In Proceedings of ICPhS '95, Stockholm Sweden (vol. 2, pp.474-477)

SIMULATION OF TONGUE SHAPE VARIATIONS IN THE SAGITTAL PLANE BASED ON A CONTROL BY THE EQUILIBRIUM-POINT HYPOTHESIS.

Payan Yohan, Perrier Pascal & Laboissière Rafael
Institut de la Communication Parlée, URA CNRS 368, INPG & Université Stendhal
46, avenue Félix Viallet, 38031 GRENOBLE Cedex 01, France
E-mail : payan@icp.grenet.fr

ABSTRACT

A 2D biomechanical Finite Element model of the tongue was developed. It integrates four extrinsic muscles and three intrinsic ones, and is controlled by the Equilibrium Point Hypothesis. The ability of this model to generate realistic VV transitions is evaluated and implications for the control of speech are discussed.

INTRODUCTION

Very early on, speech scientists were convinced of the need for models at different levels of speech production process, in order to understand the link between physical aspects of speech and the linguistic task. By modelling the relationship between vocal-tract geometry and the acoustic signal, Stevens and House [1] and lates Fant [2] emphasised the importance of vocal-tract constriction in vowel production; Wood [3] proposed a correspondance between vowel classes and constriction location. With an articulatory model Maeda [4] underlined tongue and jaw coordinations used for the production of the same vowel in different contexts. More recently, Maeda et al [5] proposed a model to account for muscle synergies in the control of position and shape of the tongue.

A speech production model should of course include such synergies. In this perspective, a first approach could involve the learning process [6]: synergies would emerge from the search for the most efficient motor strategies suitable for the production of given articulatory patterns. Another interesting approach consists in moving a step closer to the Central Nervous System (CNS), and studying how the motor control signals to the articulators are organised. From this point of view, Feldman's Equilibrium Point hypothesis [7] is very appealing: in this theory, muscles are not controlled individually, but in respect to global control variables. Muscle coordination is thus implicitly described in the model.

Laboissière et al. [8], using a physiological jaw/hyoid bone model [9] have already emphasised some of the contributions this hypothesis offers for the understanding of the control of speech production.

Our purpose in this paper is to show how this hypothesis sheds light on the control of the tongue, in VV sequences.

The TONGUE MODEL

State of the art.

The tongue is a non-rigid body, capable of bearing large deformations, in order to precisely shape the oral cavities for an accurate production of sounds. By contrast to the jaw (which is an undeformable bone, moved by the action of external muscles), the tongue has embedded muscles, whose shape will be modified under their own activity.

In order to understand the respective contribution of each muscle in the tongue's shaping, Perkell [10] showed the interest of a physiological model of the tongue, which takes into account the intrinsic physiological and anatomical structure which underlines tongue movements. The elastic properties of the tissues were represented by distributed, individually controlled, second order systems. A better analytical description of the continuous elastic structure of the tongue is proposed by the Finite Element Method (FEM). However, up to 1993 Finite Element tongue models were only able to produce static tongue shapes for given EMG inputs [11], [12], [13]. Important progress has been made by Wilhelms-Tricarico [14], who elaborated a 3D biomechanical Finite Element model integrating inertia and force generation properties. This model is able to solve the equations of motion.

(a) (b) (c)

Figure 1: Tongue muscles action : (a) Posterior Genioglossus; (b) Hyoglossus; (c) Styloglossus. For a better representation, the tongue model is embedded in a drawing of the sagittal outlines of the vocal tract.

Such a model is most suitable for evaluating motor control theories. However, attention should be paid to the high degree of computational complexity, which could involve enormous simulation times. In order to overcome this possible drawback and to permit an extensive evaluation of the EP hypothesis for the control of speech movements, we propose to first limit the description of the tongue to the sagittal plane. Even with this limitation, important progress can be made by working on the numerous articulatory data available in this plane (X-ray, microbeams, electromagnetometer data).

A 2D Finite Element model of the tongue

In brief, FEM is a classical technique, in which interpolation functions are developed in order to integrate, across regions of interest, continuous material properties, such as mass, stiffness and deformation capabilities. For this aim, the non-rigid body is described in terms of *discrete nodal* points. The global deformation of the body is analytically calculated from the displacements of these nodes.

The proposed model is described by 63 nodes. The tongue structure is thus defined by 48 isoparametric elements, each of them being delimited by four nodes. This structure satisfies two main criteria : ① the distribution of the nodes, should reflect the internal muscle structure in the sagittal plane, but ② the number of nodes should be compatible with sufficiently low computational costs. The model includes extrinsic muscles —Genioglossus posterior (GGp) and anterior (GGa), Hyoglossus (HG) and Styloglossus (SG)— and intrinsic ones —superior (SL) and inferior (IL) Longitudinalis, Verticalis (V)—. Each muscle is defined by specific sets of elements. Force distribution within each element is determined by the direction of the corresponding muscle fibres. Each element can be shared between many muscles and is able to account for fibres interdigitating.

Figure 1 shows the action of three extrinsic muscles: GGp, HG and SG, respectively mostly recruited for the production of the extreme cardinal vowels [i], [a] and [u].

Finally, the Runge Kutta method is used to solve the equation of motion.

CONTROL OF THE MODEL

Input commands to the physiological models of the tongue are for the most part EMG signals [10], [12], [14]. These models are thus essentially controlled in terms of force level inputs.

Such simulations are very interesting as long as they are only intended to describe the influence of each muscle on the tongue shape. However they seem to be unsuitable for a correct description of tongue control. First of all, as mentioned in the introduction, individual commands to muscles require an additional control layer in order to specify the synergies between muscles (about 20 muscles act together during tongue movement). Moreover, EMG activation, and hence muscular force levels, are the consequences of an interaction between central activation from the CNS to the motoneuron pool, and reflex activation, related in particular to muscle lengths, from the muscle to the motoneuron pool [15]. EMG activation can therefore not be directly controlled by the CNS.

From this point of view, the Feldman EP Hypothesis is very appealing.

Motor innervation to muscles arises from α MNs which innervate the main body of the muscle and from γ MNs which contribute to α MN excitation through reflexes. The basis of the model is the suggestion that movement arises from changes to neural control variables which shift the equilibrium point of the motor system. The essential control variables are independent changes in the membrane potentials of α and γ motoneurons (MNs) which establish a threshold muscle length (λ) at which the recruitment of MNs begins. As the system changes λ, muscle activation, and hence force, vary in relation to the difference between the actual and threshold muscle lengths. Moreover, due to reflex damping, this activation also depends on the rate of muscle length changes. Thus, by shifting λ by changes to the central facilitation of MNs, the system can produce a movement to a new equilibrium position.

The dependence of active muscle force on muscle activation is approximated by an exponential function, estimated from empirical force-length relations for cat gastrocnemius muscle [16].

The equation of motion for the complete system has the following form :

$$M\ddot{U} + f\dot{U} + K(U,\dot{U},\lambda)U = F(U,\dot{U},\lambda) + P$$

where U, $\dot{U}$ et $\ddot{U}$ are displacement, velocity and acceleration nodes.

M is a global mass matrix and f the global passive damping matrix.

F represents the active muscle force and P the external force, corresponding here to gravity.

K is the stiffness matrix, which determines the inner forces of the FE structure; this matrix accounts for the distribution stiffness amongst elements. For active muscles this stiffness depends on node displacements and velocity as given by the slope of the exponential Force/Activation relationship.

SIMULATIONS

Our aim here is to show that tongue movements measured for the transition between two vowels, can be accounted for, by linear shifts of the control variables of the system, between equilibrium positions close to the actual vowel tongue shapes. To accomplish this, we worked on cineradiographic data showing sagittal vocal tract outlines for a French subject [17].

Usually, normal vowel transitions require both jaw and tongue movements. In order to take into account the influence of jaw displacements on the tongue movements, a simple command specifying the mandible position is introduced. For this, the model is embedded in a complete vocal tract model, where jaw position is specified via the inferior incisive. The jaw command is directly obtained from the data by measuring the position of this incisive.

In order to evaluate the Equilibrium Point Hypothesis applied to the model, we worked on [i-a] transition for which tongue movement, as observed in the sagittal plane, is simple and relevant. Moreover, muscle activity during this transition is essentially due to GGp and HG : GGp was supposed to be the only active muscle for the production of the vowel [i], while HG accounted for the production of the vowel [a].

Starting from the rest position, [i] is obtained by a single shift of the posterior GGp lambda command. The amount of force involved to maintain [i] tongue posture is just sufficient to counteract gravity and inner tongue forces: in this condition, GGp lambda is close to the actual length of the muscle. The same strategy is used for HG lambda in the [a] configuration.

Movement from [i] to [a] is generated by linear shifts of the lambda command for the two muscles mentioned above. The command shift rate is the same for both lambdas and is chosen in order to correctly fit tongue shape variations over time.

Figure 2 plots the movement simulated for the whole sequence. The real [i-a] contours measured on the cineradiographic database are superposed to the simulated ones. The adequacy between simulations and data is quite good. Further analysis are in course in order to compare more precisely the kinematics properties of specific points of the tongue .

Figure 2 : simulation of [i-a] transition (dotted) and measured contours (dark).

CONCLUSIONS

By elaborating a 2D physiological model of the tongue, integrating basic elasticity and force generation principles, it was possible to propose preliminary tests of Feldman's theory applied to speech production control. This theory allows, with simple commands, a generation of realistic movement in a VV sequence. Moreover, the intended equilibrium positions are consistent with the shape effectively reached for each of the vowels. This feature is interesting to understand the relationships between vocal tract geometry and motor control space.

ACKNOWLEDGEMENT

The authors acknowledge David Ostry for many discussions. This research was supported by the European Union (ESPRIT-BR Project n° 6975), by the Cooperation France-Québec .

REFERENCE

[1] **Stevens K.N. & House A.S. (1955).** Development of a Quantitative Description of Vowel Articulation. *J. Acoust. Soc. Am..*, 27, 484-493.

[2] **Fant G. (1960).** Acoustic Theory of Speech Production., Mouton, The Hague.

[3] **Wood S. (1979).** A Radiographic Analysis of Constriction Locations for Vowels. *J. of Phonetics*, 7, 24-44.

[4] **Maeda, S. (1990).** 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 (Eds.), *Speech Production and Speech Modeling* (pp. 131-149). Dordrecht, The Netherlands: Kluwer Academic Publishers.

[5] **Maeda, S., Honda K., and Kusawaka N. (1993).** From EMG to Vowel Formant Patterns. *3rd Seminar on Speech Production: Models and Data*. Old Saybrook, CT., 11-13 May.

[6] **Jordan, M.I. (1990).** Motor Learning and the Degrees of Freedom Problem. In M. Jeannerod (Ed.), *Attention and Performance* (Chapter XIII). Hillsdale, NJ: Erlbaum.

[7] **Feldman, A.G. (1966).** Functional Tuning of The Nervous System with Control of Movement or Maintenance of a Steady Posture – II Controllable Parameters of the Muscles. *Biophysics*, *11*, 565-578.

[8] **Laboissière R., Ostry D.J, and Perrrier P. (1995).** *In these Proceedings.*

[9] **Laboissière R., Ostry D.J, and Feldman A.G.** (submitted). The Control of a Multi-Muscle Systems: Human Jaw and Hyoid Movements.*J. of Neurophysiology.*

[10] **Perkell, J.S. (1974).** *A Physiologically-Oriented Model of Tongue Activity in Speech Production*. Ph.D. Thesis. Boston, Ma.: Massachussetts Institute of Technology.

[11] **Kiritani S., Miyawaki, K., and Fujimura, O. (1976).** A Computational Model of the Tongue. *Annual Bulletin*, *10* (pp. 243-252). Tokyo: Research Institute of Logopedics and Phoniatrics, University of Tokyo.

[12] **Kakita, Y., Fujimura, O., and Honda, K. (1985).** Computation of Mapping from Muscular Contraction Patterns to Formant Patterns in Vowel Space. In V.A. Fromkin (Ed.), *Phonetic Linguistics* (pp. 133-144). Orlanda, Florida: Academic Press.

[13] **Hashimoto K. and Suga S. (1986)** Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue. *Journal of Acoustic Society of Japon* (E) 7,1,39-46.

[14] **Wilhelms-TricarioR**.(in press) Physiological Modeling of Speech Production: Methods for Modeling Soft-Tissues Articulators. *J. Acoust. Soc. Am..*

[15] **Feldman A. (1986).** Once more on the Equilibrium-Point Hypothesis for motor control", *Journal of Motor Behavior,Vol. 18,n°1.*

[16] **Feldman A. and Orlovsky G.N. (1972).** The influence of different descending systems on the tonic reflex in the cat. *Experimental Neurology* 37:481-494.

[17] **Badin P., Gabioud B., Beautemps D., Lallouache T.M., Bailly G., Maeda S., Zerling J.P.& Brock G.(1995).** Cineradiography of VCV sequences : articulatory-acoustic data for a speech production model. *15th International Congress on Acoustics,* Trondheim, Norway.