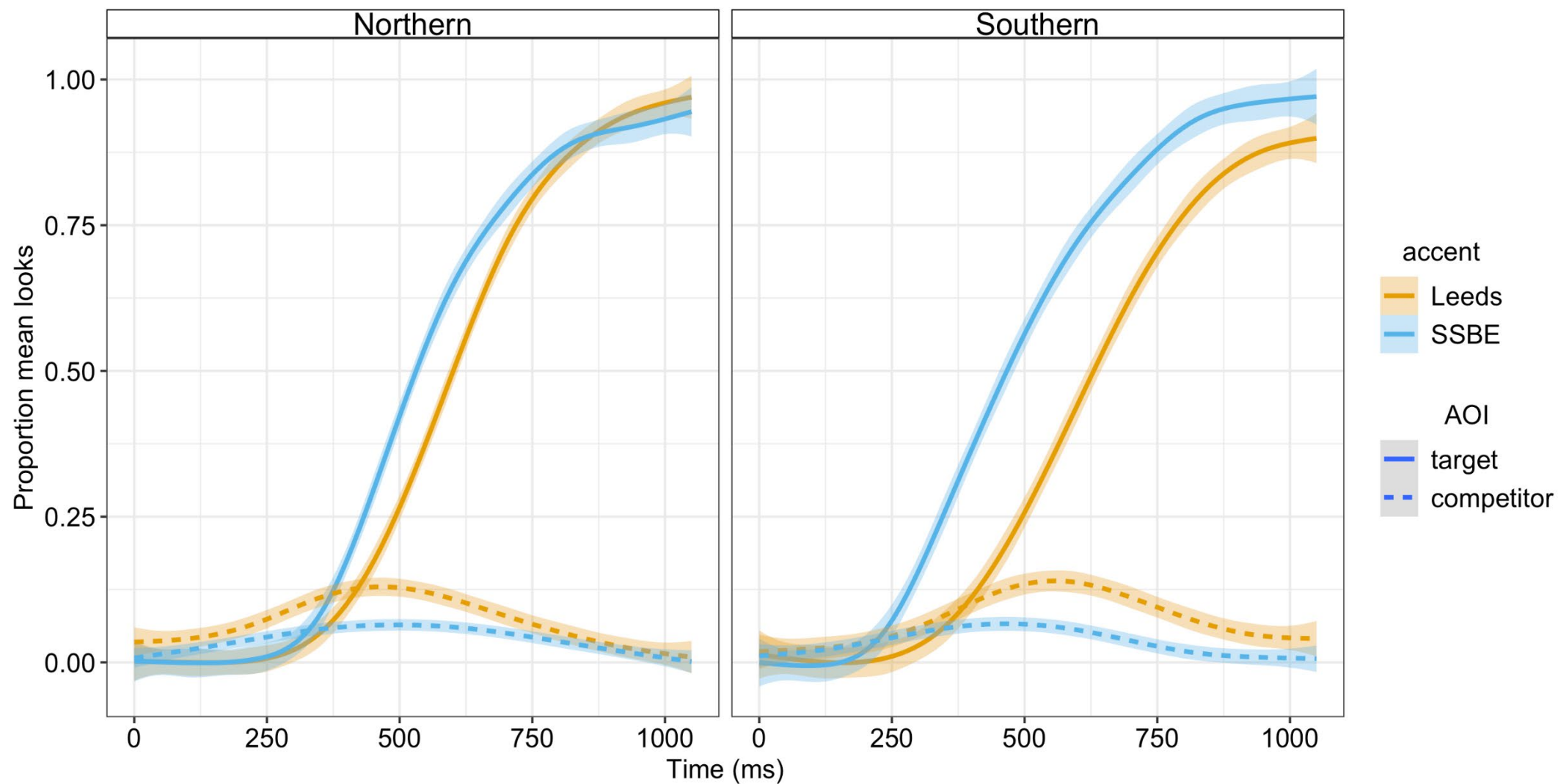
TRAP-BATH statistical analysis

Generalised additive mixed models (GAMMs), *Wieling, 2018*.

SSBE was processed significantly faster than Leeds by northern listeners.
The two accents were not processed differently by the two listener groups.



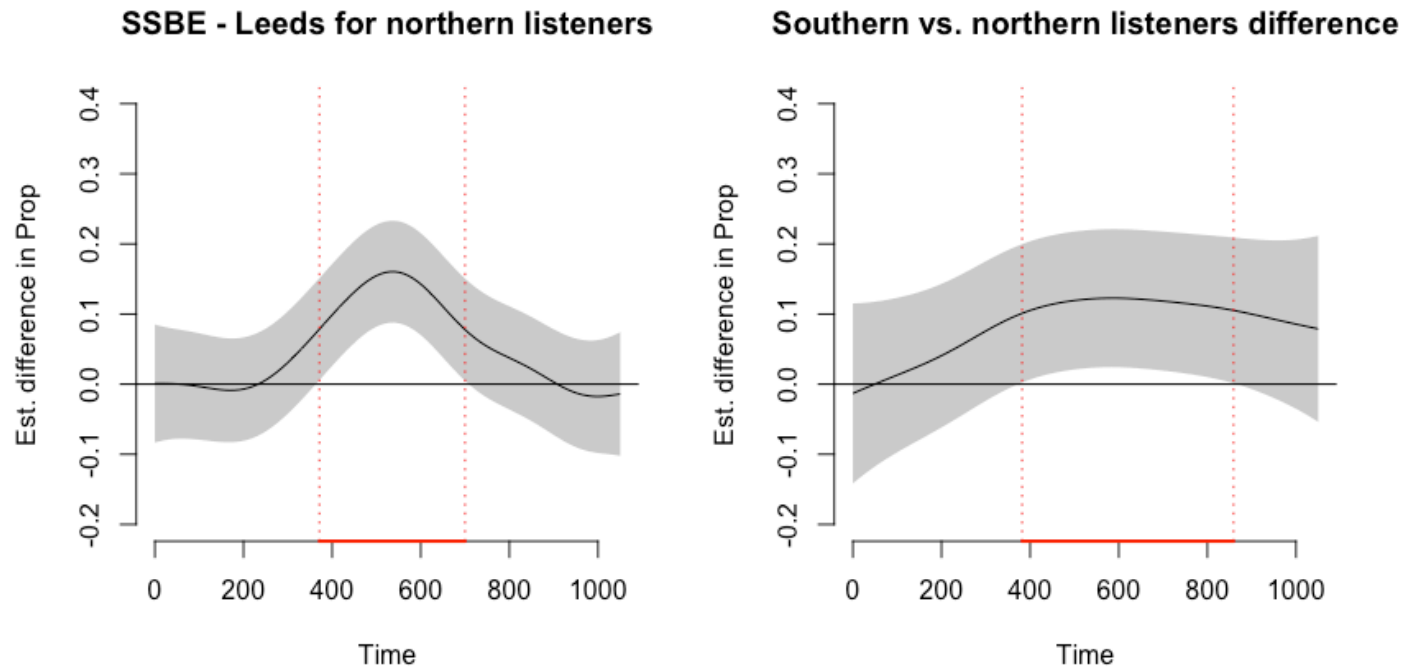
FOOT-STRUT raw data



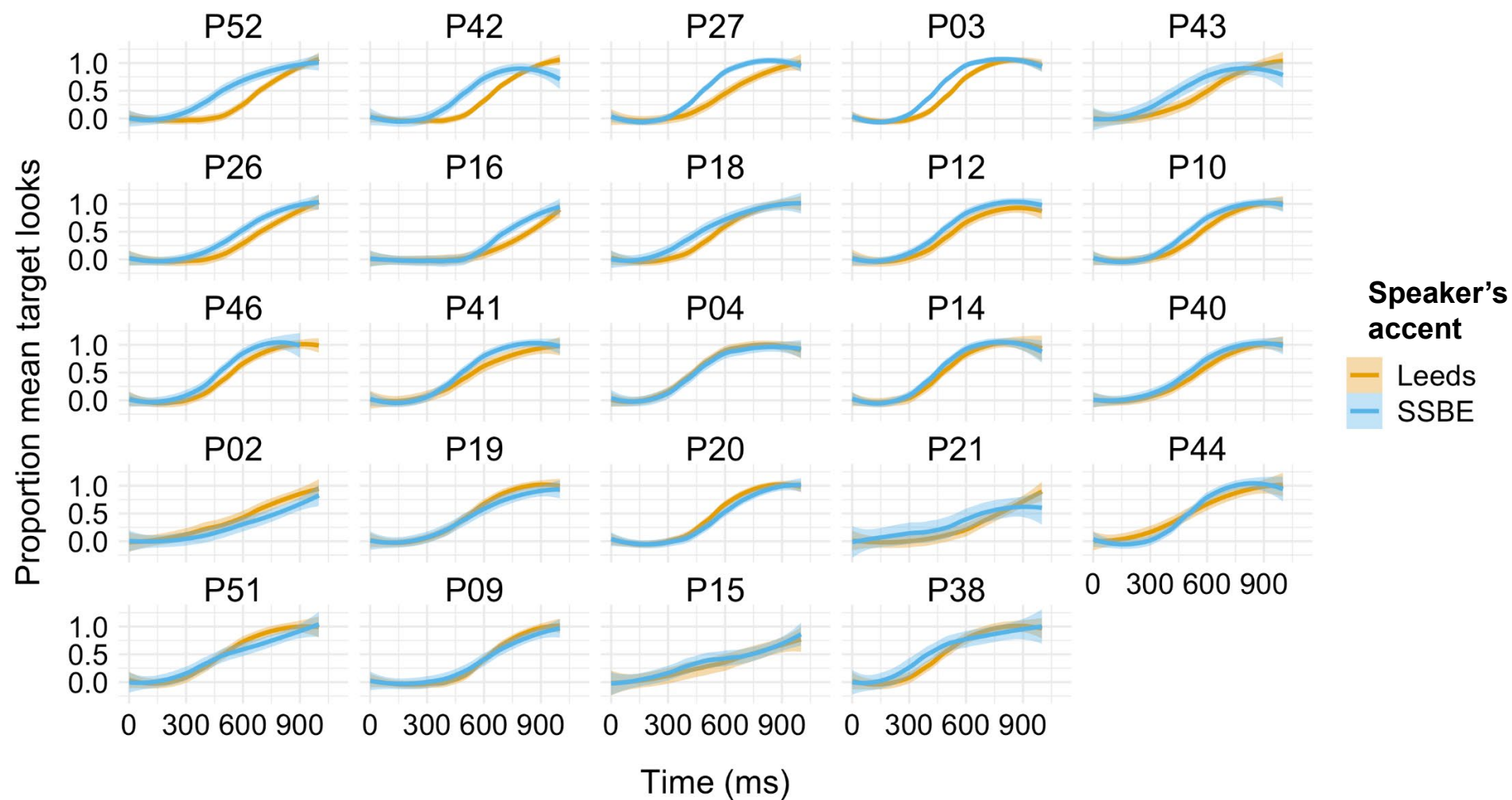
TRAP-BATH statistical analysis

Generalised additive mixed models (GAMMs), *Wieling, 2018*.

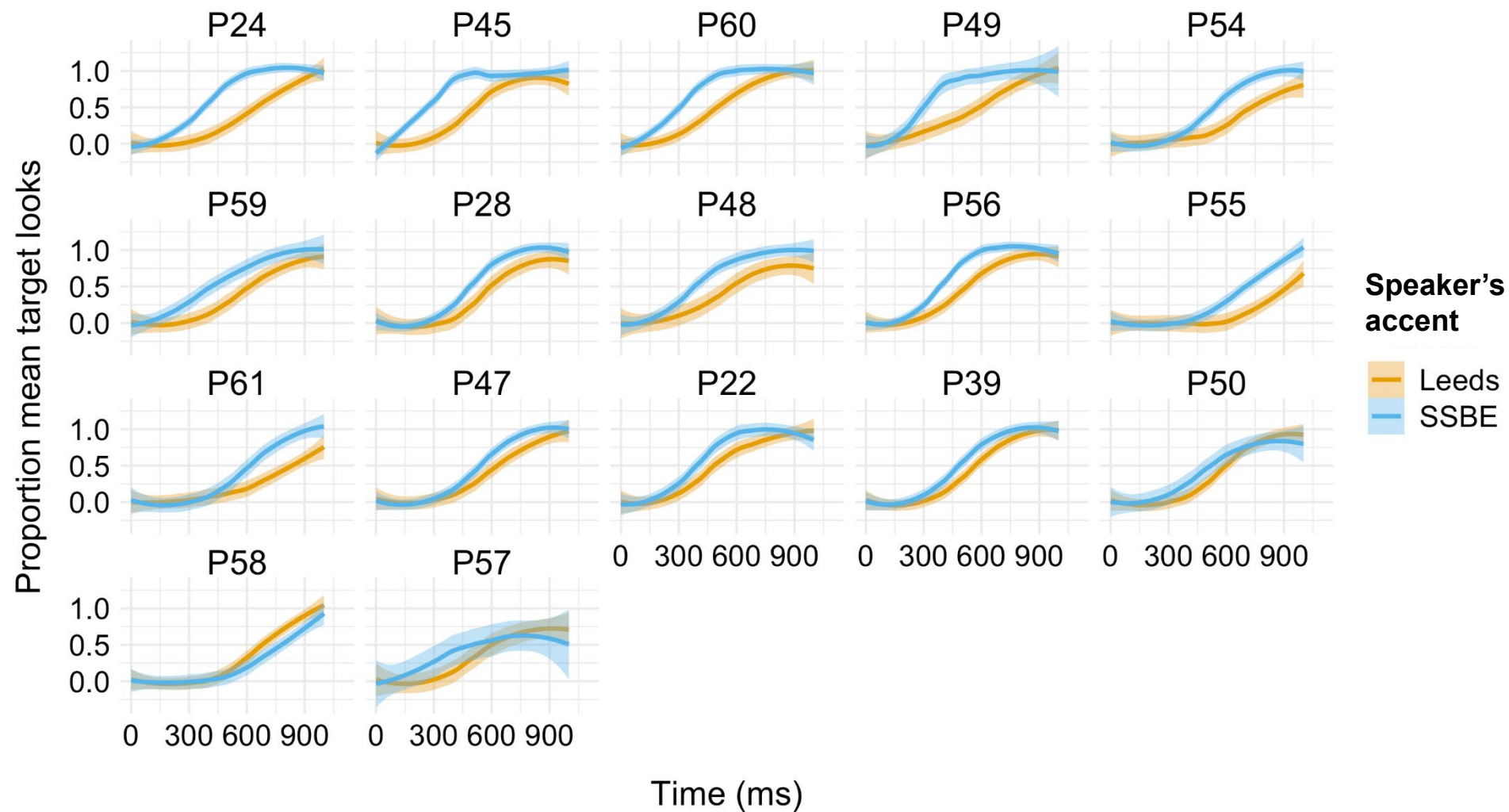
SSBE was processed significantly faster than Leeds by northern listeners.
The two accents were not processed differently by the two listener groups.



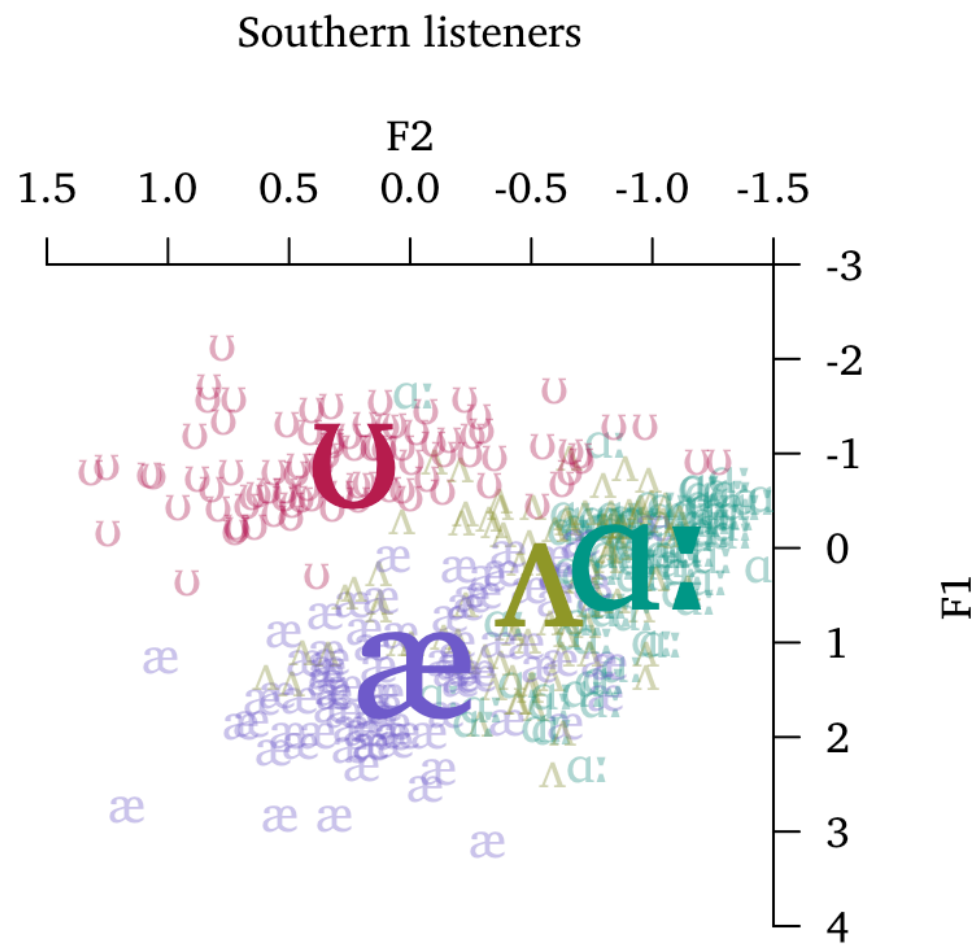
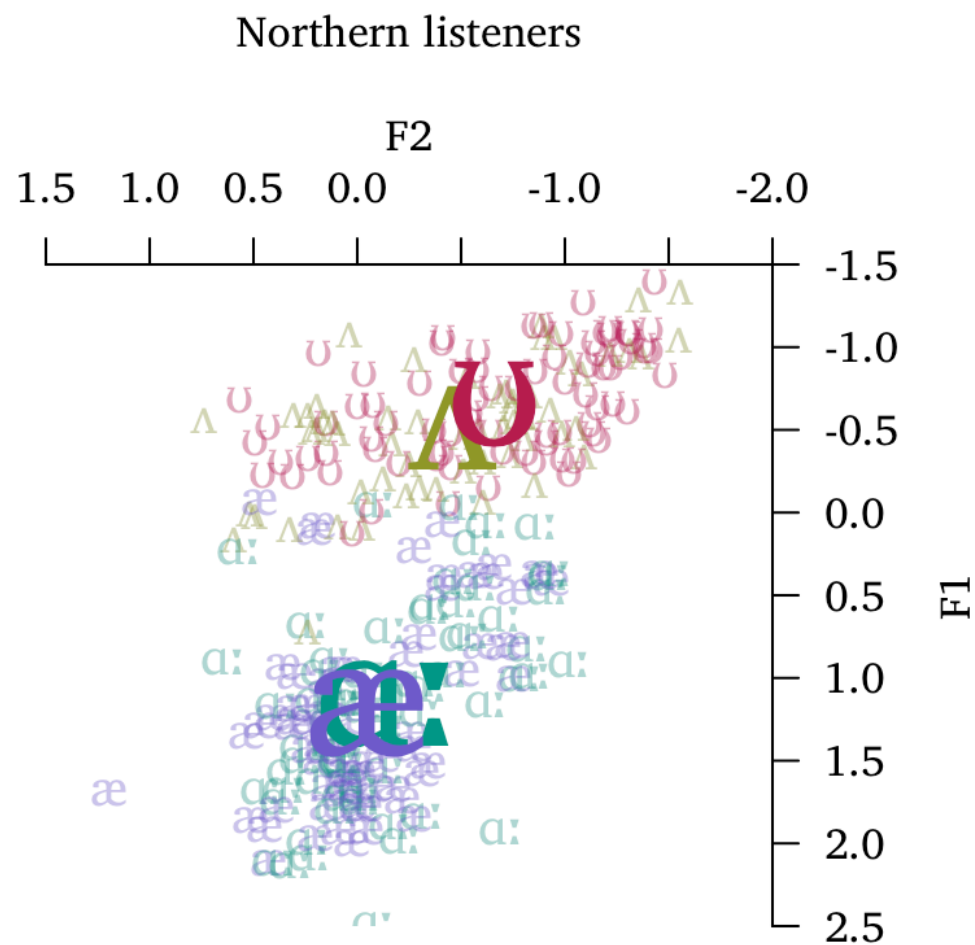
FOOT-STRUT northern listeners



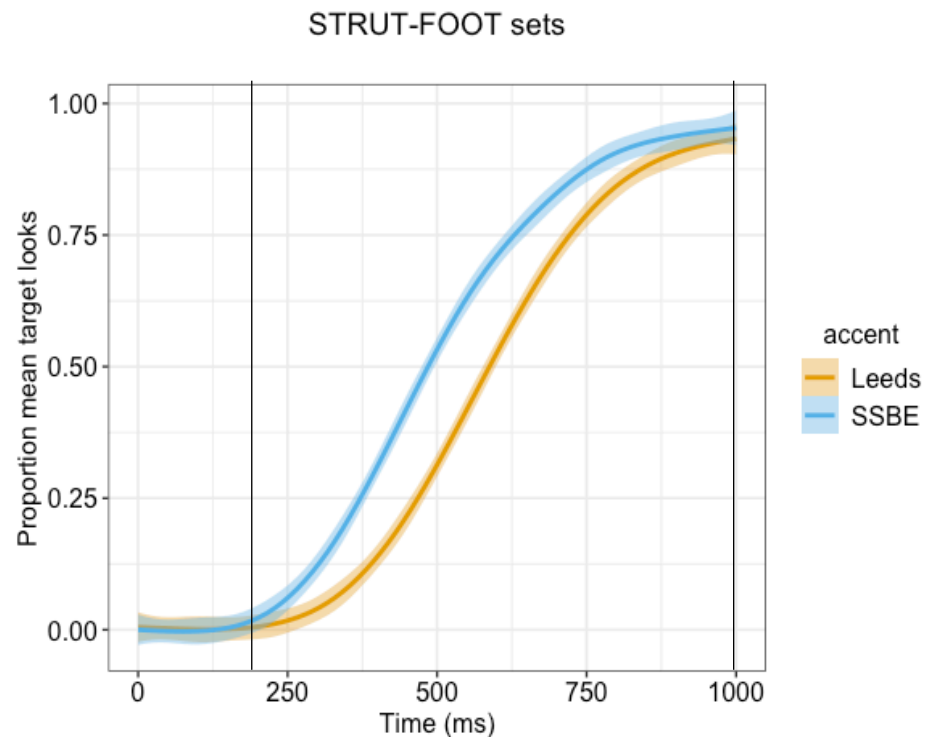
FOOT-STRUT southern listeners



Northern vs southern listeners: Production



Perception vs. production



‘Perceptual benefit’: Difference between the Proportion looks to target in the SSBE condition and the Leeds condition for each pair of sets and participant. Smaller time window 200ms -1000ms.

Pillai scores: degree of overlap between two distributions.

Discussion

Discussion

The TRAP-BATH distinction in SSBE will help Northern listeners disambiguate the words earlier; they will look at the target earlier in the SSBE condition.

CORRECT!

Both listener groups disambiguated the words earlier in the SSBE condition, where there was a difference between TRAP and BATH.

There doesn't seem to be a difference between the listener groups.

Discussion

The FOOT-STRUT distinction in SSBE will not necessarily help Northern listeners disambiguate the words earlier, as this is not a native contrast; both accent conditions will be similar.

INCORRECT...

Both listener groups were faster in the SSBE condition.

Northern listeners were able to use the distinction between STRUT and FOOT, even if they do not have this contrast in their native repertoire.

It is possible that with more data the difference between listener groups for STRUT-FOOT becomes apparent, but not supported by the statistical modelling for now.

Discussion: production vs. perception

Perception and production representations are not necessarily the same, but there is debate about the nature of the link...

Directly linked: Evans and Iverson (2007), Fridland and Kendall (2012), Pinget, Kager & Van de Velde (2019)

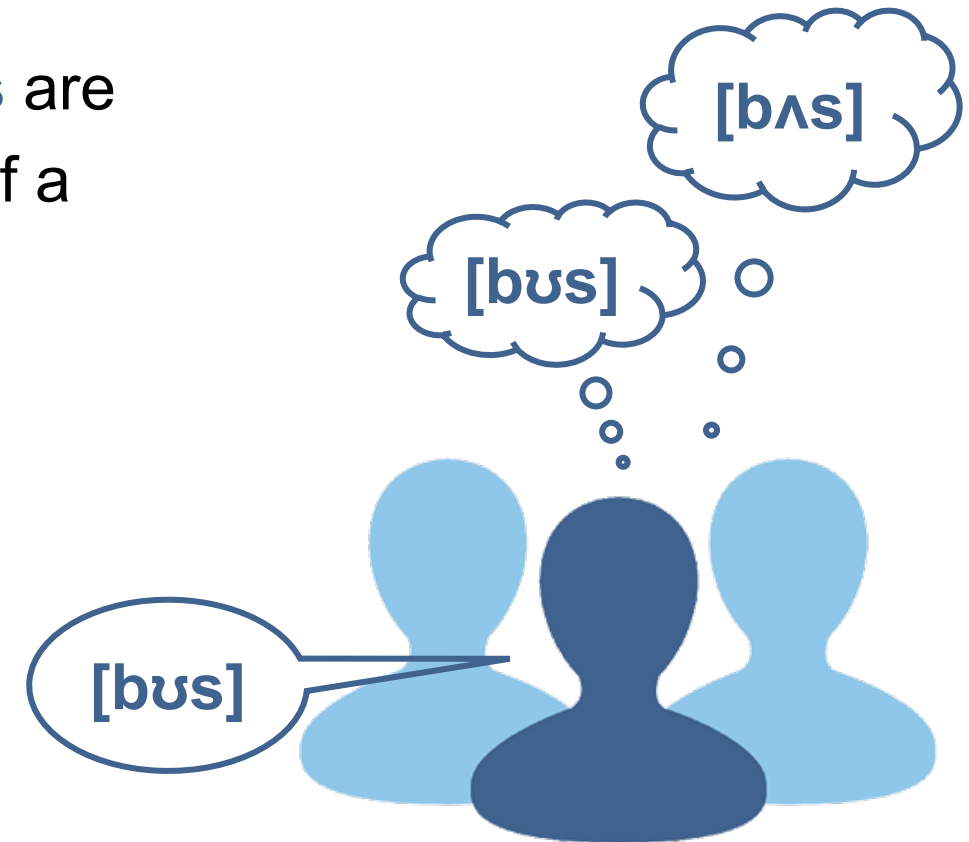
Our study: no apparent link in these tasks

Discussion: production vs. perception

Who are the listeners who can use categories they don't have in their "repertoire"?

Sumner & Samuel (2009: 498): **fluent listeners** are those “able to handle multiple regular variants of a particular word across dialects”

Does efficient mapping require being able to encode/store relevant representations?



Conclusion

Listeners are able to use contrasts that don't exist in their native repertoire when processing or adapting to a different accent, at least for a familiar accent.

Brief exposure to accent-specific features may help?

Next steps:

- Replication with SSBE listeners in the South
- Non-blocked conditions
- Other accents, other populations, e.g., L2 learners

Thanks to...

School of Languages, Cultures and Societies at the University of Leeds for funding the project.

Chris Norton, Cat Davies and Jamie Lingwood (Leeds) for help with eye-tracking.

Jieun Song (UCL) for help with recordings.

Bissera Ivanova (Leeds) for help with data collection.

Megan Lloyd and Abigail Folmer (Leeds) for help with participant recruitment and speech annotation.

Thank you all for listening!