Modeling adaptation in speech motor control

Benjamin Parrell[1*] & Adam C. Lammert[2*]
[1]University of Wisconsin–Madison, [2]Worcester Polytechnic Institute, [*]equal contribution

Introduction

One of the hallmarks of biological sensorimotor control is the ability to change or adapt over time. Such *adaptive control* is evident in human speech production: the speech system readily adapts to gradual changes in vocal tract size over development, to abrupt changes to vocal tract shape (e.g., to dental retainers), to alterations in perceived auditory feedback, and to the presence of perturbations applied to the jaw. Some types of adaptation can occur without learning, such as the compensatory changes in articulatory position seen when the lips or jaw is perturbed, or when speaking with a bite block. However, the majority of these changes need to be learned over the course of repeated speech production.

Developers of the DIVA model[1] have shown how specific adaptive control phenomena can be implemented by adding time-shifted motor commands generated by the auditory and somatosensory feedback control subsystems to the feedforward motor command (weighted by a learning rate parameter). However, such an implementation applies only to optimal control frameworks, like DIVA, that rely on a pre-specified motor trajectory, and is thus unsuitable for models where movements are not pre-specified, but rather emerge from the interplay between task goals and the current state of the body, such as in dynamical systems[2,3] or optimal feedback controllers. Here, we develop and compare three ways of implementing adaptive control in a dynamical systems framework for speech, based upon: tuning dynamical parameters, adding dynamical primitives, and adjusting controlled-space coordinate transformations.

Framework

In dynamical systems models of speech, task-level goals (i.e., articulatory gestures) are typically modeled as point-attractors. These task-level goals are related to lower-level articulator movements in accordance with a control law (here, taken from Task Dynamics[2]):

$$\ddot{\phi} = J^{*}\left(M^{-1}\left[-BJ\dot{\phi} - K(z-g)\right]\right) - J^{*}\dot{J}\dot{\phi} \qquad [1]$$

Where $\ddot{\phi}$ is the articulator-space acceleration of the system state, $\dot{\phi}$ is the current velocity, ϕ is the current position. The task-space position of the system is z, and the Jacobian, J, relates changes in task space to those in articulatory space. The variable g is the target position or spatial goal. M, B, and K are the dynamical parameters that represent, respectively, the inertial, damping, and stiffness coefficients of the gesture.

Given this equation, it is apparent that there are at least three methods to implement motor adaptation. The **first method** is by tuning the dynamical parameters of the gestures, which can have a profound influence on the global movement trajectories that has been utilized in dynamical system control modeling of speech motor control since its inception. The **second method** is by adding a forcing function, f, to Eq. 1 that can influence the existing forces in the control law:

$$\ddot{\phi} = J^{*}\left(M^{-1}\left[-f(z_0-g) - BJ\dot{\phi} - K(z-g)\right]\right) - J^{*}\dot{J}\dot{\phi} \qquad [2]$$

for the initial position z_0. The shape of the forcing function may be specified by Dynamical Movement Primitives (DMPs), which employ a set of state-dependent (i.e., autonomous) Gaussian basis functions that can be flexibly composed and optimized according to a variety of criteria[4]. The **third method** is by adjusting the way that desired task-level changes are transformed into motor commands that control the mobility-space (i.e., articulator-space) state of the plant. This transformation is accomplished through the Jacobian and its inverse J^{*}. Typically, the Jacobian itself and the method of inverting it are taken to be static functions of the system state. However, breaking this static assumption affords an alternative way to model adaptation in this framework, modeled on recent developments in adaptive control in robotics[5].

Method

We explore these alternatives in the context of modeling adaptation to a force-field perturbation applied to the jaw during production of /i/-/æ/[6]. As a proof of concept, we use a highly simplified model of the vocal tract where 1) jaw protrusion and retraction are the dimensions of control (mobility space), 2) the tongue rides passively on the jaw, and 3) task-level goals are taken to be constrictions in the vocal tract (palatal for /i/, pharyngeal for /æ/). We build on recent results that have shown that DMPs can account for the imposed perturbation and return the jaw to a relatively straight-line trajectory, as seen in the pre-perturbation behavior of the model[4] (Fig 2). This was accomplished by optimizing the basis function weighting that shapes the forcing function. In the Trajectory Optimization approach, we estimate the dynamics of the control system and environment to directly estimate the forcing function such that a straight-line trajectory is achieved. In the Effort Optimization approach, we iteratively estimate the kernel weights to minimize the total force generated over the course of the movement. Both optimization approaches yield similar results, suggesting that the DMP approach is applicable to a range of optimization strategies.

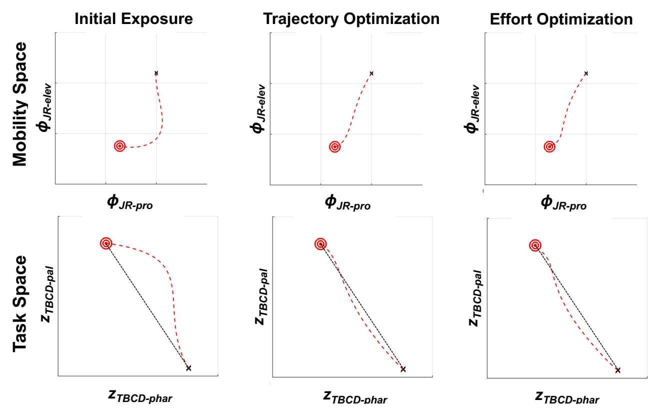

Figure 1: Jaw trajectories for /i/-/ae/ in mobility space (top row) and task space (bottom row). After initial exposure creates deviations from the expected straight-line trajectories (left column), optimization on both the trajectory and effort result in a return to baseline behavior, in agreement with empirical results in human speakers.

DMPs provide one way to model adaptation to perturbations of articulator dynamics in a dynamical-systems control framework. However, it is unclear whether this approach will be sufficient to account for adaptation to perturbations of other types, such as altered auditory feedback, or whether having DMPs directly alter task-level dynamics is desirable from a theoretical standpoint. We compare the DMP approach with the two alternative approaches mentioned above. We test the ability of adaptive dynamical parameter tuning and Jacobian-based controlled-space coordinate transformation adjustments to replicate behavior in a jaw-dependent forcefield demonstrated above. Intriguingly, because updating the Jacobian relies on estimates of the mobility state, it is possible that this approach will also be applicable to sensory perturbations which affect the internal estimate of that state, such as altered auditory feedback for vowel formants.

1. Guenther, F. H. *Neural control of speech*. (The MIT Press, 2016).
2. Saltzman, E. & Munhall, K. G. A Dynamical Approach to Gestural Patterning in Speech Production. *Ecol. Psychol.* **1**, 333–382 (1989).
3. Parrell, B., Ramanarayanan, V., Nagarajan, S. & Houde, J. The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLOS Comput. Biol.* **15**, e1007321 (2019).
4. Parrell, B. & Lammert, A. C. Bridging Dynamical Systems and Optimal Trajectory Approaches to Speech Motor Control With Dynamic Movement Primitives. *Front. Psychol.* **10**, (2019).
5. DeWolf, T., Stewart, T. C., Slotine, J.-J. & Eliasmith, C. A spiking neural model of adaptive arm control. *Proc. R. Soc. B Biol. Sci.* **283**, 20162134 (2016).
6. Tremblay, S., Shiller, D. & Ostry, D. Somatosensory basis of speech production. *Nature* **423**, 866–869 (2003).