Speech Mode Manipulation in Stable Dysarthria

Heejin Kim¹, Mark Hasegawa-Johnson²

¹ Beckman Institute, University of Illinois, Urbana-Champaign, USA
² Department of ECE, University of Illinois, Urbana-Champaign, USA
{hkim17, jhasegaw}@illinois.edu
Presentation preference: either oral presentation or poster

Dysarthria, a neuromotor speech disorder, can result from a variety of causes such as traumatic brain injury (TBI), degenerative disease (e.g. amyotrophic lateral sclerosis, Parkinson’s disease) or a congenital disorder (e.g. cerebral palsy). The characteristics of dysarthria may vary, depending upon which brain region is affected, but they generally include deviant speech patterns such as imprecise consonants, reduced vowel space, and a decreased range of pitch and loudness (Darley et al. 1975, Duffy 2005). Speech mode manipulations (e.g., loud speech, slow speech, clear speech) are frequently used as intervention techniques for dysarthria, but little is known about the effectiveness of such treatment especially for speakers with stable dysarthria. Research on the speech mode effect in dysarthria is vital, as findings may provide objective bases for therapeutic decisions. In this work we discuss the feasibility of speech mode manipulations as a mean for enhancing speech intelligibility and improving ASR performance for speakers with cerebral palsy (CP). Discussion will be made by couching evidence from prior studies on stable dysarthria and from the new data in our ASR interface experiment.

Substantial experimental work has been conducted to discover the relationship between acoustic measures and reduced intelligibility in dysarthria. A primary research question concerns the magnitude of acoustic-perceptual changes in loud speech and slow speech modes wherein global loudness and rate changes, respectively, are expected to improve segmental clarity. In clear speech, speakers are not requested to modulate a specific prosodic cue, thus speakers may engage both loudness and rate modulations to produce additive effects; this has not been tested yet. Less research attention has been given to stable dysarthria (e.g., CP, stroke, and TBI), compared to progressive dysarthria associated with progressive disease such as Parkinson’s disease (Palmer and Enderby 2007). Given the converging finding that the effectiveness of speech modes differs depending on the type and severity level of disorder, a question remains largely unanswered for speakers with stable dysarthria. We conducted an ASR interface experiment, where a speaker repeated target words if the interface recognized his or her production incorrectly. Our analyses reveal that a speaker with CP increased loudness and expanded vowel space when he or she repeated the target word. Furthermore, the ASR accuracy was improved for the second repetitions of target words, compared to the first repetitions. Taken together with the evidence that speakers with CP utilize loudness for marking prosodic contrasts (Kim et al. 2010, Le Dorze et al. 1994, Patel 2003), our findings suggest the feasibility of speech mode manipulation to benefit speakers with CP. Last, conflicting findings will be discussed regarding rate reduction effects. These findings together will lay the groundwork for future research.