

A comparison of hand-measured and automated speech rate measurements in the authoritative speech of Marine Corps Drill Instructors

Speech rate, while an important property of human speech with implications ranging from automatic speech recognition systems to second language fluency testing, has seldom been measured due to the time-consuming nature of the transcription and measurement tasks (de Jong & Wempe 2009). Further, there seems to be both disagreement on the most accurate method by which speech rate is calculated, be it phonemes per unit of time (cf. Cucchiarini, Strik, and Boves 2002) or syllables per unit of time (cf. Kormos and Dénes 2004), as well as whether the goal of the accuracy is subjectivity, and therefore more representative of human perception, or objectivity, and therefore more representative of a static truth. While the concept of objectivity is appealing in the speech sciences, how a speaker is perceived sociophonetically cannot be overlooked in favor of objectivity for objectivity's sake.

In a sociophonetic study of the Command Voice of Drill Instructors (hereafter DI) in the U.S. Marine Corps, four speech styles, two authoritative and two non-authoritative, were identified, recorded, and acoustically analyzed for six speakers—three female and three male, at the DI School in Parris Island, SC (Hicks Kennard 2006). Because of the importance of the “sense of urgency,” or the DI belief that the way they speak directly affects the speed and accuracy of their recruits’ reactions to orders, measures of speech timing are particularly relevant. Two measures of timing, vowel duration and speech rate, were used to explore this “sense of urgency.” The data were analyzed using Praat (Boersma & Weenik 2001). Vowel duration measurements were taken on the first 200 non-high stressed vowels in each speech style for each speaker. Speaking rate measurements were obtained via syllable counts per second from breath groups from the same data that the vowel duration measurements were taken, totaling approximately one minute of continuous speech in each speech style.

For vowel duration, all three female speakers had the longest vowel duration in the most authoritative speech style, called ‘the teachback’, precisely where the male speakers had their shortest vowel duration. Interestingly, the vowel duration results did not affect speaking rate. Because longer vowels typically cause slower speaking rates, one would expect the teachback style, with its longer vowels for the female speakers, to be their slowest speech style. However, this was not the case for any of the female speakers. Yet for the male speakers, both authoritative speech styles, including the teachback, had the fastest two speaking rates, as well as the shortest average vowel durations.

With the emergence of the Praat script of de Jong and Wempe (2009), which was not available at the time of the analysis of the speech rate of the DI speech data, the teachback data will be reevaluated for speech style, comparing the hand-measured speech rates of the teachback of both female and male speakers with results from de Jong and Wempe’s Praat scripts. Because the teachback speech style is a recitation of a script that must be delivered verbatim in the Command Voice for evaluation purposes, this comparison will shed light on the accuracy and suitability of such automated scripts for speech rate for sociophonetic field data, where human perception and not machine recognition is central to research goals.

References

- Boersma, P. & D. Weenik. 2001. Praat, a system for doing phonetics by computer. *Tech. Rep.* 132, Institute of Phonetic Sciences of the University of Amsterdam.
- Cucchiariini, C., Strik, H., & Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, 111, 2862-2873.
- de Jong, N. H., & Wempe, T. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41, 2. 385-390.
- Hicks Kennard, C. 2006. Gender and Command: A Sociophonetic Analysis of Female and Male Drill Instructors in the United States Marine Corps. Ph.D. dissertation, University of Arizona.
- Kormos, J., & Dénes, M. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145-164.