Talker and sentence variability attenuate speaking rate normalization in speech perception

Caleb J. King [a], Chloe M. Sharpe [b], Anya E. Shorey [a], Christian E. Stilp [a]

[a] Department of Psychological and Brain Sciences, University of Louisville

[b] School of Psychology, Xavier University

Introduction

Acoustic context effects influence speech perception

- **Spectral Contrast Effects** (**SCEs**) – perceiving larger changes in *spectral* properties between two sounds than are physically present
- **Temporal Contrast Effects** (**TCEs**, aka speaking rate normalization) – perceiving larger changes in *temporal* properties between two sounds than are physically present

Variability challenges speech perception

- Speech perception is often faster and/or more accurate when hearing one talker compared with hearing multiple talkers (i.e. talker adaptation / normalization)

Does this variability impact acoustic context effects?

- SCEs were smaller when context sentences were spoken by different talkers (Assgari & Stilp, 2015)
- Impact of variability on TCEs is unknown, and is the objective of this study

Method

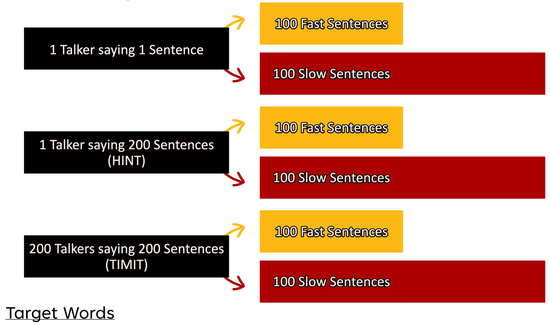
Participants

20 normal-hearing native English-speaking undergraduates

Stimuli

Context Sentences

- Taken from Assgari & Stilp (2015)
- Speaking rates manipulated in PRAAT to make half of the sentences fast (50% of original duration) and the other half slow (150% of original duration)



Target Words

- Natural ten-step series varying in VOT from “deer” to “tier”; same talker as in the 1 Talker/1 Sentence condition
- Each trial = one context sentence followed by one target word

Procedure

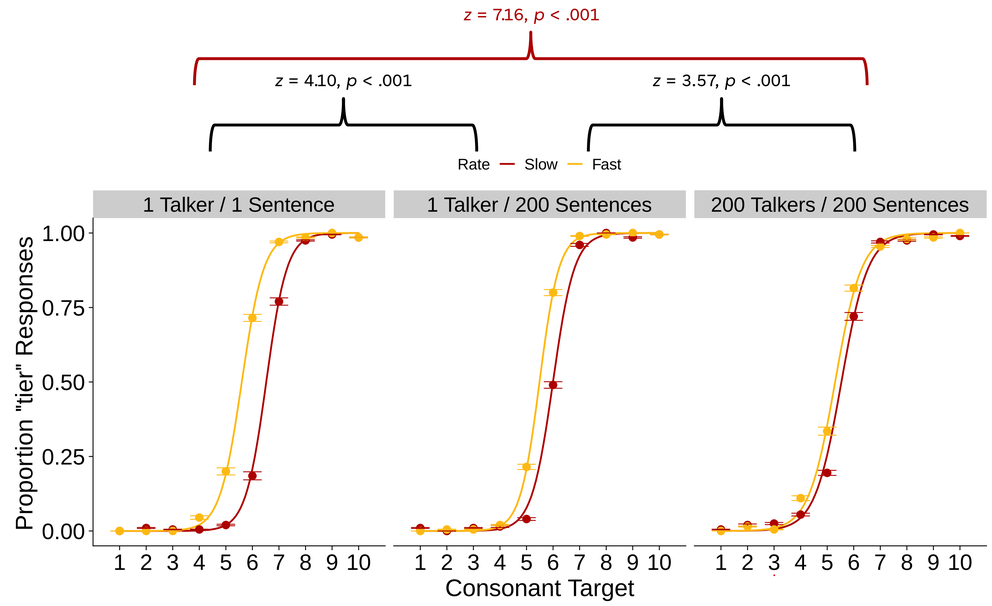
60-minute online study using Gorilla

- Headphone screen (Woods et al., 2017)
- Practice categorizing “deer”/“tier” target endpoints (80% correct required to continue)
- Main task: Three blocks, each with 200 trials in random orders

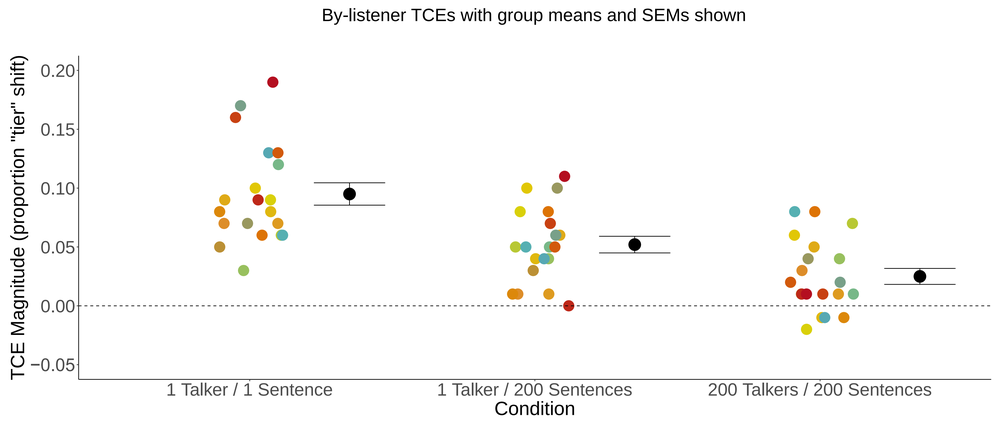
Results

Generalized linear mixed-effects model predicting “tier” responses

- *Fixed effects*: Target, condition, speaking rate, and their interactions
- *Random effects*: Random slopes for target and condition, random intercepts for participants



Sentence variability decreases TCE magnitudes; sentence *and* talker variability decrease TCE magnitudes even further



Discussion

TCE magnitudes decreased when a different sentence was heard on each trial, and decreased again when a different talker spoke a different sentence on each trial. This pattern of results is similar but not identical to Assgari and Stilp (2015), which reported smaller SCEs due to talker variability only using these same context sentences.

Why are these patterns of results (slightly) different?

- Different context effects (SCEs vs. TCEs)
- Different mechanisms behind the context effect
  - **SCEs**: neural adaptation (Stilp, 2020)
  - **TCEs**: either entrainment to amplitude envelope oscillations (Bosker & Ghitza, 2018) or evoked responses to rapid amplitude increases (acoustic edges; Oganian & Chang, 2019)
- Different target stimuli
  - Assgari & Stilp (2015): isolated vowels
  - Here: words with initial consonant

What is it about the variability that diminishes context effects?

- **SCEs**: f0 variability (Assgari et al., 2019)
- **TCEs**: unlikely to be f0 variability (TCEs differed across the one-talker conditions)
  - Different sentences (lexical content) and different talkers (prosody) can each vary the amplitude envelopes of context sentences (and thus vary their oscillations/acoustic edges)
- Future study: isolate talker variability to understand its unique contribution to these results (200 Talkers/1 Sentence)

Psychometric function slopes patterned slightly differently from that of TCEs

- **TCEs**: 1 Talker/1 Sentence > 1 Talker/200 Sentences > 200 Talkers/200 Sentences
- **Slopes**: (1 Talker/1 Sentence = 1 Talker/200 Sentences) > 200 Talkers / 200 Sentences
- Link between task difficulty and TCE magnitude not entirely clear

References

Assgari, A. A., & Stilp, C. E. (2015). Talker information influences spectral contrast effects in speech categorization. *The Journal of the Acoustical Society of America*, *138*(5), 3023–3032.

Assgari, A. A., Theodore, R. M., & Stilp, C. E. (2019). Variability in talkers' fundamental frequencies shapes context effects in speech perception. *The Journal of the Acoustical Society of America*, *145*(3), 1443–1454.

Bosker, H. R., & Ghitza, O. (2018). Entrained theta oscillations guide perception of subsequent speech: behavioural evidence from rate normalisation. *Language, Cognition and Neuroscience*, *33*(8), 955–967.

Oganian, Y., & Chang, E. F. (2019). A speech envelope landmark for syllable encoding in human superior temporal gyrus. *Science advances*, *5*(11), eaay6279.

Stilp, C. E. (2020). Evaluating peripheral versus central contributions to spectral context effects in speech perception. *Hearing Research*, *392*, 107983.

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*(7), 2064–2072.