Dynamic Aspects of Articulating with a Virtual Vocal Tract in Dysarthria

Rationale & Specific Purpose

Broadening our understanding of the components and processes of speech sensorimotor learning is crucial to forwarding methods of speech neurorehabilitation. Haith and Krakauer [1] suggest that, while the bulk of sensorimotor learning research has utilized sensorimotor adaptation paradigms, learning sensorimotor control within an unfamiliar working space (without perturbation) is characteristic of much real human experience and a poignant research consideration. Several researchers have studied limb sensorimotor control using virtual environments to create novel sensorimotor working spaces [2-6]. However, the study of novel sensorimotor learning has yet to be undertaken for the auditory-motor transformations that are important to speech sensorimotor learning.

We present kinematic and acoustic data describing participant efforts to produced phonemes within an unfamiliar articulatory-acoustic working space using a virtual vocal tract. Typically-functioning participants and participants with dysarthria were asked to produce diphthongs using an electromagnetic articulograph (EMA) driven articulatory speech synthesizer to provide participant-controlled auditory feedback. The aim of this work is to characterize performance similarities and differences within and between groups to generate hypotheses regarding future experimental manipulations that may support neurorehabilitation. While typical methods for eliciting involuntary speech sensorimotor adaptations using LPC resynthesis are not viable for many individuals with dysarthria because they require robust a speech-acoustic signal from the participant, the current "virtual vocal tract" method provides a viable alternative for eliciting sensorimotor adaptation that circumvents this limitation [7].

Methods

The NDI Wave EMA system was used to register the real time movements of the tongue, lips, and jaw. Five sensors were attached along the midsagittal plane (two on the dorsal surface of the tongue, one on each lip, and one at the juncture of the central mandibular incisors near the gingival border). Reference sensors corrected for participant head movements. Sensor movements were transformed into control parameters for an articulatory speech synthesizer [8, 9]. The mathematical method for transforming articulator movements was speaker-independent, though speaker-specific calibrations based on differences in the available physical working space within the oral cavity were necessary. The "speech" heard by all participants was indistinguishable, except for differences in vowel quality reflecting idiosyncrasies in articulatory movement and differences in proficiency between participants in reliably controlling the synthesizer.

Four typically-functioning speakers and four speakers with dysarthria have currently completed the experimental protocol. All participants in the dysarthria group met the inclusionary criteria of being survivors of traumatic brain injury with functionally significant speech-intelligibility deficits [10], no indicated deficits in speech discrimination [11], and sufficient sustained attention and short-term auditory memory to participate in the experiment [12]. All participants passed a brief audiometric pure-tone screening to assure typical hearing.

Following a calibration procedure that created a simple linear mapping between articulatory movements and movements of the articulatory synthesis parameters, participants were instructed to use their articulators to control the speech synthesizer to produce repetitions of the diphthongs /e/ and /o/. Formant and articulatory sensor positions and movements during the diphthongs were obtained from the synthesized productions by indexing intervals defined by on-glide steady-state, transition, off-glide steady state sequences. Comparable kinematic and acoustic measures of performance included: on-glide position, average transition speed, peak speed, peak acceleration, and off-glide position.

Results

Results revealed that both typically-functioning speakers and speakers with dysarthria demonstrated the ability to learn a novel articulator-acoustic mapping and produce discriminable speech sounds. Comparisons of kinematic and acoustic measures between groups revealed the participants with dysarthria produced significantly more centralized (reduced range of motion) and variable articulatory positions compared to controls (Figure 1). Moreover, formant transition rates and sensor speeds were reduced on average for the dysarthria group compared to typically-functioning speakers. However, individual results indicated that some controls were significantly more variable with respect to both static and dynamic aspects of articulation than some participants with dysarthria.

Discussion

Taken together, these results suggest that, while talkers with dysarthria may display limits in the articulatory working space and constraints on the dynamics of articulation, it is not necessarily true that typically-functioning talkers are less variable than talkers with dysarthria in learning to articulate with a virtual vocal tract. It is hypothesized that experiment protocols that perturb the articulatory-acoustic mapping by reducing the acoustic sensitivity of articulatory movement may elicit compensatory increases in articulatory speed and range-of-motion for participants with dysarthria.

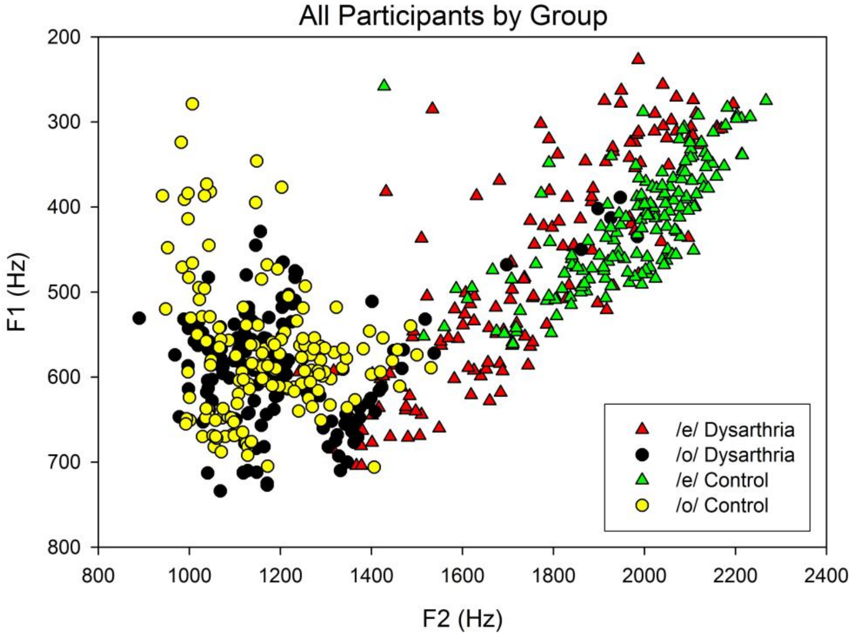

Figure 1: Acoustic data for diphthong off-glides. As a group, participants with dysarthria exhibited greater variability than controls. On average, the dysarthria group also exhibited a more centralized working space than the control group.

References (not included in 750 word limit)

[1] Haith & Krakauer (2013). Model-based and model-free mechanisms of human motor learning. *Advances in experimental medicine and biology*, 782, 1-21.

[2] Mussa-Ivaldi, F. A., Casadio, M., Danziger, Z. C., Mosier, K. M., & Scheidt, R. A. (2011). Sensory motor remapping of space in human–machine interfaces. *Progress in brain research*, 191, 45.

[3] Mosier, K. M., Scheidt, R. A., Acosta, S., & Mussa-Ivaldi, F. A. (2005). Remapping hand movements in a novel geometrical environment. *Journal of neurophysiology*, 94(6), 4362-4372.

[4] Liu, X., Mosier, K. M., Mussa-Ivaldi, F. A., Casadio, M., & Scheidt, R. A. (2011). Reorganization of finger coordination patterns during adaptation to rotation and scaling of a newly learned sensorimotor transformation. *Journal of neurophysiology*, 105(1), 454-473.

[5] Nagengast, A. J., Braun, D. A., & Wolpert, D. M. (2009). Optimal control predicts human performance on objects with internal degrees of freedom. *PLoS computational biology*, 5(6), e1000419.

[6] Sternad, D., Abe, M. O., Hu, X., & Müller, H. (2011). Neuromotor noise, error tolerance and velocity-dependent costs in skilled performance. *PLoS computational biology*, 7(9), e1002159.

[7] Berry, J. J., North, C., Meyers, B., & Johnson, M. T. (2013). Speech sensorimotor learning through a virtual vocal tract. *Proceedings of Meetings on Acoustics*, 19, 060099, 1-8.

[8] Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and speech modelling* (pp. 131-149). Springer Netherlands.

[9] Huckvale, M. (2009). VTDemo – Vocal Tract Acoustics Demonstrator [computer software]. University College London.

[10] Yorkston, K.M., & Beukelman, D.R. (1984). *Assessment of Intelligibility of Dysarthric Speech.* Tigard, Oregon: C.C. Publications, Inc.

[11] Ross, M., Lerman, J., & Cienkowski, K.M. (2004). *Word Intelligibility by Picture Identification – WIPI, 2nd Edition.* St. Louis, MO: Auditec.

[12] Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities.* Itasca, IL: Riverside.