The effect of regional variation on speech processing: evidence from an eye-tracking experiment.

Gisela Tomé Lourido[1], Robert Lennon[1] & Bronwen G. Evans[2]
University of Leeds[1], University College London[2]
g.tomelourido@leeds.ac.uk | r.w.lennon@leeds.ac.uk | bronwen.evans@ucl.ac.uk

UNIVERSITY OF LEEDS

UCL

RESEARCH QUESTION

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

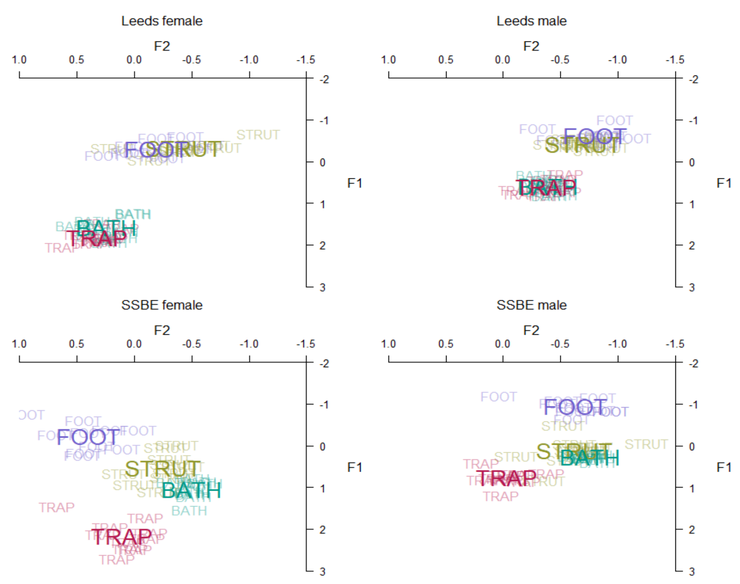
INTRODUCTION

- Hearing speech produced in an unfamiliar accent has a processing cost (Adank, et al. 2009; Floccia, et al. 2006), although listeners can rapidly adapt to novel talkers and accents (e.g., Bradlow & Bent, 2007, cf. Shaw et al. 2018).
- However, 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.
- These experiments often use pictures or words to cue a specific social category (e.g., gender, age, region) explicitly, but it is unclear whether **brief exposure to accent-specific phonetic features in the speaker's speech alone** would also influence speech processing.

AUDIO STIMULI

Audio stimuli

- Naturally-produced words recorded by 2 Leeds & 2 SSBE speakers (2 f, 2m)
- Embedded in carrier sentence: "I'm asking you to access _____" (Evans & Iverson, 2004)
- Leeds accent: [aɪm ˈæskɪŋ ju tə ˈækses]
- SSBE accent: [aɪm ˈɑːskɪŋ ju tə ˈækses]



DESIGN

Visual stimuli (following Best et al, 2013)

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

20 Test sets
- 10 BATH-TRAP contrasts
- 10 FOOT-STRUT contrasts
- Controlled for frequency

BUS + BOOK | PATH + PACK

20 Filler sets
- Contrasts are 'acoustically similar', e.g., DRESS-KIT, LOT-THOUGHT

LEG + LID | POP + PORT

Procedure

- Eyelink 1000 Plus eye-tracker (500 Hz sampling rate).
- Each trial consisted of a target-competitor pair.
- 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 in the centre which triggered the audio and clicked on the word they heard.

PREDICTIONS

BATH-TRAP

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

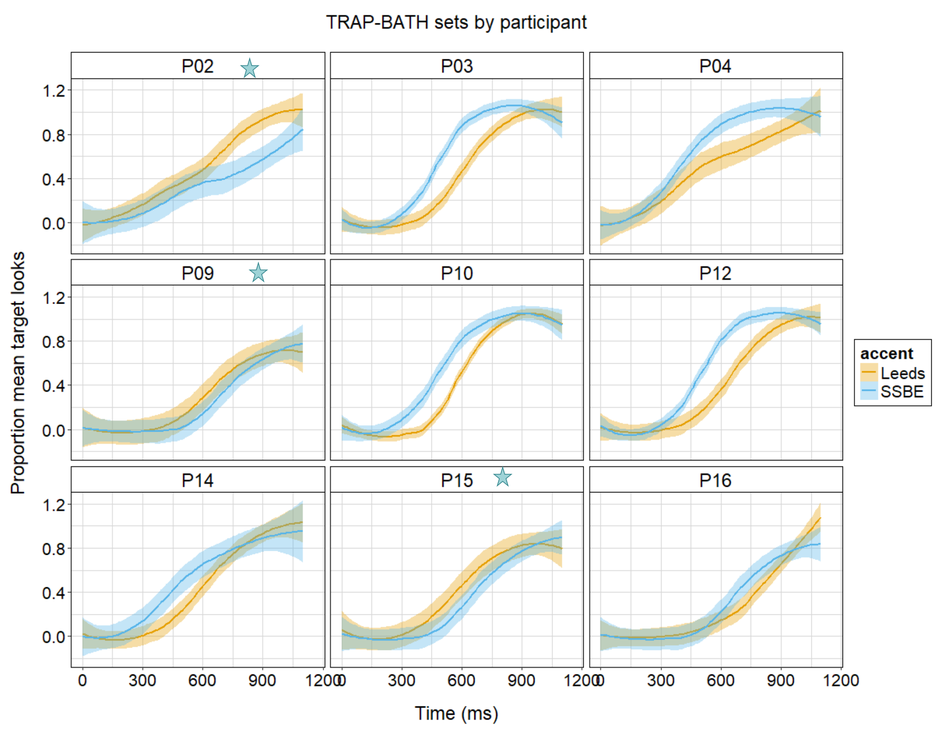
FOOT-STRUT

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

PARTICIPANTS

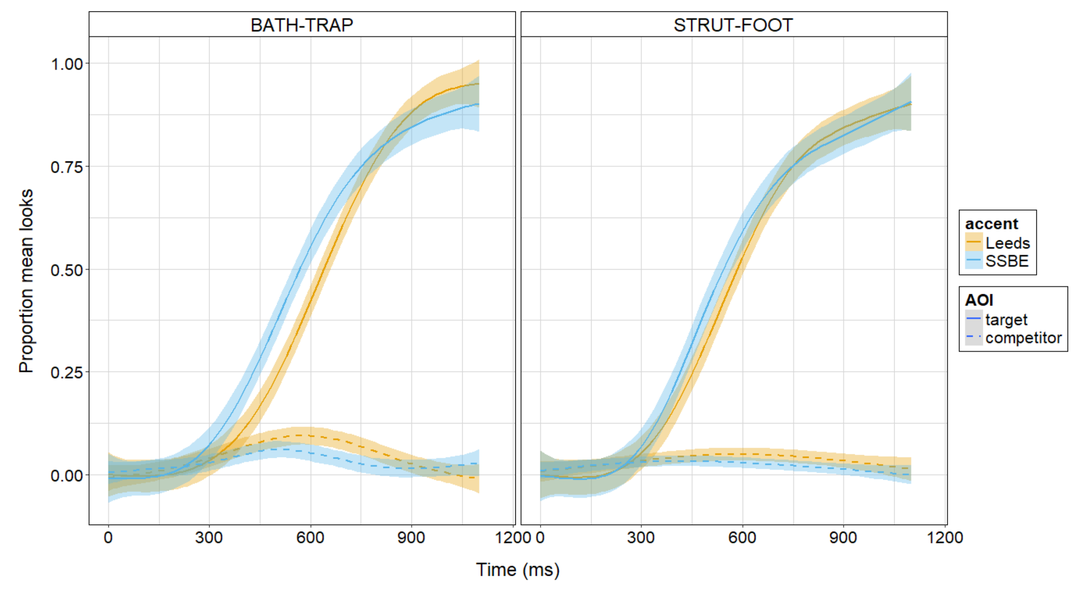
17 participants were tested (8 were excluded). The remaining 9:
- were 18-44 years old (mean = 28.9), 7 f, 2m
- were monolingual, born and raised in the North of England
- had not lived elsewhere for more than 8 months
- had parents who were monolingual and born in the North of England
  - only the parents of one participant lived elsewhere (P16)
- reported no speech, language, hearing or visual impairments

INDIVIDUAL DIFFERENCES



- Not all participants show a clear distinction between conditions
- P02, P09 and P15 reported not to have a contrast between TRAP and START. They used a front vowel for both lexical sets.

RESULTS



- Separate GAMM models were fitted for each variable following Sóskuthy (2017).
- Model comparison suggested that the inclusion of the parametric term and the smooth difference term for accent significantly improves the model fit for the BATH-TRAP model, but not STRUT-FOOT.

DISCUSSION

- **BATH-TRAP:** Despite the non-native vowel distribution, overall, Northern listeners are able to use this contrast to facilitate processing.
  - However, listeners who do not have the contrast (P02, P09, P15) perform similarly with Leeds & SSBE - there is no SSBE advantage.
- **STRUT-FOOT:** Northern listeners do not use this contrast to facilitate processing.
  - They don't have a STRUT vowel. It is possible they don't have a robust representation of this category.
- Overall, listeners were faster with STRUT-FOOT in both Leeds and SSBE and BATH-TRAP in SSBE than BATH-TRAP in Leeds.

Future work

- Finish data collection (Northern listeners), control group of SSBE listeners.

ACKNOWLEDGEMENTS. We thank School of Languages, Cultures and Societies (University of Leeds) for funding the study, Dr Chris Norton for his help with the eye-tracker and Dr Jieun Song for her assistance with the recordings.

REFERENCES. ADANK, P., EVANS, B. G., STUART-SMITH, J., & SCOTT, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Exp. Psychology. Human Perception and Performance*, 35(2), 520–529. BRADLOW, A. R., & BENT, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2), 707–729. EVANS, B. G., & IVERSON, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *JASA*, 115(1), 352. FLOCCIA, C., GOSLIN, J., GIRARD, F., & KONOPCZYNSKI, G. (2006). Does a regional accent perturb speech processing? *Journal of Exp. Psychology: Human Perception and Performance*, 32(5), 1276–1293. KOOPS, C., GENTRY, E., & PANTOS, A. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics*, *14*(Selected Papers from NWAV 36), Article 12. SHAW, J. A., BEST, C. T., DOCHERTY, G., EVANS, B. G., FOULKES, P., HAY, J., & MULAK, K. E. (2018). Resilience of English vowel perception across regional accent variation. *LabPhon*, 9(1), 1–36. SÓSKUTHY, M. (2017) Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. Working Paper. STRAND, E. A. (1999). Uncovering the Role of Gender Stereotypes in Speech Perception. *Journal of Language and Social Psychology*, *18*(1), 86–100.TANENHAUS, M. K., SPIVEY-KNOWLTON, M. J., EBERHARD, K. M., & SEDIVY, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.