MASTER Cognitive Science - 2019-2020

Deep Tongue: speech and predictive control of tongue movements by the brain

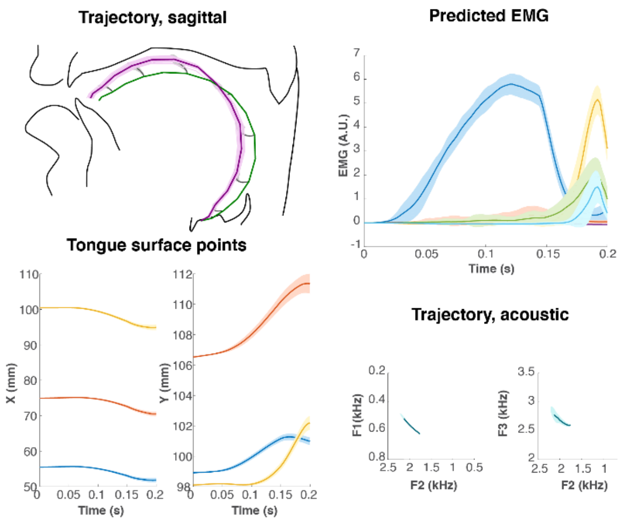
Abstract

One can talk while running, jumping, chewing a gum, or the head upside down, situations that are bound to perturb the dynamics of the articulators that produce speech. Where does this resilience come from? This internship aims at answering this question through modeling. The objective is to propose an explanation for how the brain produces speech in various dynamical conditions and to account for the experimental results obtained so far in this domain. We will test the assumption that the brain implements an internal model of the dynamics of the speech articulators in order to predict their instantaneous state, and integrates this prediction in real time with the noisy and delayed sensory feedbacks in order to ensure the correct pronunciation of sounds in all circumstances, while minimizing effort.

The internship aims at framing this hypothesis in terms of an Optimal Feedback Control model. This work is part of the "*Bayesian Cognition and Machine Learning for Speech Communication*" chair of the *Multidisciplinary Institute of Artificial Intelligence* of University Grenoble Alpes, and may be extended to a PhD funded by the Institute.

Context

Speech production is a task requiring great dexterity. Some sounds like /i/, /s/ or /l/ need a very precise positioning of the tongue in the oral space. Nevertheless, save pathological conditions, every human can accomplish this motor task, from youth to old age. We can speak with precision in variable conditions: walking, running, jumping, eating, lying down, in weightlessness… Besides, speech is a very rapid task, some transitions between sounds lasting a few dozens of milliseconds, the same order of magnitude than the delays affecting sensory and motor signals, and their cerebral processing These delays exclude that speech be purely feedback-controlled in a reactive mode. Rather than simply trying at each moment to zero the discrepancy between the expected and produced



Predictions of an effort-optimal tongue control for the production of a diphthong, i.e. a transition between two wowels (/ə/→/i/)

sounds, it is more likely that the brain predicts the sensory consequences of its actions, while factoring the characteristics of the environment and the likelihood of perturbations.

Furthermore, speech motor control is highly redundant: the brain controls many muscles that need to be coordinated, and there exists an infinite number of ways to achieve a given movement of the articulators.

For these reasons, we hypothesize that the brain works like an optimal feedback controller. In short, optimal feedback control posits that the controller considers 1) the sensory feedbacks which are delayed and noisy; 2) a prediction of the instantaneous state of the production system through an internal model of the dynamics of this system; 3) the minimization of a cost (effort, imprecision…) that ensures a djudicious choice of motor commands, task stability, and optimal correction of motor commands in case of an unexpected external perturbation. With simplifying hypotheses, a formal solution of this problem has been shown (Todorov and Jordan 2002).

The internship will aim at exploring the predictions of such a model applied to speech production, as a function of the assumptions on the cost to minimize, the imprecision of sensory and motor signals, and the available sensory modalities (hearing, proprioception, touch). According to the mathematical skills of the candidate, (s)he could try more or less complex formulations of optimal feedback control. The behavior of these models will then be tested when exposed to perturbations of the auditory feedback or mechanical perturbations of the tongue. These predictions will be compared to the published (or our own unpublished) experimental data on the compensation of unexpected perturbation of speech, and may motivate new experimental work to be carried out in the lab.

This work follows up a PhD thesis in Cognitive Science (2018) that modeled the response to sudden mechanical perturbations of speech production, and a preliminary study presented at the *Society for Neuroscience* annual meeting (see Figure). The internship can lead to a PhD combining modeling and experimental validation; this PhD will be funded by the "*Bayesian Cognition and Machine Learning for Speech Communication*" chair of the Multidisciplinary Institute in Artificial Intelligence (see http://miai.univ-grenoble-alpes.fr).

Contacts

Pascal Perrier, PR Grenoble-INP, Pascal.Perrier@gipsa-lab.fr, 04 76 57 48 25
Pierre Baraduc, CR CNRS, Pierre.Baraduc@gipsa-lab.fr, 04 76 82 71 50

Funding

MIAI chair "Bayesian Cognition and Machine Learning for Speech communication".

References

Todorov E, Jordan MI (2002). Optimal feedback control as a theory of motor coordination. *Nat Neurosci*, *5*(11):1226-1235.

Guigon E, Baraduc P (2008). Optimality, stochasticity, and variability in motor behavior. *J Comput Neurosci.* 24(1):57-68.

Baraduc P, Perrier P (2017) Motor control of the tongue during speech: predictions of an optimization policy under sensorimotor noise. *Society for Neuroscience Annual Meeting*, Washington, USA, and *Speech Units Workshop: Perceptuo-motor relationships in speech communication*, Geneva, Switzerland (2018).