AN INTEGRATED MODEL OF THE BIOMECHANICS AND NEURAL CONTROL OF THE TONGUE, JAW, HYOID AND LARYNX SYSTEM

Vittorio Sanguineti[1], *Rafael Laboissière*[2] *and David J. Ostry*[3]
[1] DIST, Università di Genova (Italy)
[2] Institut de la Communication Parlée, Grenoble (France)
[3] McGill University, Montréal (Canada)

Abstract

A model of the sagittal plane motion of the tongue, jaw, hyoid bone and larynx is presented, based on the λ version of equilibrium point hypothesis. The focus is on the organization of control signals underlying vocal tract motions. A number of muscle synergies or 'basic motions' of the system are identified. It is shown that systematic sources of variation in a X-ray data base of midsagittal vocal tract motions can be accounted for with six independent commands, each corresponding to a direction of articulator motion. It is further shown that hyoid position and orientation can be predicted from the application of other vocal tract commands and need not be explicitly controlled. The dynamics of individual commands are also assessed. It is shown that the dynamic effects are not neglectable in speech-like movements because of the different dynamic behavior of soft and bony structures.

10.21437/Eurospeech.1997-536

1. INTRODUCTION

Investigations of the physiology of speech production have shown that the oral apparatus – the 'plant' that needs to be controlled by the nervous system to produce speech sounds – has an intricate structure that involves a huge number of mechanical degrees of freedom, and a mechanical behavior that is complex and often counter-intuitive. This suggests that the mechanics of vocal tract structures have to be taken into account when investigating the neural control of speech production.

The bony parts are relatively simple to characterize, but until recently [3] there have been no attempts to understand empirically observed movements of the jaw [6] by means of a dynamic model in which muscle properties, reflexes and neural signals are represented. Moreover, the different vocal tract structures are known to interact mechanically [10] and therefore must be considered as a whole. Attempts to investigate these interactions are however still at a preliminary stage – see, for instance, the work of [2] on the mechanical interaction between tongue and larynx.

In the case of the much better studied movements of the arm, it has been suggested that the viscoelastic properties of muscles and reflexes, in particular the stretch reflex, have a built-in capability to 'compensate' for the effects of inertia, dynamic interaction forces, gravity. This would imply that muscle properties and the peripheral circuitry are important determinants of the dynamic behavior of the body, and therefore needs to be accounted for in formulating a biomechanical model. Reflex mechanisms are often neglected in biomechanical studies but have been explicitly accounted for in our recent models of the jaw-hyoid system [3] and tongue [8]. In the present paper, we report initial simulation results of a sagittal plane biomechanical model of the jaw, tongue, hyoid and larynx systems, which accounts for dynamics at a mechanical and a neuromuscular level.

2. METHODS

The tongue has been modelled by means of Finite Elements (FE) techniques; the bony structures (i.e. the jaw and the hyoid bone) are approximated as rigid bodies. Vertical movements of the larynx which is represented as a point mass are modelled as well; see Fig. 1. The modelled muscles include: genioglossus, hyoglossus, styloglossus, superior and inferior longitudinalis, verticalis, the jaw opener group (geniohyoid + anterior belly of digastric) and masseter, anterior and posterior temporalis, superior and inferior heads of lateral pterygoid, mylohyoid, the hyoid retractors (stylohyoid + posterior belly of digastric), sternohyoid, sternothyroid, and thyrohyoid. The modelled dynamic behavior of muscles and reflexes is based on the Equilibrium-Point (EP) hypothesis of motor control – λ-model [1]; see Fig. 2.

The λ model assumes that neural control signals produce voluntary movement by acting on motoneurone (MN) membrane potentials. The effect at the level of the muscle is to change the threshold muscle length (λ) at which α MN recruitment begins [1]. By changing the values of λs over time, the musculoskeletal system may be caused to move to a new equilibrium position.

EUROSPEECH '97
5th European Conference on Speech
Communication and Technology
Rhodes, Greece, September 22-25, 1997

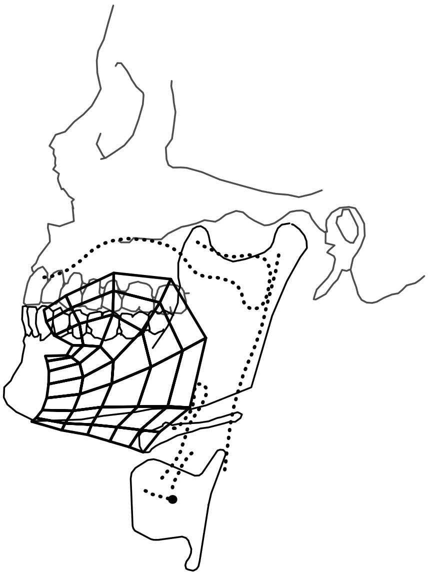

Figure 1: Biomechanical model of the oral cavity

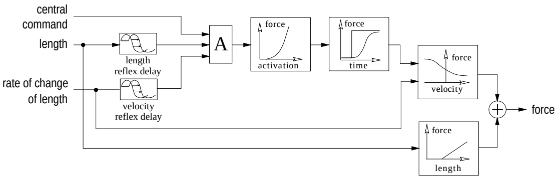

Figure 2: Block diagram of the muscle model

A number of additional assumptions may be made concerning the organization of control signals. The λ model proposes that central control variables can be interpreted as geometric quantities, namely, threshold muscle lengths. In the case of multiple muscle systems, because of their spring-like behavior, the set of λs associated with individual muscles (or muscle compartments) specify an equilibrium configuration for the system. This does not mean that the individual λs are independently controlled. Indeed, control is presumably organized into a relatively small number of different combinations of λ changes, or muscle synergies, each corresponding to an elementary or primitive motor behavior. All possible movements may result from the combination of these basic motions.

How can such muscle synergies or basic motions be identified for the jaw-hyoid-tongue-larynx system? In the case of jaw motion, the observation of a variety of different patterns of coordination between jaw protrusion and rotation has suggested [6] that during speech its mechanical degrees of freedom can be controlled independently.

Moreover, the data of [11] suggest that the observed patterns of motion of the hyoid bone are largely uncorrelated with jaw movements. On the other hand, simulation studies [2] have demonstrated a close mechanical coupling between the hyoid bone, the larynx and the tongue. These findings suggest that the hyoid-larynx system can move independently of the jaw. It is however unclear whether hyoid bone position and orieintation result from active control or are a consequence of mechanical interaction with the surrounding structures.

[5] and [9] have suggested that the tongue system consists of a number of separately controlled muscle groups. In particular, it appears that the tongue tip can move independently of the tongue body. However, unlike jaw movements, there is no a priori basis for the identification of functional degrees of freedom of tongue motions [4, 10].

3. RESULTS

3.1. Determination of independent commands

The commands were derived by fitting the model to a cineradiographic data base and assuming that control signals to muscles are organized to produce maximally independent motions [10]. This resulted in independent commands associated with: jaw protrusion (JP) and rotation (JR), larynx elevation (LY), tongue dorsum (arching/flattening, TD), tongue body (front/back, TB), tongue tip (raised/lowered, TT), that are related to geometric arrangement of muscles; the effects of the individual commands are displayed in Fig. 3.

Moreover, we also found that predicted changes to articulator positions resulting from the commands derived above were found to be largely independent of the initial vocal tract configuration. That is, when a given command was applied at different initial vocal tract configurations, similar changes in configuration were produced. This means that a postural change which results from a combination of the above commands can be interpreted as the combination of the changes elicited by the individual commands.

3.2. Control of the hyoid bone

It is our assumption that hyoid bone motion is not explicitly controlled. To test this possibility, at least for this particular data set, we assessed the proportion of variance in hyoid position and orientation which is accounted for by the six commands shown in Figure 3. The basic idea here was to compare the variance associated with the actual hyoid position and orientation to the residual variance after the application of the first six commands derived from the fitting procedure. There are no direct ways to compute these variance estimates, because in general the three hyoid coordinates (two positions and one orientation) are correlated. To avoid this problem, the raw data were first rotated and scaled to give three indepen-

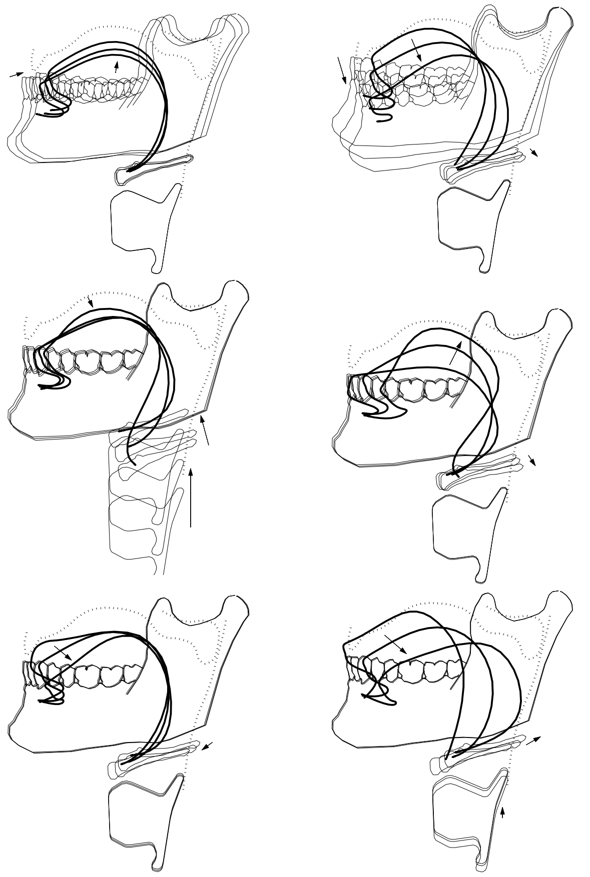

Figure 3: Effect of individual commands on vocal tract configuration. Top, from left to right: Jaw protrusion (JP), Jaw rotation (JR). Middle: Larynx elevation (LY), Tongue dorsum (TD). Bottom: Tongue tip (TT), Tongue body (TB)

dent coordinates which were normalized to have standard deviation equal to one. The residual was projected onto this new coordinate system. The spread of the data is accounted for by a volume which may be obtained from the principal axes of the covariance matrix. The computed volume for the residual (projected on the new coordinate system) was 0.3522 and hence the proportion of variance not accounted for by the six commands was 12.4%, which is the squared value of the volume of the residual.

In conclusion, the six commands account for 87.6% of the variance in hyoid position and orientation. This is consistent with the suggestion that hyoid position and orientation result from the application of other vocal tract commands.

3.3. Inter-articulatory coarticulation

A crucial issue in understanding phenomena such as inter-articulator coordination and coarticulation is how such elementary motions or central commands combine to yield the intricate temporal structure of the observed speech movements. To this regard, a consequence of system dynamics is that movements are delayed with respect to the hypothetic central commands; such a delay may be at least partially responsible for the observed coarticulatory effects [7], without need for explicit planning.

The issue has been addressed with the model in the case of the sequences /igi/ and /ugu/. For each sequence, the configurations corresponding to the vowels (/i/ and /u/) and to /g/ are displayed, as well as the time course of commands TD and TB, and the coordinates of the highest point of tongue surface (*xtd* and *ytd*); see, respectively, Fig. 4 and Fig. 5.

It can be observed that, due to the dynamic effect and to mechanical interaction between tongue and palate, while in both simulations the central commands for /g/ are the same, the place of articulation of /g/ is substantially different in /igi/ and /ugu/; this is consistent with the observations of [5].

4. CONCLUSION

This simple finding emphasizes the need to model muscle properties, reflexes and dynamics before drawing conclusions about the characteristics of vocal tract movement which are explicitly accounted for in planning orofacial motions. These preliminary studies have shown that plant dynamics, i.e. inertial and centrifugal/Coriolis forces are almost completely compensated for by muscle visco-elasticity. This is consistent with the idea that commands may indeed have a relatively 'simple' structure. The present model allows us to assess the plausibility of this hypothesis in detail.

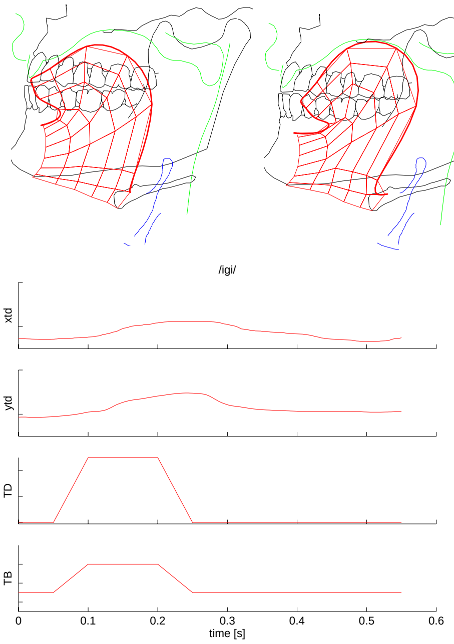

Figure 4: Simulation of sequence /igi/

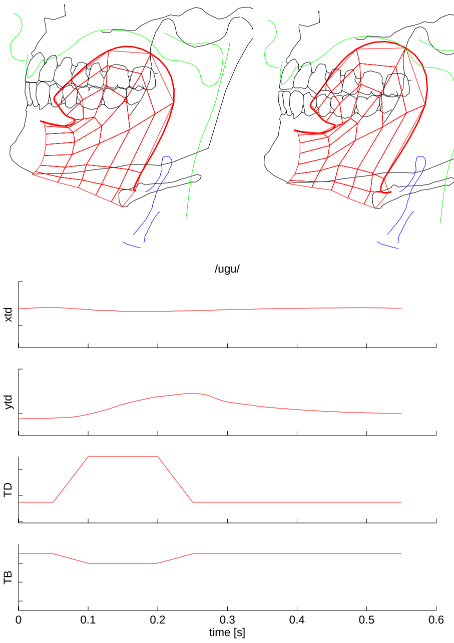

Figure 5: Simulation of sequence /ugu/

5. REFERENCES

[1] A. G. Feldman, S. V. Adamovich, D. J. Ostry, and J. R. Flanagan. The origins of electromyograms – Explanations based on the equilibrium point hypothesis. In J. M. Winters and S. L.-Y. Woo, editors, *Multiple Muscle Systems – Biomechanics and Movement Organization*. Springer-Verlag, New York, 1990.

[2] K. Honda, H. Hirai, and J. Dang. A physiological model of speech production and the implication of tongue-larynx interaction. In *International Congress on Spoken Language Processing*, pages 175–178, Yokohama, 1994.

[3] R. Laboissière, D. J. Ostry, and A. G. Feldman. Control of multi-muscle systems: Human jaw and hyoid movements. *Biological Cybernetics*, 74(3):373–384, 1996.

[4] S. Maeda. Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In W. J. Hardcastle and A. Marchal, editors, *Speech production and speech modelling*, pages 131–149. Kluwer Academic Publishers, The Netherlands, 1990.

[5] S. E. G. Öhman. Numerical model of coarticulation. *Journal of the Acoustical Society of America*, 41:310–320, 1967.

[6] D. J. Ostry and K. G. Munhall. Control of jaw orientation and position in mastication and speech. *Journal of Neurophysiology*, 71(4):1515–1532, 1994.

[7] D.J. Ostry, P.L. Gribble, and V.L. Gracco. Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned? *Journal of Neuroscience*, 16(4):1570–1579, 1996.

[8] Y. Payan, P. Perrier, and R. Laboissière. Simulation of tongue shape variations in the sagittal plane based on a control by the equilibrium point hypothesis. In *Proceedings of the XIII International Conference on Phonetic Sciences*, volume 2, pages 474–477, Stockholm, Sweden, 1995.

[9] J. Perkell. *Physiology of Speech Production*. MIT Press, Cambridge, MA, 1969.

[10] V. Sanguineti, R. Laboissière, and Y. Payan. A control model of human tongue movements in speech. *Biological Cybernetics*, in press.

[11] J. R. Westbury. Mandible and hyoid bone movements during speech. *Journal of Speech and Hearing Research*, 31:405–416, 1988.