Spectral contrast effects in vowel categorization by listeners with sensorineural hearing loss

2aSC14

Christian Stilp, Department of Psychological and Brain Sciences, University of Louisville
Joshua Alexander, Department of Speech, Language, and Hearing Sciences, Purdue University
christian.stilp@louisville.edu

INTRODUCTION

- Phoneme categorization is influenced by *spectral contrast effects* (SCEs), the perceptual magnification of spectral differences between sounds. For example (after Ladefoged & Broadbent, 1957):

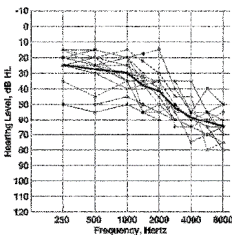
Precursor	More likely to hear
Sentence (unmodified)	/ɪ/ or /ɛ/ vowel target
Sentence with /ɛ/-like (high F_1) frequencies emphasized	/ɪ/ (low F_1)
Sentence with /ɪ/-like (low F_1) frequencies emphasized	/ɛ/ (high F_1)

- SCEs are larger when F_1-emphasized spectral peaks in the preceding sentence are higher-amplitude or broader-bandwidth (Stilp *et al.*, 2015).
- Despite their widespread influence on speech perception for normal-hearing (NH) listeners, SCEs have never been measured in hearing-impaired (HI) listeners.
- Listeners with sensorineural hearing loss (SNHL) may display broadened auditory filter tuning and/or abnormal suppression. This would broaden the effective bandwidths of suprathreshold spectral peaks in speech.
- Given that broader spectral peaks produce larger SCEs (Stilp *et al.*, 2015), we predict that listeners with SNHL will exhibit larger SCEs than NH listeners in a vowel categorization task.

METHODS

Participants

- HI: 14 listeners (ages 51-87) with mild to moderate SNHL from the Greater Lafayette community
  - Mean (thick line) and individual audiometric thresholds for test ear shown
- NH: 25 undergraduates with self-reported normal hearing from the University of Louisville
- All were native English speakers



Stimuli

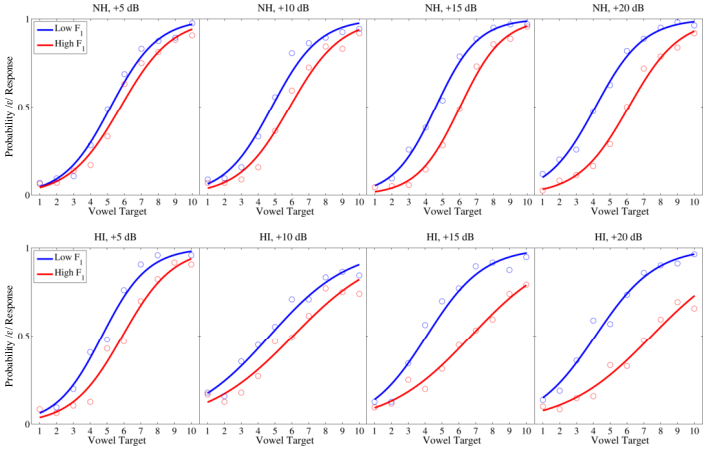
- Precursor sentence: "Please say what this vowel is", spoken by CS (2174 ms) [same as in Stilp *et al.* (2015)]
  - Low F_1 (100-400 Hz) or high F_1 (550-850 Hz) region emphasized by +5, +10, +15, or +20 dB using an FIR bandpass filter
- Target vowels: 10-step series of resynthesized natural tokens varying from /ɪ/ to /ɛ/, spoken by CS (246 ms) [same as in Stilp *et al.* (2015)]
- For HI listeners, stimuli were linearly amplified offline by a hearing aid simulator (Alexander & Masterson, 2015) to scale the output to levels prescribed by Desired Sensation Level *m*(I/O) algorithm v5.0a

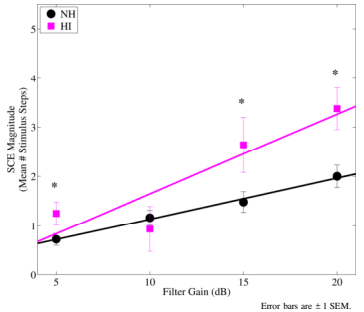
Procedure

- Each trial presented a precursor sentence then a vowel target (50-ms ISI) monaurally (SNHL) or diotically (NH) over circumaural headphones
- NH listeners completed 160 trials/level of filter gain, HI listeners did 200

ANALYSES

- Results from 5 NH and 2 HI listeners were excluded from analyses due to an inability to consistently identify unambiguous vowel endpoints, making the final samples 20 NH and 12 HI listeners.
- For both listener groups, the first two repetitions of each stimulus were treated as practice trials and are not included in analyses.
- For each listener, for each level of filter gain, logistic regressions were fit to vowel identification data associated with each precursor (low- vs. high-F_1 filter peak). 50% points (equal probability of responding "ih" and "eh") were calculated from the regression equations.
- SCE = difference in 50% points across the regressions (*i.e.*, translation of psychometric function along the abscissa).



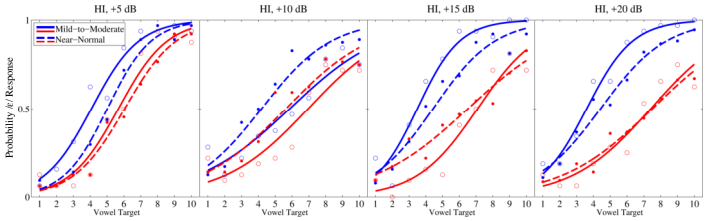


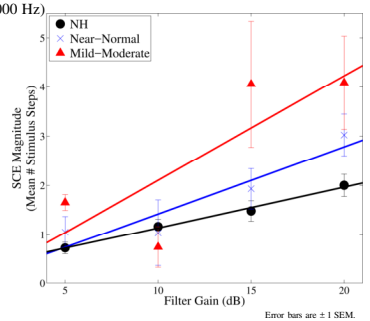
2 (listener group; between-subjects) x 4 (filter gain; within-subjects) mixed ANOVA:

- Main effect of listener group ($F_{1,30} = 7.12$, $p < .025$): HI listeners exhibited larger SCEs than NH listeners
- Main effect of filter gain ($F_{2.42,72.50} = 21.28$, $p < .001$): larger filter gains produced larger SCEs
- Significant interaction ($F_{2.42,72.50} = 4.08$, $p < .05$): HI listeners exhibited larger SCEs at +5, +15, and +20 dB

Audiometric data suggests the 12 HI listeners formed two subgroups:

- n=8 with near-normal low-frequency hearing (< 25 dB HL at 250 & 500 Hz, < 35 dB HL at 1000 Hz)
- n=4 with mild-to-moderate low-frequency hearing loss (25-55 dB HL at 250 & 500 Hz, 35-55 dB HL at 1000 Hz)





Trends suggest that hearing loss in F_1 frequency regions may be responsible for larger SCEs (mild-to-moderate mean SCE = 2.63, near-normal mean SCE = 1.76), but small sample sizes and intersubject variability limit statistical power (independent-samples *t*-test on HI subgroups: $t_{10} = 1.55$, $p = .15$).

DISCUSSION

- SCEs are reported in HI listeners' speech perception for the first time
  - SCEs were larger for HI listeners than for NH listeners
  - Like NH listeners, SCEs increased with larger spectral peaks in the preceding sentence, but grew more quickly for HI listeners
- This extends previous speech perception research with HI listeners, which focused on *intrinsic* cues to vowel identity (*e.g.*, fundamental frequency, F_1, F_2, *etc.*). Here we report HI listeners' sensitivity to *extrinsic* cues to vowel identity (*e.g.*, long-term average spectrum of preceding sounds).
- Results were suggestive of differential processing of spectral context depending on the degree of hearing impairment (near-normal low-frequency hearing vs. mild-to-moderate low-frequency hearing loss), but further study with larger samples is needed to confirm this.
- What are the potential mechanisms behind larger SCEs for HI listeners?
  - Broadened tuning of auditory filters. Broadened filtering would result in broader spectral peaks in the precursor sentence, which produce larger SCEs (Stilp *et al.*, 2015).
  - Steeper growth of loudness in F_1 regions
- Why is it bad to exhibit larger SCEs than NH listeners?
  - If category boundaries are far apart, perception is biased toward one response option.
    - NH listeners correctly labeled vowel target 10 (/ɛ/ endpoint) irrespective of whether the +20 dB peak in the preceding sentence was in low-F_1 or high-F_1 frequencies.
    - HI listeners became less accurate as high-F_1 filter gain increased. Following a +20 dB high-F_1 peak, HI listeners labeled this vowel as /ɛ/ only 65% of the time.
  - Previously unambiguous vowels became more ambiguous when SCEs were overly large, increasing confusions in speech sound categorization.
- Relevance to DSP in hearing aids and cochlear implants:
  - Speech sound recognition by HI listeners is influenced by both short-term and long-term properties of the listening context (Alexander & Kluender, 2009).
  - This argues strongly against the exclusive use of short time constants in hearing aid filtering (*e.g.*, Van Dijkhuizen *et al.*, 1987, 1989).

REFERENCES

Alexander, J. M., & Kluender, K. R. (2009). JSLHR, *52(3)*, 653-670.
Alexander, J. M., & Masterson, K. M. (2015). *Ear Hear*, *36(2)*, e35-e49.
Ladefoged, P., & Broadbent, D. E. (1957). *JASA*, *29(1)*, 98-104.
Stilp, C. E., Anderson, P. W., & Winn, M. B. (2015). *JASA*, *137*(6), 3466-3476.
Van Dijkhuizen, J. N., Anema, P. C., & Plomp, R. (1987). *JASA*, *81*(2), 465-469.
Van Dijkhuizen, J. N., Festen, J. M., & Plomp, R. (1989). *JASA*, *86*(2), 621-628.