Modeling Voice Onset Time in English: Factors and their cross-speaker variability

Jeff Mielke and Kuniko Nielsen

Voice Onset Time (VOT) is an important characteristic of stop consonants that plays a large role in perceptual discrimination in many languages, and is widely used in phonetic research. However, it has not been modeled in a phonetically comprehensive manner. The current paper aims to present an analysis of VOT in English in terms of a model of phonetic knowledge. Previous research has shown that VOT in English voiceless stops is sensitive to place of articulation, to contextual factors such as the height, tenseness, and duration of the following vowel, the voicing of coda consonants, and speaking rate. We analyzed 120 /p/- and /k/-initial words produced by 123 Canadian English speakers (n = 17742). VOTs of the initial stops were measured semi-automatically and all other segment durations were measured using forced alignment. The results of a mixed-effects regression support earlier findings that VOT is longer in /k/, directly related to following vowel duration, inversely related to speech rate, longer before tense vowels, and shorter before voiceless codas. Additionally, we find that VOT is shorter when the following syllable starts with a voiceless obstruent (the effect is greater for plosives than fricatives), and that the most relevant measure of vowel duration includes the duration of postvocalic liquids, even those that are typically analyzed as onsets. The effect of speech rate was shown to vary across speakers, while the effect of voiceless coda was relatively consistent across speakers.