Short-term, not long-term, average spectra of preceding sentences bias consonant categorization

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

INTRODUCTION

Perception of a given 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.

METHODS

Participants

• 17 native English speakers with no known hearing impairments

Procedure

• Practice: 20 sentences from the AdBrio corpus (Spahr et al., 2012) paired with endpoint consonants; >80% categorization accuracy needed to continue to test
• Test: 160 trials (in random order) in each of four blocks (illustrated below: presented in counterbalanced orders)
  1. Two blocks presented unfiltered context sentences; the other two blocks presented filtered contexts with MSDs that matched unfiltered sentence MSDs
  2. Trial structure: sentence, 50 ms ISI, then target CV which listeners identified as “da” or “ga” (see schematic in introduction)

SCE

• For each block, measured as the mean number of stimulus steps separating 50% points on logistic regressions fit to responses following each context sentence
• Group data are shown below, which are consistent with the mean SCEs listed in each figure title (all of which were significantly greater than zero)

RESULTS

Early vs. Late, Unfiltered

“Father forgot the bread” (2200 ms) from TIMIT
• Positive Late MSD = more high-F3 /d/ responses
• Negative Late MSD = more low-F3 /g/ responses
• Early MSDs had no influence on performance, whether they exceeded ±10 dB (Early vs. Late) or were < ±10 dB
• Early vs. Late Filtered tested against Nothing vs. Late
  • Early vs. Late Filtered tested against Nothing vs. Late
  • Early vs. Late Filtered tested against Nothing vs. Late Filtered in paired t-test: t(16) = 0.76, p = 0.46
• MSDs of entire unfiltered sentences cannot predict these results
• Early vs. Late: entire-sentence MSDs were large but of the opposite sign of Late MSDs (“She looked in the mirror” MSD = −10.69; “The family bought a house” MSD = 8.39)
• Nothing vs. Late: entire-sentence MSDs were extremely small (“Father forgot the bread” MSD = 0.60, “A tree fell on the house” MSD = −2.46), yet these materials biased performance to a similar degree as the Early vs. Late Unfiltered stimuli

Early vs. Late, Filtered

“Correct execution of my instructions is crucial” (2200 ms) from TIMIT
• Unfiltered
  • t = 0.44, p = 0.67
  • Two blocks presented unfiltered context sentences; the other two blocks presented filtered contexts with MSDs that matched unfiltered sentence MSDs
  • Early vs. Late: entire-sentence MSDs were large but of the opposite sign of Late MSDs (“She looked in the mirror” MSD = −10.69; “Father forgot the bread” MSD = 0.60, “A tree fell on the house” MSD = −2.46), yet these materials biased performance to a similar degree as the Early vs. Late Unfiltered stimuli

DISCUSSION

• Unfiltered context sentences produced SCEs that biased consonant categorization, consistent with Stilp and Assgari (2018)
• Late MSDs predicted performance
  • Positive Late MSD = more high-F3 /d/ responses
  • Negative Late MSD = more low-F3 /g/ responses
• Early MSDs had no influence on performance, whether they exceeded ±10 dB (Early vs. Late) or were < ±10 dB (Nothing vs. Late)
• Early vs. Late Unfiltered tested against Nothing vs. Late
  • Early vs. Late Filtered tested against Nothing vs. Late
  • Early vs. Late Filtered tested against Nothing vs. Late Filtered in paired t-test: t(16) = 0.76, p = 0.46
• MSDs of entire unfiltered sentences cannot predict these results
• Early vs. Late: entire-sentence MSDs were large but of the opposite sign of Late MSDs (“She looked in the mirror” MSD = −10.69; “The family bought a house” MSD = 8.39)
• Nothing vs. Late: entire-sentence MSDs were extremely small (“Father forgot the bread” MSD = 0.60, “A tree fell on the house” MSD = −2.46), yet these materials biased performance to a similar degree as the Early vs. Late Unfiltered stimuli
• SCEs were numerically smaller in Unfiltered conditions than in Filtered Conditions, but these differences were not statistically significant (t-tests < 1, p > 0.33)
• This difference was significant in Stilp and Assgari (under review), but that was across eight vowel categorization experiments. The comparison here is likely underpowered
• Variability in duration, phonetic context, and many other properties across unfiltered sentences likely contribute
• Results deviate from time course work by Holt (2005, 2006); particularly with Early MSDs failing to nullify the influence of Late MSDs
• Sizeable differences in how speech versus nonspeech (tone) contexts sample frequency regions over time
• MSDs in the last 500 ms of context sentences were a poor predictor of vowel categorization in Stilp and Assgari (under review)
• Are the present results specific to consonant (d/-)/g/) categorization? Parallel research examining Early/Late windows and vowel (i/-)/e/) categorization needed

REFERENCES

5. Spade A; et al. (2012) Ear Hear 33(2), 112-117
7. Stilp CE, Assgari AA (under review) Attn Percept Psychophphys
8. Stilp CE, Assgari AA (2017a) JASA, 141(2), 1140-1158
9. Stilp CE, Assgari AA (2017b) JASA, 142, 2707