The Acoustics of Conversational Fricatives in Mid-Western American English

Benjamin V. Tucker, Viktor Kharlamov and Daniel Brenner

This present study examines the acoustic dimensions of fricative reduction in mid-Western American English. Our goals are to determine (i) which acoustic parameters signal the identity of the fricatives and to what extent, (ii) how the discriminative utility of these parameters may vary in conversational and read speech, and (iii) the extent to which speakers demonstrate idiosyncratic reduction schemes.

We investigate these questions using an extensive list of acoustic measures drawn from prior work with laboratory data (Jongman et al., 2000; McMurry & Jongman, 2011). The measures include duration, center of gravity, skew, kurtosis, peak power, etc. Conversational data comes from the Buckeye Corpus of Conversational Speech (Pitt et al., 2005) that contains interview responses gathered in the context of conversation from 40 speakers (not including the interviewers). Laboratory speech comes from the TIMIT Corpus (Zue, 1988) which is composed of 10 read sentences each from a large number of speakers (e.g., in the three dialect regions of interest, there are 2,200 sentence recordings from 220 different speakers). To ensure dialectal comparability (the Buckeye Corpus is drawn exclusively from the Ohio area), TIMIT recordings come from speakers in three of eight TIMIT regions centering around Ohio (regions 2 - 4). In all, more than 100,000 fricative tokens are surveyed, both in terms of the stylistic task-based difference between conversation and careful read speech, and also in terms of the individual speakers and their speaking styles.

Preliminary results suggest that speakers produce a wide range of acoustic variation and that acoustic cues differ depending on the speech style and the individual. As individual speech cues reduce (e.g. the intensity dips less in stop closures (Warner & Tucker, 2011)), speakers appear to make use of large constellations of cues to convey their message. We also discuss the challenges of unbalanced data in corpus-based research with spontaneous speech.

References

Jongman, A., Wayland, R., and Wong, S. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108(3):1252–1263.

McMurray, B. and Jongman, A. 2011. What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*, 118(2):219–246.

Pitt, M. A., Johnson, K., Hume, E., Kiesling, S., and Raymond, W. 2005. The Buckeye corpus of conversational speech: labeling conventions and a test of transcriber reliability. *Speech Communication*, 45:89–95.

Warner, N., & Tucker, B. V. (2007). Categorical and Gradient Variability in Intervocalic Stops. In *Talk given at the annual meeting of the Linguistic Society of America*.

Zue, V. and Seneff, S. 1988. Transcription and alignment of the TIMIT database. In *Proceedings of the 2nd Meeting on Advanced Man – Machine Interface through Spoken Language*, pages 11.1–11.10.