INTRODUCTION

- Perception of a speech sound is heavily influenced by surrounding sounds. When spectral properties differ between earlier (context) and later (target) sounds, this can produce spectral contrast effects (SCEs) that bias categorization of later sounds.

- SCEs affect vowel categorization in simulations of cochlear implant (CI) processing (Stilp, 2017) and for CI users (Feng & Osenholtz, 2018).
- Also, normal-hearing listeners’ SCEs are smaller when they hear 200 different talkers compared to one talker (Assgari & Stilp, 2015).
- Finally, CI users can struggle with talker discrimination (e.g., Fu et al., 2004, Massida et al., 2011, Stickle et al., 2004).
- This study tested the interaction of these three threads. We predicted that talker variability would not affect SCE magnitudes in simulated CI (noise-vocoded) vowel categorization.

METHODS

Participants

- 40 native English undergraduate students with self-reported normal hearing

Stimuli

1. Context Sentences

- 1 male talker reading 200 different sentences (HINT: Nilsson et al., 1994); 200 talkers reading 200 different sentences (TIMIT: Garofolo et al., 1990)
- Filtered to amplify low-F1 frequencies (100-400 Hz) or high-F1 frequencies (550-850 Hz) by 5 dB to produce SCEs (after Assgari & Stilp, 2015)
- Noise-vocoded from 100-5000 Hz using 4, 8, 12, or 24 channels

2. Target Vowels

- 10-step series perceptually varying from “ih” (as in ‘bit’) to “eh” (as in ‘bet’); as tested in Assgari & Stilp (2015)

Procedure

- On each trial, participants heard a sentence and categorized the following target vowel:
  - a) 1 Talker, 4 Channels
  - b) 1 Talker, 8 Channels
  - c) 200 Talkers, 4 Channels
  - d) 200 Talkers, 8 Channels

RESULTS

SCE Calculation

- In each block, SCEs were measured as the difference in mean percentage “eh” responses across filtering conditions (low-F1-amplified versus high-F1-amplified)

Experiment 1 (n = 17)

- Experiment 1 was more challenging than anticipated. Response curves were flat lines (means and standard errors shown above), reflecting participants’ inability to distinguish target vowels. Measuring categorization shifts (SCEs) due to preceding spectral context is problematic when listeners do not have clear categories to begin with. This led to Experiment 2.

Experiment 2 (n = 18)

- Performance improved substantially in Experiment 2 owing to testing higher spectral resolution (more channels than Experiment 1; response means and standard errors shown above)

- SCEs were calculated for each participant in each block (means and standard errors shown at right), then analyzed in a 2 (number of talkers) x 2 (number of spectral channels) repeated measures ANOVA
  - No main effect of number of talkers: F(1,17) = 0.03, p = 0.88
  - No main effect of number of channels: F(1,17) = 0.57, p = 0.46
  - No interaction between talkers and channels: F(1,17) = 0.00, p = 0.97

- The interaction between the number of talkers and the number of spectral channels in noise vocoding was not statistically significant. SCE magnitudes were relatively constant across conditions. This supports the prediction that talker variability would not affect SCE magnitudes. However, the challenge of interpreting a null result remains.

- Results may be due in part to impoverished pitch cues in noise vocoding. Assgari et al. (2019) showed that variability of acoustically dissimilar talkers (high F0 variability) resulted in smaller SCEs. Talker variability does not appear to affect spectral context effects in categorization of noise-vocoded vowels. This patterns differently from talker variability affecting SCEs for normal-hearing listeners.

- Experiment 1: spectral channels (4 / 8) x number of talkers (1 / 200)
- Experiment 2: spectral channels (12 / 24) x number of talkers (1 / 200)

REFERENCES

Assgari AA, Theodore RM, Stilp CE (2019) JASA, 143(6), EL468-EL473
Feng L, Oxenham AJ (2018) JASA, 144(6), 253-260
Garofolo JS et al. (1990) DARPA TIMIT CD-ROM, NIST
Stilp CE (2017) JARO, 19(1), 465-483