A new model of speech motor control based on task dynamics and state feedback

Benjamin Parrell[1,†], *Vikram Ramanarayanan*[2,†], *Louis Goldstein*[3], *Shrikanth Narayanan*[4], *Srikantan Nagarajan*[4], *and John Houde*[5]

[1]Department of Psychology, UC Berkeley
[2]Educational Testing Service R&D
[3]Department of Linguistics, USC
[4]Department of Electrical Engineering, USC
[5]Department of Radiology, UCSF
[6]Department of Otolaryngology, UCSF

ben.parrell@berkeley.edu, vramanarayanan@ets.org

Abstract

We present a model of speech motor control based on articulatory targets that explicitly incorporates acoustic sensory feedback using a framework for state-based control. We do this by combining two existing, complementary models of speech motor control – the Task Dynamics model [**?**] and the State Feedback Control model [**?**]. We demonstrate the effectiveness of the combined model by simulating a simple formant perturbation study, and show that the model qualitatively reproduces the behavior reported in human subjects, producing online compensation for unexpected perturbations of F1.

Index Terms: speech motor control, auditory feedback, task dynamics, state feedback control, feedback perturbation

1. Introduction

2. Response of model to altered feedback

One of the strongest pieces of evidence that the speech-production system uses acoustic feedback to control ongoing speech comes from studies which perturb the spectral components of speech in real time [**?**, **?**, **?**]. In these studies, subjects repeatedly produce either a single word with an extended vowel or a sustained vowel while listening to feedback of their own voice played back in real time via headphones. On a random subset of trials, their speech is perturbed (either F1 alone or F1 and F2). Subjects compensate somewhat, though not completely, for this unexpected perturbation by shifting their own formants in the opposite direction (e.g. a positive shift in F1 played back to the subjects induces a negative F1 shift in the subjects production). These compensations generally being roughly 200 ms after the onset of the perturbation or, for experiments with continuous perturbation throughout the production of a word, the vowel onset.

These results are generally taken as evidence that vowels have an explicit acoustic target [**?**]. We hypothesized that this compensatory behavior could, alternatively, be produced by a system with articulatory, rather than explicitly acoustic, goals. Although the targets in such a system might be in articulatory space, the actual articulatory state of the system cannot be directly known; rather, the current state must be estimated from 1) the expected outcome of produced motor command and 2) sensory feedback. In such a system, a given motor command issued to achieve a particular articulatory task would generate an **expected sensory expectation**, which could then be compared with incoming sensory feedback. Any discrepancy between the expected and actual sensory feedback would generate a sensory error, which could be used to correct the estimate of the current articulator state.

[†]The first two authors contributed equally to the work.

3. Acknowledgements

4. References