Investigation of Tongue Motor Control During Speech Production
A Model Based Study

Saeed Dabbaghchian

Department of Speech, Music, and Hearing, KTH Royal Institute of Technology, Stockholm, Sweden
saeedd@kth.se

1. Introduction

Speech production is probably the most complex motor task which needs a coordinated operation of several tens of muscles. How our brain learns to do such a complex task? This is really a difficult question to answer. Once trying to seek answer for this question, some other critical questions may arise. I will discuss some of these questions briefly and then will state a proposal for my thesis.

2. Nature of Speech Goals

What is the nature of motor goals during the speech production? Is it reaching some articulatory gestures? Is it acoustic or auditory goals?

It has been a debate about the nature of speech goals. Fujimura [1] specified the goals as articulatory movement while other researchers like Fowler [2] proposed vocal tract shape as the speech goal or Stevens [3] who considered spectral characteristic as speech goals. There are several evidences which support or discard some aspects of these hypothesis. A bite-block experiment (for e.g. see [4]) is an evidence which discards the absolute position of articulators as targets. Lip-tube experiment [5] is another interesting study which supports auditory goals and using of orosensory feedback. Another study which supports auditory goals is auditory feedback perturbation experiment by Purcell et al. [6]. This experiment shows that when the auditory feedback is artificially modified, subjects change their articulatory configuration to reach the same formants. Honda et al. [7] conducted an experiment where they showed the role of tactile feedback in speech production by using a time-varying inflated palate. Another study by Tremblay et al. [8] was done which supports the articulation goals. In this study a complex mechanical load alters the subject's jaw while the acoustic output does not change perceivably. For further discussion to refer the reader to [9] where the author concludes: "Speech goals have both articulatory/motor and acoustic/auditory components, but that there is a hierarchy between these two components."

3. Control Variables

How motor cortex and central nerve system control fibres and muscles? Is that possible or even necessary to control a single fibre or the cortex always activate the fibres in a group?

Of course controlling a single fibre creates more flexibility but it also adds lots of redundancy. This question can be answered from two different perspectives : physiology and motor redundancy.

4. Redundancy

Even if we assume, just for simplicity, that the cortex always activate the fibres in a group, and if we agree on the goals of speech task, still there are lots of redundancy in the system. How the brain deals with this redundancy?

The motor task redundancy is not excluded for speech production and almost every motor task has such a redundancy. However we can see more redundancy in speech production task. Lindblom [10] suggested that the speech gestures are selected based on motor "economy" which means that redundancy is removed by minimizing the energy. However Harris et al. [11] proposed another hypothesis based on hand movement experiment. According to this study our brain find a solution to minimize the variability of the movement instead of energy.

5. Thesis Proposal

Using of a model can help scientists to test hypothesis, and to propose new hypothesis by raising new questions. During the past years, some models were developed to study speech production. In order to study motor control, a biomechanical model of speech articulator is needed. Buchaillard et al. [12] developed a 3D tongue model to study speech motor control; and recently another model was developed by Anderson et al. [13]. One of the advantages of the model-based study over subject-based study is that in a model-based study we can do experiments which are difficult or even impossible to do with subjects.

In this thesis, an existing model developed in ArtiSynth [14] will be used to investigate some aspects of speech motor control especially the redundancy problem and two dominant hypothesis about the redundancy reduction, as it was already discussed, will be investigated. The mentioned model has only 11 muscles for the tongue. In order to keep the study possible, regarding computational time, the study will focus only on tongue muscles and one control variable for jaw position. At first step, a parameter space including 11 muscles of the tongue and jaw position will be discretized and a parameter study will be conducted. Then the results will be analyzed by considering energy requirements and movement variation for different solutions which generate the same articulatory movements or similar acoustic output. The result of this study may reveal some aspects of redundancy reduction and may also be used to propose motor control strategy.

A continuous speech is usually considered as a production of a sequence of speech units known as phoneme or syllable. This hypothesis requires a coarticulation model to explain continuous speech (see for e.g. Kent [15]). According to the author's knowledge this problem has not been investigated from the neuromuscular perspective and it's relation with the nature of speech goals and redundancy problem. So, this is also a good research question to pursue in this thesis.

6. References

[1] O. Fujimura, "Relative invariance of articulatory movements: An iceberg model," in *Invariance Var. Speech Process.*, J. S. Perkell and D. H. Klatt, Eds. Psychology Press, 1986, pp. 226–242.

[2] C. A. Fowler, "An event approach to the study of speech perception from a direct-realist perspective," Tech. Rep., 1986.

[3] K. Stevens, "The quantal nature of speech: Evidence from articulatory-acoustic data," in *Hum. Commun. A Unified View*, 1972, pp. 51–66.

[4] T. Gay, B. Lindblom, and J. Lubker, "Production of bite-block vowels: Acoustic equivalence by selective compensation," *J. Acoust. Soc. Am.*, vol. 69, no. 3, pp. 802–10, 1981.

[5] C. Savariaux, P. Perrier, and J. P. Orliaguet, "Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production," *J. Acoust. Soc. Am.*, vol. 98, no. 5, p. 2428, 1995.

[6] D. W. Purcell and K. G. Munhall, "Adaptive control of vowel formant frequency: evidence from real-time formant manipulation." *J. Acoust. Soc. Am.*, vol. 120, no. 2, pp. 966–977, 2006.

[7] M. Honda, A. Fujino, and T. Kaburagi, "Compensatory responses of articulators to unexpected perturbation of the palate shape," *J. Phon.*, vol. 30, no. 3, pp. 281–302, 2002.

[8] S. Tremblay, D. M. Shiller, and D. J. Ostry, "Somatosensory basis of speech production," *Nature*, vol. 423, no. 6942, pp. 866–869, 2003.

[9] P. Perrier and S. Fuchs, "Motor Equivalence in Speech Production," in *Handb. Speech Prod.* Hoboken, NJ: John Wiley & Sons, Inc, apr 2015, pp. 223–247.

[10] B. Lindblom, "Economy of speech gestures," in *Prod. Speech*. New York, NY: Springer New York, 1983, pp. 217–246.

[11] C. M. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," *Nature*, vol. 394, no. 6695, pp. 780–784, 1998.

[12] S. Buchaillard, P. Perrier, and Y. Payan, "A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning." *J. Acoust. Soc. Am.*, vol. 126, no. 4, pp. 2033–2051, 2009.

[13] P. Anderson, N. M. Harandi, S. Moisik, I. Stavness, and S. Fels, "A Comprehensive 3D Biomechanically-Driven Vocal Tract Model Including Inverse Dynamics for Speech Research," in *INTERSPEECH*, 2015, Conference Proceedings, pp. 2395–2399.

[14] "ArtiSynth: A 3D Mechanical Modeling System combining Multibody and Finite Element Simulation (http://artisynth.magic.ubc.ca/artisynth/)."

[15] R. D. Kent and F. D. Minifie, "Coarticulation in recent speech production models," pp. 651–669, 1977.