Titrating effects of acoustic variability on context effects and psychometric function slopes in speech categorization

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

Speech perception is influenced by Temporal Contrast Effects (TCEs, aka speaking rate normalization)

- Changes in speaking rate produce larger perceived changes in temporal properties between sounds than are physically present

Stimulus variability impedes TCEs (King et al., 2022, ASA)

- Relative to baseline (1 Talker / 1 Sentence), TCE magnitudes decreased when a different sentence was heard on each trial (1 Talker / 200 Sentences), and decreased again when a different talker spoke a different sentence on each trial (200 Talkers / 200 Sentences)
- Psychometric function slopes were shallower when a different talker spoke a different sentence on each trial (200 Talkers / 200 Sentences), relative to the other conditions

However, stimuli varied beyond the number of talkers and sentences

- Each condition tested slow sentences, but the slow speaking rates had different amounts of variability by condition; same for fast
- Here, we controlled speaking rate variability to test its role on TCEs and psychometric function slopes

Method

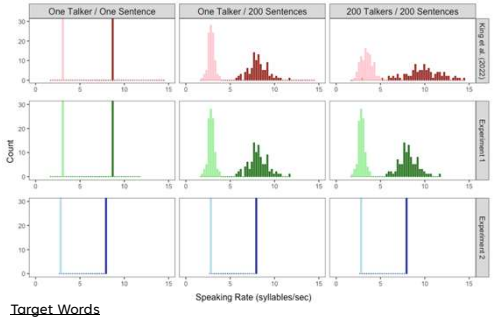
Participants

Normal-hearing native English-speaking undergraduates
King et al. (2022): n = 20; **E1:** n = 22; and **E2:** n = 24

Stimuli

Context Sentences

- **King et al. (2022):** 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)
- **E1:** Matched speaking rates of 200 Talkers / 200 Sentences to 1 Talker / 200 Sentences
- **E2:** All fast sentences set to 8.0 syllables per second, all slow sentences set to 2.67 syllables per second



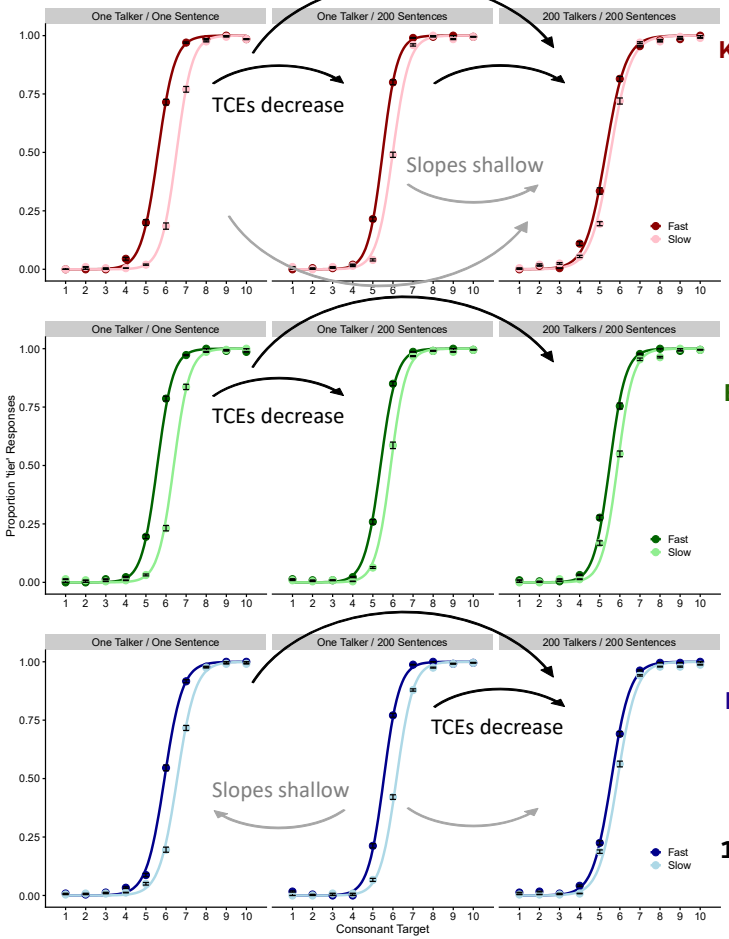
Target Words

- Natural ten-step series varying in VOT from "deer" to "tier"
- Each trial = one context sentence followed by one target word

Results

Generalized linear mixed-effects modeling predicting "tier" responses

- *Fixed effects*: Target ("deer" to "tier" continuum), condition (# Talkers / # Sentences), speaking rate (slow or fast), and their interactions
- *Random effects*: Random slopes for target and condition, random intercepts for participants



King et al. (2022): Both sentence and talker variability decrease TCE magnitudes

Talker variability produced shallower slopes

E1: Talker variability no longer decreases TCE magnitudes

Slopes are similar in all conditions

E2: Sentence variability no longer decreases TCE magnitudes

Steeper slopes in 1 Talker / 200 Sentences

Discussion

Speaking rate variability influences TCEs in speech categorization

- **E1:** As expected, between the conditions where speaking rates were matched, there was no difference in TCE magnitudes
- **E2:** Although all speaking rates were the same across conditions, TCE magnitudes still decreased with increased talker variability, patterning differently than expected

Speaking rate variability influences psychometric slopes in speech categorization

- **E1:** With further controlled speaking rate variability, psychometric function slopes were similar across all conditions, indicating no difference in task difficulty
- **E2:** In contrast, although fast and slow speaking rates were held constant across conditions, psychometric function slopes were unexpectedly steeper in One Talker / 200 Sentences

Not all sources of variability are equally consequential

- Variability in context sentences' mean f0 alters spectral contrast effect magnitudes (Assgari et al., 2019), but not so for variability in context sentences' mean F1 or mean F3 (Mills et al., 2022)
- Speaking rate variability harms performance, but amplitude variation does not (Sommers et al., 1994)
- Even while matching speaking rates, sentences differed in other (potentially perceptually salient) properties
  - Duration
  - Others also possible (lexical content, syntax, etc.)

Potential analogy to modulation informational masking (Conroy & Kidd, 2021)

- Both measured using a context-target trial structure
- Stimulus uncertainty – context rate varied from trial to trial
  - Conroy & Kidd – nonspeech masker modulation rate
  - Here – sentence speaking rate
- Performance decreased when masker modulation rate varied from trial to trial (Conroy & Kidd, 2021)
- While our paradigm did not include masking, these results might highlight the role of modulation / speaking rate uncertainty

Individual differences analyses

- Each listener group completed 3 conditions, but TCE magnitudes were not significantly correlated in any pairwise combinations of conditions in any experiment
- Stimulus variability may limit test/retest reliability of these context effects (Heffner & Myers, 2019)

References

Assgari, A. A., Theodore, R. M., & Stilp, C. E. (2019). *JASA*, *145*(3), 1443-1454.

Conroy, C., & Kidd Jr, G. (2021). *JASA*, *149*(5), 3665-3673.

Heffner, C. C., & Myers, E. B. (2019). *JASA*, *145*(3), 1790.

King, C. J., Sharpe, C. M., Shorey, A. E., & Stilp, C. E. (2022). *JASA*, *151*(4), A262-A263.

Mills, H. E., Shorey, A. E., Theodore, R. M., & Stilp, C. E. (2022). *JASA*, *152*(1), 55-66.

Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). *JASA*, *96*(3), 1314-1324.