

Dissociating contributions of talker gender and acoustic variability for spectral contrast effects in vowel categorization

Ashley A. Assgari¹, Asim Mohiuddin¹, Rachel M. Theodore² and Christian E. Stip¹

¹Department of Psychological and Brain Sciences, University of Louisville

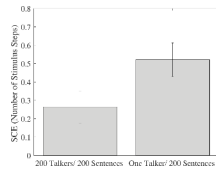
²Connecticut Institute for the Brain and Cognitive Sciences, University of Connecticut

INTRODUCTION

Spectral contrast effects (SCEs) occur when the auditory system perceptually magnifies spectral differences between sounds:

Precursor	More likely to hear
Sentence (unmodified)	/i/ or /e/ vowel target
Sentence with /e/-like (high F ₁) frequencies emphasized	/i/ (low F ₁)
Sentence with /i/-like (low F ₁) frequencies emphasized	/e/ (high F ₁)

SCEs in vowel perception were diminished when precursor sentences were spoken by multiple talkers rather than one talker (Assgari & Stip, 2015 JASA)



They suggested that greater variability in talker f₀ could be responsible for smaller SCEs, but talker f₀ was not controlled in their experiment

However, acoustic variability was confounded with higher-level variability (talker gender), which also freely varied

- Are SCEs smaller when the gender of who spoke the precursor sentence differs from who spoke the target vowel?

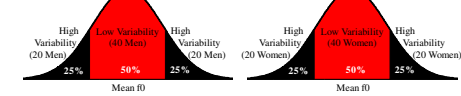
We investigated whether **gender variability** and **f₀ variability** show separable contributions to the attenuation of SCEs in vowel perception

METHODS

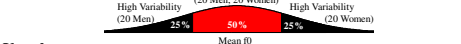
Sentences

- Measured mean f₀ in TIMIT sentences for Men and Women
- Sorted into Low f₀ Variability and High f₀ Variability conditions

EXPERIMENT 1



EXPERIMENT 2



Vowels

- Natural vowels interpolated from [i] to [e] using PRAAT
- Same stimuli as used in Assgari & Stip (2015)

Filters

- Bandpass filter added a +5 dB spectral peak to precursor sentences in low-F₁ (100-400 Hz) or high-F₁ (550-850 Hz) region

Participants

- 20 normal-hearing native English speakers in each experiment

METHODS (CONTINUED)

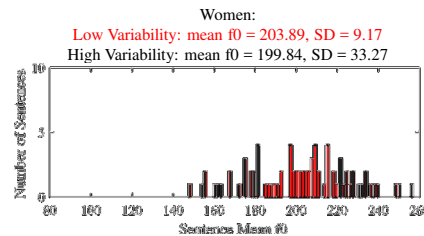
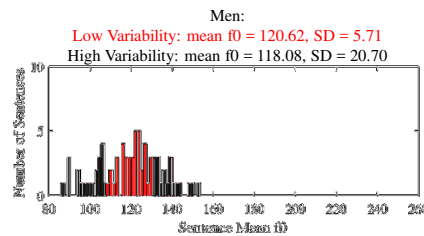
Procedure

- Each trial presented a sentence then a target vowel which listeners identified as “ih” as in “bit” or “eh” as in “bet” (see schematic in Introduction)
- Practice: 20 sentences from the AzBio corpus (Spahr *et al.*, 2012 *Ear Hear*) paired with endpoint vowels; >80% accuracy needed to continue to test session
- Test: 4 blocks (Experiment 1) or 2 blocks (Experiment 2) of 160 trials (each talker repeated 4 times per block)

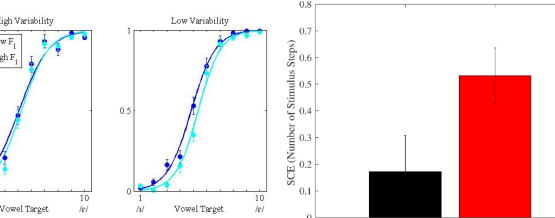
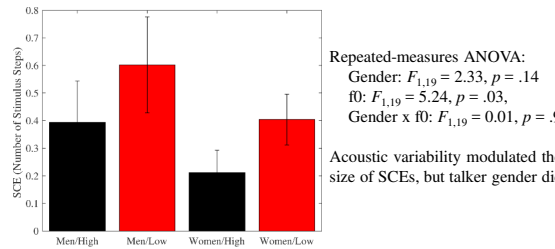
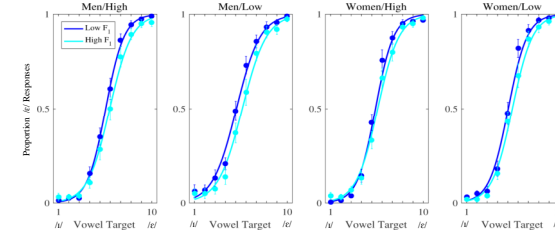
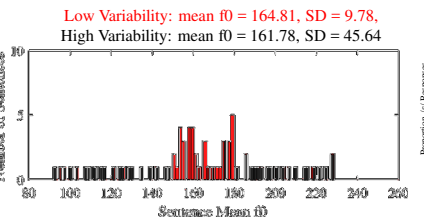
Analyses

- Logistic regressions were fit to each listener’s responses to low-F₁ and high-F₁ sentences. Midpoints were calculated from each regression equation.
- SCE magnitudes were defined as differences in midpoints between low-F₁ and high-F₁ functions (*i.e.*, number of stimulus steps along the abscissa)

EXPERIMENT 1: Blocked gender, low/high f₀ variability



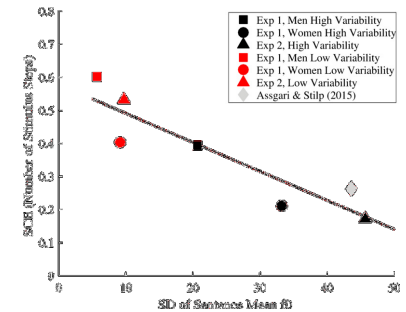
EXPERIMENT 2: Mixed gender, low/high f₀ variability



DISCUSSION

SCEs in vowel perception are smaller when mean f₀ is highly variable across sentences

- SCEs are not diminished simply by hearing multiple talkers but rather by high acoustic variability between talkers
- Across experiments where multiple talkers were heard, a strong negative relationship exists between SCE magnitude and f₀ variability ($r = -0.92, p < .01$)



Talker gender does not seem to influence SCEs

- Modest trend for gender effect in Experiment 1 is likely due to slightly larger standard deviations of f₀ for Women
 - Sentences were selected by matching z-scores across Men and Women, but this does not equate SD across Men and Women
 - Larger SDs result in smaller SCEs (see above figure)
 - Future studies should equate f₀ variability, not z-scores
- SCEs were not significantly smaller when the gender of who spoke the sentence (woman) did not match the gender of who spoke the target vowel (man)
- SCE magnitudes were similar for low f₀ variability from mixed genders (Experiment 2), one gender (Experiment 1), or one talker (Assgari & Stip, 2015)

Are these effects due to f₀ *per se*?

- f₀ is correlated with other acoustic properties (e.g., F₁, F₂, F₃)
- Manipulating f₀ variability also manipulates F₁, F₂, and F₃ variability, which might better explain these results

Future directions

- f₀ measures beyond its mean across a sentence:
 - What role does f₀ variability within a sentence play?
- Measures and manipulations of acoustic variability beyond f₀:
 - Is categorization of vowels that differ in F₁ better explained by F₁ variability than f₀ variability?
- Effects of high acoustic variability for a single talker:
 - However, this falls well short of High Variability across talkers; will smaller SCEs still be observed?

In conclusion, lower-level variability shapes context effects in speech perception more than higher-level source information.