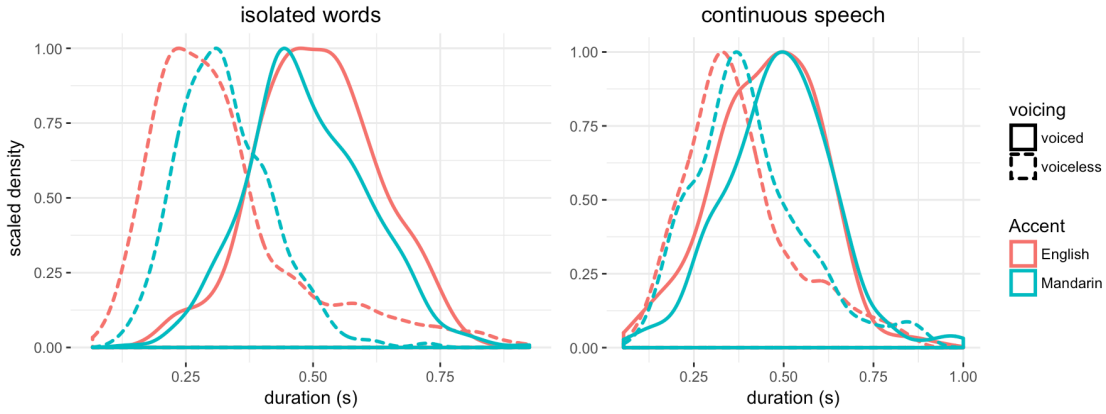
# A CORPUS OF NATIVE AND NON-NATIVE SPEECH FOR SPEECH PRODUCTION RESEARCH

Ruolan Li, Xin Xie & T. Florian Jaeger (University of Rochester)
r.li@rochester.edu

To date, controlled comparisons between native- and non-native speech are largely based small samples, partly due to a paucity of non-native speech database. We introduce a corpus of speech recordings from general American English (AE) and Mandarin-accented English (ME). Both isolated and continuous speech is available from all speakers. Our aim is to create an annotated corpus of comparable speech data from a large sample of individual speakers. Here we wish to inform the research community of the availability of such a database. We begin by providing a description of the speakers and the recording materials. Then, we present some preliminary analyses to illustrate the kinds of questions that can be investigated with this corpus.

**Speakers and speech materials.** The corpus currently contains recordings from 30 speakers, with 15 speakers (10 male and 5 female) in each accent group. All AE speakers speak a Northeastern dialect of American English. All Mandarin-accented speakers are late L2 learners of English who acquired English in mainland China. Speakers read two lists that contained isolated words and continuous speech in a laboratory setting. These lists have been widely used in speech perception studies.[1,2] The isolated word list contains 180 monosyllabic words sampling the entire English phonetic inventory. Specifically, this word list includes minimal pairs (e.g., tap and tab) that are confusing when read by a ME speaker. The continuous speech list contains 80 sentences, divided equally into 5 sets of phonetically balanced sentences. Each speaker is instructed to read each word (or sentence) three times. Additional recordings are made in case of mispronunciations or disfluencies.

**Acoustic analyses.** In this report, we present preliminary analyses on the production of word-final stops in AE and ME. We examine how voicing is distinguished by AE and ME speakers, and whether accent-specific patterns are similarly present in both isolated and continuous speech. For each speaker, we annotated and hand-measured 86 stop-final words from isolated words, and 73 words from continuous speech. Our analysis focuses on three durational measures: duration of vowel, closure, and burst. Past research shows that word-final voicing in stops is signaled by longer vowel, shorter closure and shorter burst.[3] In particular, while vowel and closure are salient cues for distinguishing voiced tokens from voiceless tokens in AE, such contrast is diminished in ME. Previous work has exclusively focused on isolated speech. Extending prior research, our data suggest that for all three acoustic cues (vowel, closure and burst), the degree of separability of the voicing contrasts is affected by accents (AE vs. ME) and types of speech (isolated vs. continuous). Fig. 1 shows the distribution of vowel duration as an example (see figure). Overall, AE speakers make greater separation between voiced and voiceless stops, compared with ME speakers. This pattern is retained in continuous speech, albeit to a lesser degree. This result lends support for to the validity of generalizing findings from isolated speech to continuous speech.




References: 1. Bench, J., & Bamford, J. (Eds.). (1979). 2. Weil (2001). JASA. 3. Flege et al. (1992). JASA.