



HAL
open science

Online compensation of auditory feedback perturbations in speech: an optimalfeedback control model of tongue biomechanics

Tsiky Rakotomalala, Pierre Baraduc, Pascal Perrier

► To cite this version:

Tsiky Rakotomalala, Pierre Baraduc, Pascal Perrier. Online compensation of auditory feedback perturbations in speech: an optimalfeedback control model of tongue biomechanics. ISSP 2024 - 13th International Seminar on Speech Production, May 2024, Autrans (Grenoble), France. hal-04738746

HAL Id: hal-04738746

<https://hal.science/hal-04738746v1>

Submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Online compensation of auditory feedback perturbations in speech: an optimal feedback control model of tongue biomechanics

Ny-Tsiky Rakotomalala^{1,4}, Pierre Baraduc^{1,2,4}, Pascal Perrier^{1,3,4}

¹Univ. Grenoble Alpes, ²CNRS, ³Grenoble INP, ⁴GIPSA-lab

name.surname@grenoble-inp.fr

Introduction. Online responses to perturbations of the auditory feedback have been extensively studied since the pioneering study of Houde and Jordan (1998). A striking finding is that subjects tend to rely preferentially on either the somatosensory or the auditory feedback to adjust speech production (Lametti, Nasir, and Ostry 2012). Another remarkable phenomenon is that even modest online compensation tend to show an asymmetry between up- and down-shifts of the speech formants (Hantzsch, Parrell, and Niziolek 2022). Here we present a modeling analysis of these experiments using GEPPETO-OC. This model integrates the neurobiomechanical model of GEPPETO (Payan and Perrier 1997) into a stochastic feedback control framework (Todorov and Jordan 2002), in order to handle high level feedback control during movement execution. This new framework also allows to account for the impact of motor and sensory variability while preserving the idea of optimal effort allocation. We here explore whether the interaction between biomechanical constraints and characteristics of the sensory feedbacks can explain human online audiomotor control.

Methods. Speech goals were considered multimodal: auditory (target region for phonemes defined as ellipsoids in F1/F2 and F3/F2 domains) and proprioceptive (ellipsoids in the 3D space of variables that characterize the tongue shape). For simplicity, the timing of these goal sequence was considered fixed.

GEPPETO-OC does not separate planning and execution: in order to determine the current motor commands, the *optimal controller* computed and adapted the trajectories of the motor commands in real time by constantly minimizing both the neuromuscular effort and the sensory inaccuracy at the phonemic targets as a hybrid cost:

$$C = \sum_{t=t_{\text{current}}}^{T-1} \|[\lambda(t+1) - \lambda(t)]^+\|^2 + \alpha \sum_{g=g_{\text{current}}}^{N_g} \text{dist}(p(T_g), p_g)$$

T is the total movement duration, λ the motor commands, N_g the number of phonemic targets; p_g the specified sensory values at the phonemic targets —together with a requirement of final stability (null velocity); $p(T_g)$ contains the actual sensory values at the time T_g set for each target; α is a trade-off parameter.

Motor and sensory signals were assumed corrupted by additive and multiplicative noises, and feedback was assumed delayed by up to 80 ms. An *optimal estimator* (EKF) estimated the state of the system through an internal model of the biomechanics based on an efferent copy of the motor commands, predicted the sensory reafference, and corrected the state estimate as a function of the sensory prediction error.

The original GEPPETO model controlled a finite-element biomechanical model of the tongue. In GEPPETO-OC, to speed up computation, a reduced model was developed. First, the dimensionality of the upper contour of the tongue, described by the position of 16 nodes in the sagittal plane, was reduced to 5D via an autoencoder. Then a LSTM network model of the plant dynamics was trained on thousands of simulations of the finite element model. Last, auditory feedback was computed from the tongue contour using a harmonic model of the vocal tract (Badin and Fant 1984).

Results. To assess whether across-subject differences in sensory precision could explain differential reliance on somatosensory or auditory feedback, we simulated trajectories aimed at the / ε / target from a neutral tongue position in three different conditions: (1) no auditory perturbation; called reference on Figure 1; (2) the auditory feedback of F1 was up-shifted by 125 *mels* from the onset of the movement; (3) F1 was conversely down-shifted by 125 *mels*.

For these three conditions, we tested two different cases: (a) the auditory signal was more noisy than the somatosensory signal; (b) the somatosensory signal was more noisy than the auditory signal. Each condition was repeated 20 times to compare trial-to-trial variability.

In the (a) case (strong auditory noise), compensation value at the end of the movement is less than 10 *mels* (8%) in upward and downward shift. In the (b) case, compensation value reach approximately 35 *mels* (28%). The dynamics of

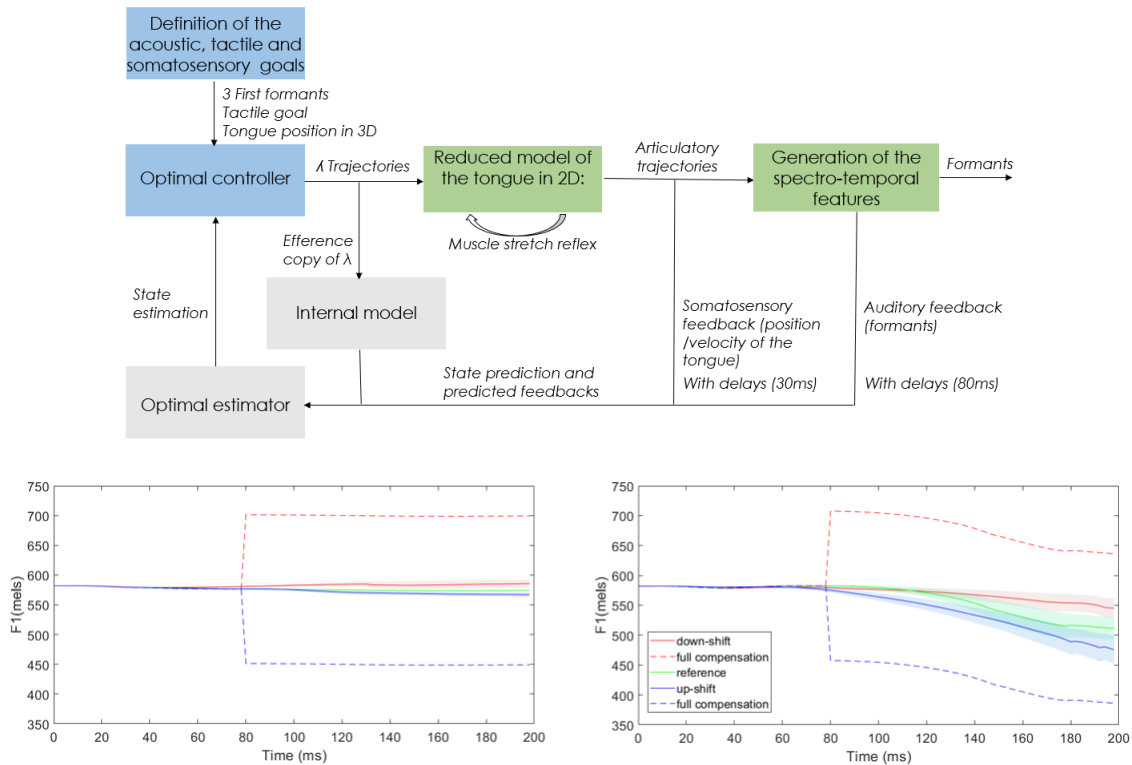


Figure 1: Upper panel : Architecture of GEPPETO-OC. Lower panels: Compensation to auditory perturbation for /ε/: left: higher amount of noise in the auditory feedback; right: higher amount of noise in the somatosensory feedbacks. Dashed lines represent a theoretical full compensation level (subtracting the perturbation from the average reference trajectory).

the compensation were also different between (a) and (b), case (a) being close to linear. Compensation of the up-shift and down-shift were also asymmetrical, but differently so in the two cases studied.

Discussion. Our model could reproduce the sensory preference effects by only assuming a difference in sensory precision. Furthermore, the predicted compensation was modest and larger for down-shifts in setting (a), as found experimentally (Hantzsch, Parrell, and Niziolek 2022); this result is likely due to concurrent optimization of precision and effort. In the "auditory preference" case (b), the dynamics of the compensation clearly reflected the asymmetry between an assistive and a resistive perturbation.

Results could also be analyzed in the kinematic domain (tongue contour motion) and in the motor domain (change in muscle activation patterns). This additional data will be presented during the conference.

Future work should compare the effect of a perturbation on either tongue movement (as here) or on stable vowels (static articulatory positions). This would check whether the model is also able to reproduce the experimental observations of a much slower compensation (at around 460 ms) under the latter experimental condition.

References.

- Badin, P. and G. Fant (1984). "Notes on vocal tract computation". In: *STL QPSR* 2.3, pp. 53–108.
- Hantzsch, Lana, Benjamin Parrell, and Caroline A Niziolek (2022). "A single exposure to altered auditory feedback causes observable sensorimotor adaptation in speech". In: *Elife* 11, e73694.
- Houde, John F and Michael I Jordan (1998). "Sensorimotor adaptation in speech production". In: *Science* 279.5354, pp. 1213–1216.
- Lametti, Daniel R, Sazzad M Nasir, and David J Ostry (2012). "Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback". In: *Journal of Neuroscience* 32.27, pp. 9351–9358.
- Payan, Yohan and Pascal Perrier (1997). "Synthesis of VV sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis". In: *Speech communication* 22.2-3, pp. 185–205.
- Todorov, Emanuel and Michael I Jordan (2002). "Optimal feedback control as a theory of motor coordination". In: *Nature Neuroscience* 5.11, pp. 1226–1235.