Talker Normalization and Acoustic Properties Both Influence Spectral Contrast Effects in Speech Perception

4pSC3

UNIVERSITY OF LOUISVILLE

Ashley A. Assgari and Christian E. Stilp

Department of Psychological and Brain Sciences, University of Louisville

INTRODUCTION

Spectral contrast effects occur when the auditory system perceptually magnifies spectral differences between the preceding acoustic context and a subsequent target sound.

There is an ongoing debate as to whether talker information influences spectral contrast effects.

- Research using sine tones as acoustic context has shown that talker information is not necessary or sufficient to produce spectral contrast effects (Holt, 2005; 2006; Laing *et al*., 2012).
- However, sine tones cannot represent the acoustic complexity and extreme variability of natural speech and are thus not ecologically valid.

There has been limited research investigating the role of acoustic variability in spectral contrast effects (Watkins, 1991).

- Typically investigations of spectral contrast present the same precursor on every trial.

The present experiments investigate the roles of talker and acoustic variability on spectral contrast effects using natural speech signals.

- Generalize this approach to large (+20 dB) and modest spectral peaks (+5 dB), as both produce spectral contrast effects (Stilp *et al*., in press).

METHODS

Sentences

- One Talker / One Sentence (2174 ms)
  - "Please say what this vowel is" used in Stilp *et al*. (in press)
- 200 Talkers / 200 Sentences (TIMIT database, *M* = 2247 ms)
  - 100 sentences comparable to One Talker / One Sentence in Low-F_1 (100-400 Hz) range, another 100 sentences comparable in High-F_1 (550-850 Hz) range
- One Talker / 200 Sentences (HINT database, *M* = 1739 ms)
  - 100 sentences comparable to One Talker / One Sentence in Low-F_1 (100-400 Hz) range, another 100 sentences comparable in High-F_1 (550-850 Hz) range

Vowels

- Same stimuli as used in Stilp *et al*. (in press)
- Natural vowels interpolated from [ɪ] to [ɛ] using PRAAT (246 ms)
- [ɪ] endpoint: $f0$=100 Hz, F_1 = 400→430 Hz, F_2 = 2000→1800 Hz
- [ɛ] endpoint: $f0$=100 Hz, F_1 = 580→550 Hz, F_2 = 1800→1700 Hz

Filters

- Reliable spectral peaks added to sentences using 300-Hz-wide bandpass filter (Low-F_1: 100-400 Hz, High-F_1: 550-850 Hz)
- Experiment 1: +20 dB filter gain, Experiment 2: +5 dB

Participants

- Native English speakers with normal hearing
- Experiment 1: n= 16, Experiment 2: n=14

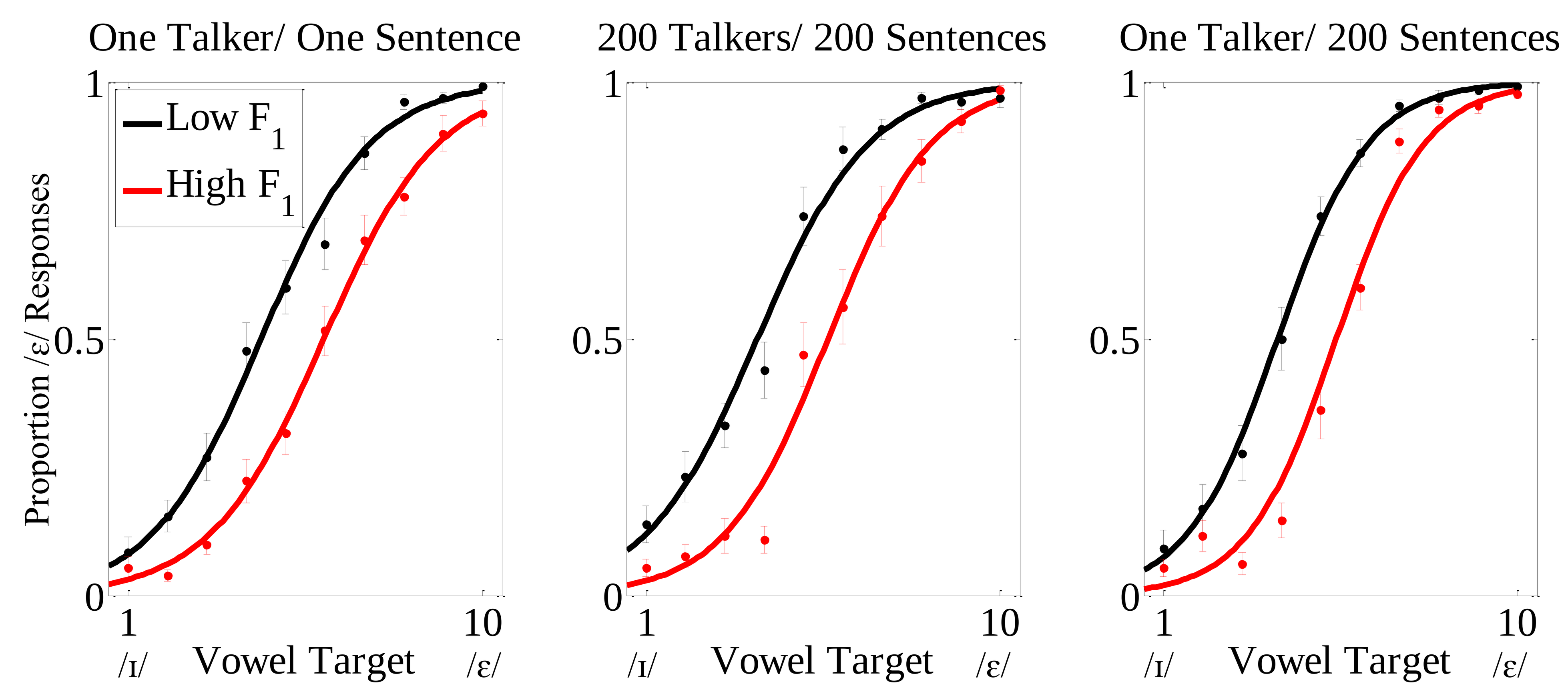
Procedure

- Sentence-vowel pairs presented diotically at 70 dB SPL via circumaural headphones in sound-isolating booths.
- Conditions were blocked and randomized across participants.

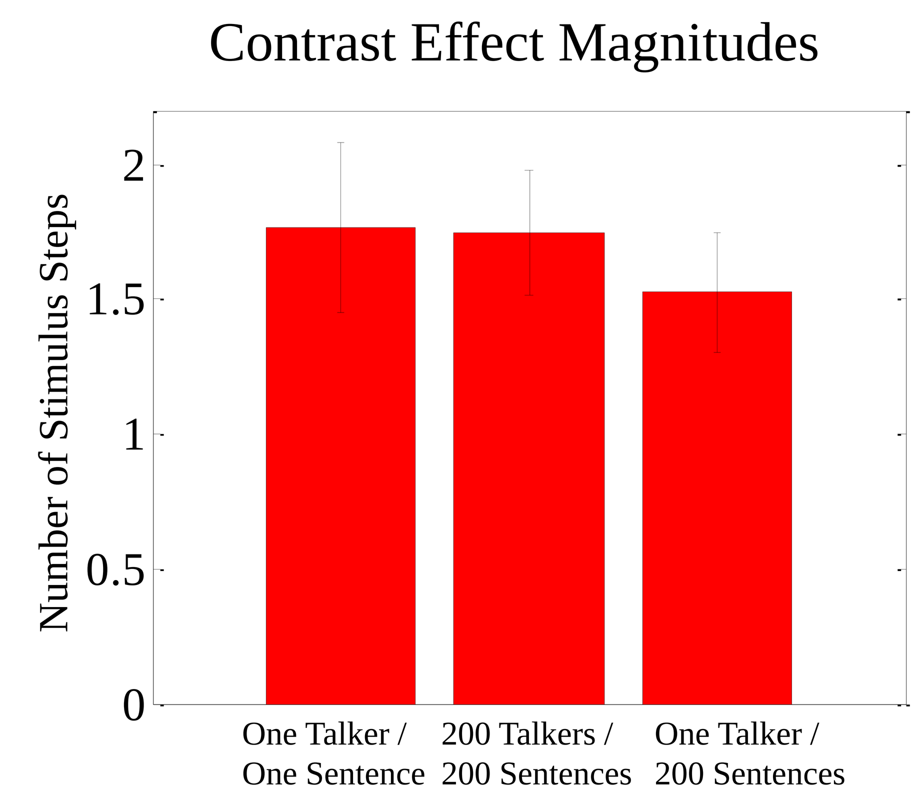
ANALYSES

- Performance criterion = mean of 80% accuracy on vowel continuum endpoints (n=3 removed from Experiment 1, n=3 removed from Experiment 2).
- Logistic regressions were fit to each listener's identification curves. Midpoints were calculated from these regressions.
- Contrast effects magnitudes defined as differences in midpoints between Low-F_1 and High-F_1 functions (*i.e.*, number of stimulus steps along the abscissa).
- Contrast effects were analyzed using paired-sample *t*-tests (Bonferroni correction for multiple analyses within each participant group).
- Error bars in all figures indicate 1 standard error of the mean.

EXPERIMENT 1 (+ 20 dB peaks added to sentences)

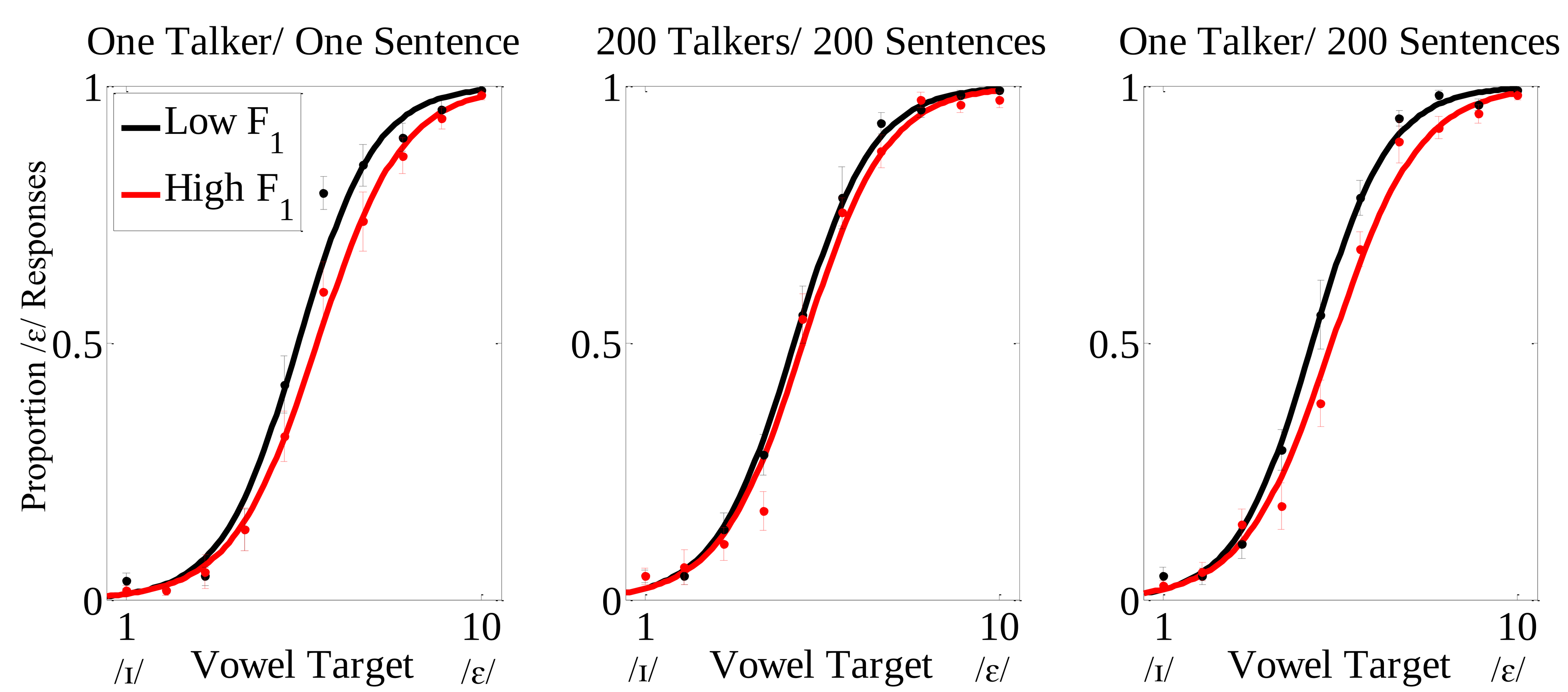


All contrast effect magnitudes greater than 0, one-way *t*-tests: all t_{12}>5.62, p<.001

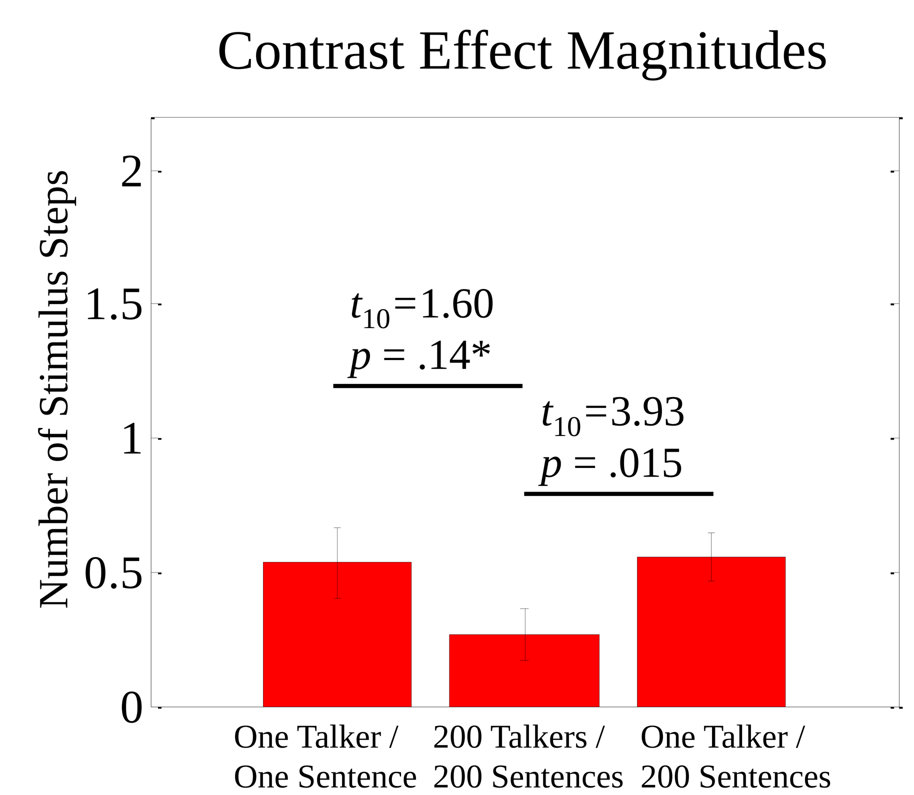


all t_{12} < 1.24, p > .22
Spectral contrast effect magnitudes comparable in all conditions

EXPERIMENT 2 (+ 5 dB peaks added to sentences)



All contrast effect magnitudes greater than 0, one-way *t*-tests: all t_{10}>2.8, p <.018



*Marginally significant result driven by one participant exhibiting contrast effect in the opposite direction in the One Talker / One Sentence condition; if removed, p = .02

DISCUSSION

The influence of talker normalization on spectral contrast effects depends on the magnitude of reliable spectral peaks in the acoustic context.

- When reliable spectral peaks were large (Experiment 1), talker information had no effect on spectral contrast effect magnitudes.
  - Supports general auditory view of spectral contrast
- When reliable spectral peaks were modest (Experiment 2), talker information influenced spectral contrast effect magnitudes.
  - Supports talker normalization in spectral contrast effects

Acoustic variability has no influence on spectral contrast effects as long as acoustic signal emanates from the same source.

- Comparable spectral contrast effect magnitudes across One Talker / One Sentence and One Talker / 200 Sentences
- Spectral contrast effects maintain when acoustic context is highly uncertain

Talker normalization is traditionally evident as faster response times and/or greater accuracy in single-talker conditions.

- Spectral contrast is not captured by accuracy and/or response times.
- Could spectral contrast be a new, low-level metric for measuring talker normalization?

Listeners could be "learning" the talker in single-talker conditions, leading to talker familiarization effects (higher accuracy for familiar over novel talkers; Nygaard *et al*., 1994)

- Learning of talker could have occurred without explicit instructions to pay attention to talker identity

Acoustic properties of the talker determine listeners' ability to process word and talker information independently

- With large reliable spectral peaks, word and talker processed independently (no talker effects in Experiment 1)
- With modest reliable spectral peaks, word and talker processed simultaneously (talker normalization in Experiment 2)
  - Supports Mullenix and Pisoni (1990)

REFERENCES

Holt, L. L. (2005) Temporally nonadjacent nonlinguistic sounds effect speech categorization. *Psych. Sci.*, *16*(4), 305-312.

Holt, L. L. (2006) The mean matters: Effects of statistically defined nonspeech spectral distributions on speech categorization. *J. Acoust. Soc. Am.*, *120* (5), 2801-2817.

Laing, E. J. C., Liu, R., Lotto, A. J., & Holt, L. L. (2012). Tuned with a tune: Talker normalization via general auditory processes. *Front. Psychol.*, *3*: 1-9. doi:10.3389/fpsyg.2012.00203

Mullennix, J. W., & Pisoni, D. B. (1990) Stimulus variability and processing dependencies in speech perception. *Percept. Psychophys.*, *47*(4), 379-390.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker contingent process. *Psych. Sci.*, *5*(1), 42–46.

Stilp, C. E., Anderson, P. W., & Winn M. B. (in press, 2015). Predicting contrast effects following reliable spectral properties in speech perception. *J. Acoust. Soc. Am.*

Watkins, A. J. (1991) Central, auditory mechanisms of perceptual compensation for spectral-envelope distortion. *J. Acoust. Soc. Am.*, *90* (6), 2942-2955.