A HIERARCHICAL STATE FEEDBACK CONTROL MODEL FOR SPEECH SIMULATES TASK-SPECIFIC RESPONSES TO AUDITORY AND ARTICULATORY PERTURBATIONS

Benjamin Parrell[1], Vikram Ramanarayanan[2], Sri Nagarajan[3], John Houde[3]
[1]*University of Delaware, Newark, United States of America*
[2]*Educational Testing Service R&D, San Francisco, United States of America*
[3]*University of California, San Francisco, United States of America*

We present a multi-level state feedback based model of speech motor control. As in other state feedback models, the controller governing behavior operates on an internal estimate of the current state of the vocal tract that combines sensory feedback with an internal prediction of the vocal tract state. In our model, this estimate is generated using an extended Kalman filter (Ramanarayanan et al., 2016). In speech, at least some the sensory feedback received by the system is not in the same domain as the output of the controller-feedback is received about speech acoustics, but motor commands must move the speech articulators. We resolve this issue by learning the non-linear mapping between sensory errors and state corrections from a set of simulated vocal tract movements using locally weighted projection regression.

In the current model, a hierarchical structure separates the control of high-level tasks from low-level control of the speech articulators. Currently, we assume tasks to be articulatory gestures, but other tasks such as desired vowel formants can be implemented in the same framework. Importantly, a hierarchical feedback controller allows for redundant control (Todorov et al., 2005), as each task may be completed with multiple configurations of the speech articulators. Such motor equivalence has been demonstrated for speech - when a downward external force is applied to the jaw during production of a /b/, the upper and lower lips compensate for the lower jaw position (Kelso et al., 1984). This response is task-specific and complete, such that bilabial closure is achieved. This contrasts with acoustic perturbations of vowels, where the response is typically a small fraction of the perturbation. The reasons for this incomplete compensation are currently unknown.

Simulations with our model are able to qualitatively reproduce task-specific response to jaw perturbations, with lower and upper lip movements compensating for a lower jaw position only for /b/, but not /z/. Similarly, we are able to reproduce the partial response to acoustic perturbations seen experimentally. However, our results indicate that while compensation is incomplete in the *acoustic* domain, we see complete compensation in the *task* domain. Interestingly, perturbation-induced differences in the estimated position of the *non-relevant* tasks is not corrected for, suggesting that our control scheme agrees with predictions of the minimal interventional principle.

References

Ramanarayanan, V., Parrell, B., et al. (2016). A new model of speech motor control based on task dynamics and state feedback. In I*nterspeech 2016.*

Todorov, E., et al. (2005). From task parameters to motor synergies: A hierarchical framework for approximately-optimal control of redundant manipulators. *Journal of Robotic Systems, 22*(11), 691-710

Kelso, J., et al. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology. Human Perception and Performance, 10*(6), 812-32.