Notches in sentence spectra bias subsequent phoneme categorization

Christian E. Stilp (christian.stilp@louisville.edu)
Department of Psychological and Brain Sciences, University of Louisville

UNIVERSITY OF LOUISVILLE

3aPP17

INTRODUCTION

- When spectral properties differ between earlier (context) and later (target) sounds, categorization of later sounds becomes biased through **spectral contrast effects** (**SCEs**; *e.g.*, Ladefoged & Broadbent, 1957).
- Holt and Lotto (2002) and Kluender *et al*. (2003) proposed that SCEs were related to **auditory enhancement effects** (**EEs**; *e.g.*, Schouten, 1940; Viemeister, 1980), where removal and reintroduction of a frequency component makes it more perceptually prominent.
- SCEs and EEs both demonstrate enhanced processing of spectral changes over time, but have been studied separately. Here we report SCEs and EEs in speech perception using the same set of materials.
- In Experiment 1, listeners categorized vowels (/ɪ/-/ɛ/) or consonants (/d/-/g/) following sentences processed by bandpass filters (predicted to produce SCEs) or bandstop filters (predicted to produce EEs).
- In Experiments 2 and 3, listeners categorized phonemes following sentences with spectral notches of varying depths. This tested whether EE magnitudes varied continuously, as SCEs do for sentences with variable-gain spectral peaks (Stilp *et al*., 2015; Stilp & Assgari, 2017).

METHODS

Participants

- 53 native English speakers with no known hearing impairments (n=15, 18, 20 in Experiments 1, 2, 3, respectively)

Stimuli (all presented at 70 dB SPL)

1. Vowel Experiments (Experiments 1 & 2)

- Context = "Please say what this vowel is" (2174 ms)
  - Bandpass/bandstop FIR filters applied to low-F_1 (100-400 Hz) or high-F_1 (550-850 Hz) regions
- Targets = 10 natural vowels from low-F_1 /ɪ/ to high-F_1 /ɛ/ (246 ms)
- Same context and vowel targets as tested in Stilp *et al.* (2015)

2. Consonant Experiments (Experiments 1 & 3)

- Context = "Correct execution of my instructions is crucial" (2200 ms)
  - Bandpass/bandstop FIR filters applied to low-F_3 (1700-2700 Hz) or high-F_3 (2700-3700 Hz) regions
- Targets = 10 natural CVs varying from high-F_3-onset /da/ to low-F_3-onset /ga/ (365 ms; from Stephens & Holt, 2011)
- Same context and CV targets as tested in Stilp & Assgari (2017)

Procedure

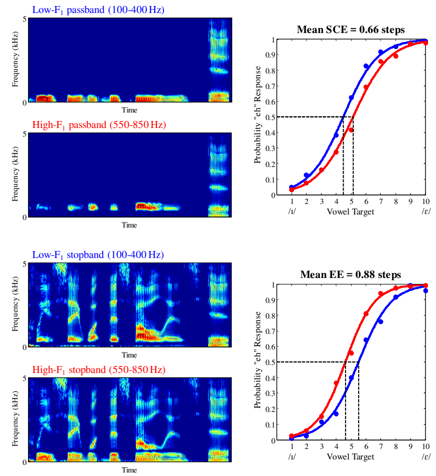
1. Practice on endpoint stimuli with feedback (80% correct required)
2. 4 blocks in counterbalanced orders; trials in random orders
   - Trial Structure: sentence, 50-ms ISI, then the target phoneme

Measuring Context Effects

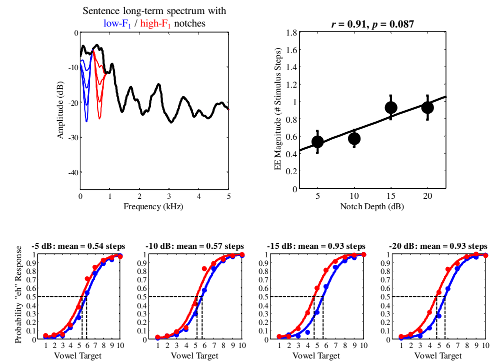
- In each block, logistic regressions were fit to each listener's responses to target phonemes following each context sentence
- 50% points were calculated for each regression and converted into stimulus step numbers (1-10, interpolated as needed; see dashed lines)
- Context effect = average number of stimulus steps separating the 50% points across logistic functions
- Group data are shown at right (and are consistent with mean data)

RESULTS

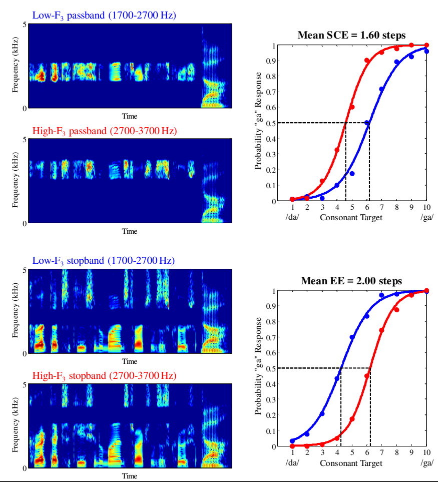
EXPERIMENT 1: VOWELS
Contrast & Enhancement Effects



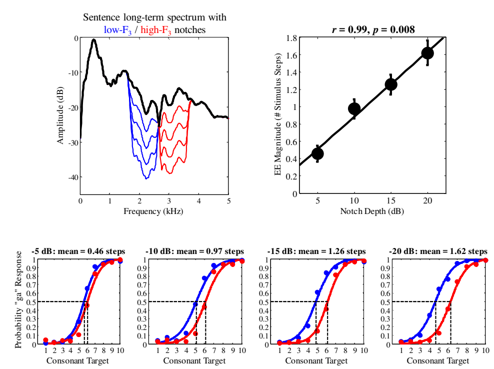
EXPERIMENT 2: VOWELS
Enhancement Effects from Variable-Depth Notches



EXPERIMENT 1: CONSONANTS
Contrast & Enhancement Effects



EXPERIMENT 3: CONSONANTS
Enhancement Effects from Variable-Depth Notches



DISCUSSION

- Passbands from context sentences produce SCEs (Experiment 1); stopband (Experiment 1) and notch filtering of context sentences produce EEs in the complementary direction (Experiments 2, 3). This replicates and extends work by Summerfield *et al.* (1984; 1987), as longer, more acoustically variable contexts produced EEs in both vowel and consonant categorization.
- EE magnitudes in vowel categorization (Experiment 2) were not strongly linearly related to notch depth. This is unlike Stilp & Alexander (2016), where SCEs were highly linearly related to the +5 to +20 dB filter gains used to add spectral peaks to the context sentence. Both studies used the same FIR filters but differed in gain (attenuation vs. amplification).
- EE magnitudes in consonant categorization (Experiment 3) were strongly linearly related to notch depth. In fact, effect magnitudes were extremely similar to SCE magnitudes in Stilp & Assgari (2017) following sentences with +5 to +20 dB spectral peaks. Again, studies used the same FIR filters but differed in attenuation / amplification.
- Why the difference? Experiment 3 used larger notch widths (1000 Hz) at higher frequency regions (1700-2700 / 2700-3700 Hz). Most EE research is at these relatively higher frequencies. Experiment 2 used narrower notches (300 Hz [but wider in ERBs]) at lower frequency regions (100-400 / 550-850 Hz). This might have resulted in asymmetric (adaptation of) suppression due to no energy being below the low-F_1 notch. 50% points on these curves were constant across notch depths, unlike high-F_1-notch data or consonant data.
- Results are consistent with Coady *et al.* (2003) and Holt (2006), who reported EE-type results in phoneme categorization following contexts with "complementary" spectrograms (energy replaced with silence, silence replaced with harmonic spectra or noise).
- Overall sentence level did not alter EEs. Experiments 2 and 3 were replicated when sentences had variable presentation levels (70 dB SPL before notch filtering, then level not reset to 70 dB SPL after).
- Results are consistent with adaptation-related mechanisms: bandpass sentences producing adaptation which resulted in SCEs; bandstop and notch-filtered sentences producing adaptation of suppression / inhibition which resulted in EEs.

REFERENCES

1. Coady JA, Kluender KR, Rhode WS (2003) *JASA*, *114(4)*, 2225-2235
2. Holt LL (2006) *JASA*, *120(5)*, 2801-2817
3. Holt LL, Lotto AJ (2002) *Hear Res*, *167(1-2)*, 156-169
4. Kluender KR, Coady JA, Kiefte M (2003) *Sp Comm*, *41(1)*, 59-69
5. Ladefoged P, Broadbent DE (1957) *JASA*, *29(1)*, 98-104
6. Schouten JF (1940) *Proc Kon Ned Akad Wetensch*, *43*, 356-365
7. Stephens JDW, Holt LL (2011) *Sp Comm*, *53(6)*, 877-888
8. Stilp CE, Alexander JM (2016) *POMA*, *26*, doi:10.1121/2.0000233
9. Stilp CE, Anderson PW, Winn MB (2015) *JASA*, *137(6)*, 3466-3476
10. Stilp CE, Assgari AA (2017) *JASA*, *141(2)*, EL153-EL158
11. Summerfield Q, Haggard M, Foster J, Gray S (1984) *P&P*, *35(3)*, 203-213
12. Summerfield Q, Sidwell A, Nelson T (1987) *JASA*, *81(3)*, 700-708
13. Viemeister NF (1980) in *Psychophysical, Physiological and Behavioural Studies in Hearing*, 190-198.