

A MODEL OF HUMAN JAW AND HYOID MOTION AND ITS IMPLICATIONS FOR SPEECH PRODUCTION

R. Laboissière¹, D.J. Ostry*, P. Perrier[†]

[†]Institut de la Communication Parlée
URA CNRS 368 / INPG / Univ. Stendhal

46, av. Félix Viallet 38031 Grenoble CEDEX 1 France

*Dept. of Psychology, McGill University
1205 Dr. Penfield Avenue, Montreal QC Canada H3A 1B1

E-mail: rafael@icp.grenet.fr

ABSTRACT

We present a model of sagittal plane jaw and hyoid motion based on the λ version of the equilibrium point hypothesis. This hypothesis suggests that movements arise from shifts in the equilibrium position of the speech articulator. The equilibrium is described as a consequence of the interaction of central neural commands, reflex mechanisms, muscle properties and external loads, but it is under the control of central neural commands. These commands act to shift the equilibrium via centrally specified signals acting at the level of the motoneuron (MN) pool. In the context of the model, we focus on a number of issues in speech research. We consider the implications of the model for the notion of articulatory targets. We suggest that simple linear control signals may underlie smooth articulatory trajectories. We explore as well the phenomenon of intra-articulator coarticulation in jaw movement.

INTRODUCTION

A major difficulty in inferring control strategies in speech from the kinematic data characterizing human orofacial motion is the lack of physiologically based models of the underlying control. Without such models, it is

difficult to assess the adequacy of descriptive accounts in which the underlying control is contaminated by factors such as dynamics and muscle mechanical properties. Physiological models permit the separation of aspects of the kinematics due to neural control from those due to the biomechanical properties of the system. In this paper, we present one such physiological model, the equilibrium point hypothesis of motor control, and consider its specific application to issues in speech control.

The model is introduced in detail below. However, very briefly, the hypothesis suggests that movements arise from shifts in the equilibrium position of the speech articulator. The equilibrium is a consequence of the interaction of central neural commands, reflex mechanisms, muscle properties and external loads, but it is under the control of central neural commands. These commands act to shift the equilibrium via centrally specified signals acting at the level of the motoneuron (MN) pool.

In the sections which follow, we introduce the basic concepts of the EP hypothesis. We describe its application to the development of a seven muscle, four degree of freedom model of jaw and hyoid motion (Laboissière, Ostry, and Feldman submitted). With the aid of simulations we show (1) how the

concept of equilibrium position provides insights into the concept of articulatory targets and (2) how at least some of the observed variability associated with intra-articulator coarticulation can be attributed to dynamics rather than central control.

THE JAW MODEL

Biomechanical Structure One of the main goals in modelling orofacial function is to study the form of CNS commands which underlie the kinematic observables. Consequently, in order to understand control on the basis of kinematics, it is necessary to be able to separate control signals from the system's biomechanics. Modelling the elaborate geometry and dynamics of the orofacial articulators is a necessity, not because of a specific interest in their characteristics, but because it is otherwise exceedingly difficult to relate control signals to the resulting kinematics which may be measured empirically.

With this aim, we have recently developed a model of the jaw and hyoid system (see Fig. 1) (Laboissière, Ostry, and Feldman submitted). The model, which is implemented as a computer simulation, has seven muscles (or muscle groups) and four kinematic degrees of freedom. Movements are not controlled directly in terms of commands to individual muscles. Rather, consistent with empirical evidence (Bothorel 1975; Ostry and Munhall 1994), control signals, which are based on different combinations of commands to muscles, are organized at the level of the system kinematic degrees of freedom. This enables independent production of jaw rotation, horizontal jaw translation, vertical hyoid translation and horizontal hyoid translation. The level of co-contraction is also controlled. These control signals may act alone or in combination.

Control of Jaw Motion: The Equilibrium Point Hypothesis Motor innervation to muscles arises from α MNs which innervate

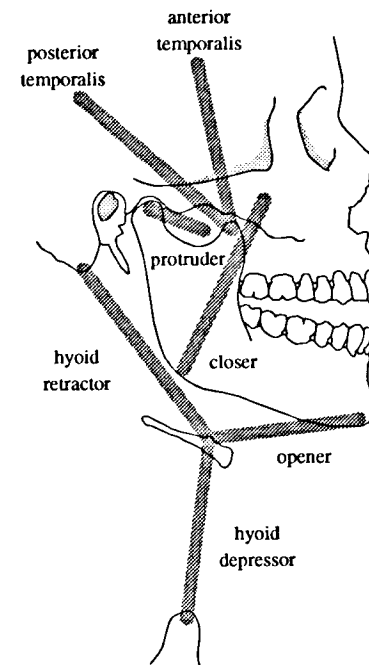


Figure 1: Schematic representation of modelled muscles with their attachments to the jaw and hyoid bone.

the main body of the muscle and from γ MNs which contribute to a 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 state 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. Muscle activation and hence force vary in relation to the difference between the actual and the threshold muscle lengths and the rate of muscle length change. Thus, by shifting λ through changes to the central facilitation of MNs, the system can produce movement to a new equilibrium position.

The central notions associated with the ba-

asic control mechanism are shown in Fig. 2 in the context of a single jaw muscle and load. For simplicity, we will focus on the jaw closer muscle masseter (shown in black) and on the load due to the gravitational force. The panel on the left shows a number of different jaw configurations. The corresponding depolarization of α MNs is shown at the top right. The horizontal line gives the threshold for MN recruitment. The descending input to the MN provides the level of central facilitation which may be specified independent of muscle length. Afferent facilitation also contributes to the depolarization of the MN and varies directly with muscle length. Thus, while the equilibrium position is essentially under central control, the activation level of the MN reflects a net contribution which includes both the direct descending input to the MN and indirect input due to afferent pathways.

In (a), we see, in the left hand panel, the subject resting horizontally with the system in equilibrium at a position near to occlusion. The level of total depolarization, as seen in the top right hand panel, exceeds the threshold level and the load due to gravity is supported by an overall level of central and afferent activity. When the subject changes to a vertical position (b), the load acting to extend the masseter increases. This increases the level of muscle-length dependent afferent facilitation of the MN pool which in turn acts to establish a new equilibrium position. Note, that the level of central facilitation is unchanged by these changes in load even though the total level of MN depolarization is changed.

The lower right hand panel of Fig. 2 demonstrates these characteristics in terms of the muscle's force-length curve. The variable λ gives the muscle length at which motoneuron recruitment begins. The exponential shape of the force-length relation reflects the well known size-principle for MN recruitment such that as the difference between the actual and threshold muscle length progressively increases, progressively larger

motor units with larger force outputs are recruited. At muscle length l , a force equal to F is generated which balances the load (a). A change in the position of the head relative to the gravitational force loads the jaw and stretches the muscle to length l' . The length dependent afferent facilitation results in the recruitment of new motor units which increases force to F' (b). At this point, the muscle force balances the load force. To summarize, changes in load which result in muscle stretch (or unloading) lead to the recruitment (or derecruitment) of motor units as a result of changes in length dependent facilitation to the MN pool. The measure of independent central control, λ , is unaffected even though both force and muscle length are changed.

As shown by comparing (b) and (c), the model suggests that voluntary movement arises as a consequence of increases in the level of central facilitation to the MN pool. Increases in facilitation depolarize MNs and result in the recruitment of additional motor units. This increases total force and results in muscle shortening. As the muscle shortens, the facilitation to the MN pool due to length dependent afferent input decreases and a new equilibrium is established. Voluntary movements are depicted in the lower right hand panel in terms of the muscle's force-length relation. At (b), where threshold muscle length is λ , the weight of the jaw is supported by muscle force F' at muscle length l' . Increasing central facilitation serves to reduce the threshold length for MN recruitment from λ to λ' . As λ shifts, the difference between the actual and threshold muscle length increases, more MNs are recruited and the muscle begins to actively shorten. As the threshold length reaches λ' , the jaw achieves a new equilibrium state in which muscle force is F' and muscle length is l . The movements which arise from changes to the independently specified parameter λ thus depend on both direct central facilitation to the MN pool and facilitation arising from afferent input to the MN.

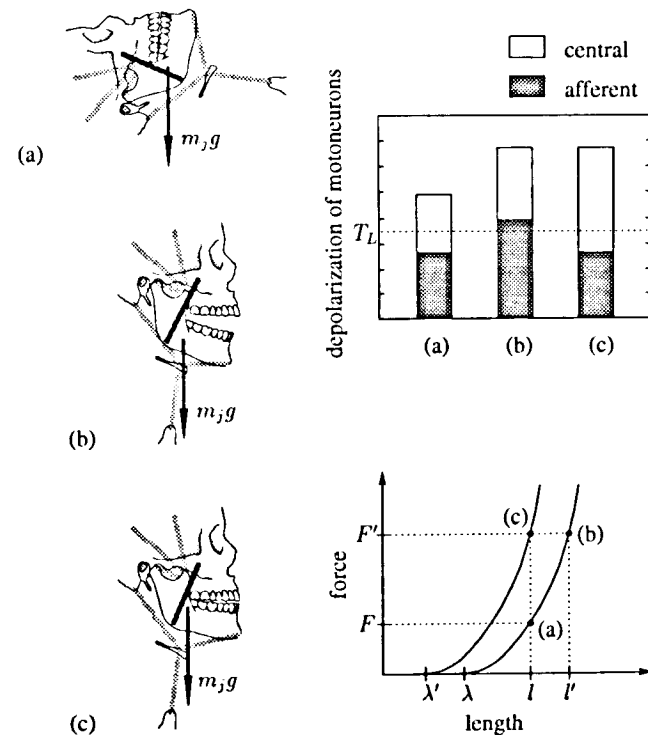


Figure 2: Jaw positions (left hand panel), levels of depolarization of MNs (upper right), and force-length curves (lower right) are shown (see text for details).

Control signals which are associated with changes in values of the system's four degrees of freedom are each mapped onto control signals at the level of individual muscles (λ). This corresponds to a time varying facilitation to the MNs of each muscle. As described above, the muscle activation and force depend on the difference between the threshold muscle length and the current length as well as the rate of muscle length change. Fig. 3 shows the layout of the model and its control mechanisms. Muscle length and velocity information are provided by muscle spindle afferent input to MNs following a reflex delay. The elements in each

of the muscle blocks in the figure correspond to individually modelled muscle mechanical properties. These include muscle properties such as the dependence of force on muscle length and on passive elastic properties and the graded development of muscle force due to calcium dependent muscle kinetics. Mechanical damping is provided by velocity dependent reflex inputs and muscle intrinsic properties. The force arising in each muscle contributes to the production of jaw and hyoid forces and torques. These act through the system's equations of motion to produce changes in jaw and hyoid position and orientation. There are separate jaw and hyoid dy-

namics and realistic musculo-skeletal geometry.

The model thus provides the means to study human jaw motions in speech in a manner which explicitly integrates biomechanical characteristics and the underlying control. In the following section we will explain the significance of a number of properties of the λ model and then, using simulations with the jaw model, show their implications for the control of speech

THE FORM OF CENTRAL COMMANDS, INVARIANCE AND VARIABILITY IN SPEECH

The Form of Central Commands One of the fundamental problems in motor control is obtaining an understanding of the form of the neural signals which underlie movements. In speech, the form of the control signal is of particular importance as it contributes to our understanding of the relationship between the phonological level and the corresponding organization at the level of the vocal tract. By exploring speech at the level of control, we can assess extent to which the regularities observed kinematically and hence acoustically correspond to invariances postulated at the linguistic level. The EP hypothesis, by its very nature, allows us to address in what ways the equilibrium as specified at the level of motor system might correspond to aspects of spatial targets, which serve as landmarks in the control of the speech sequence.

According to the model, movements are effectively changes in posture, that is, shifts in the equilibrium state of the system. We suggest that the control of speech may be related to specific postures of articulators and hence that posture and successive changes in posture correspond to a representation of the articulatory task at the level of control. This idea that articulatory movements are intended towards spatial positions is related to MacNeilage's (1970) proposals of spatial articulatory targets.

Fig. 4 lets us explore the concept of speech targets in the context of the jaw model. The

figure shows empirical and simulated jaw motions during repetitions of /isisa/. The empirical data are shown with solid lines, the predicted jaw kinematics with dots and the presumed underlying equilibrium shifts with alternating dots and dashes. The jaw orientation angle is shown in the upper panel and horizontal jaw position is shown below.

In fitting the data, we have assumed that the jaw equilibrium angle and equilibrium horizontal position both shift at a constant rate. Changes in the rate and duration of the equilibrium shift are the two controlled variables. Examination of the data shows that the correspondence between empirical and model data is generally good. Note that constant rate equilibrium shifts in the model produce the smooth movements which are observed kinematically. This suggests that smoothness of movements may not be explicitly planned or controlled but rather may arise from the dynamics.

The equilibrium shifts, particularly in the case of horizontal jaw translation, are often observed to extend beyond the kinematic endpoints of the movement. The overshoot of the actual trajectory by the equilibrium arises in the model from the need to produce the sufficiently large accelerations which are required to move the jaw in a continuous fashion at speech rates. The need to have the equilibrium position overshoot the actual spatial goal to produce rapid movement has also been demonstrated in simulations of multi-joint arm movements (Hogan 1984).

The idea that articulatory movements are intended towards spatial positions has been proposed by MacNeilage (1970). Our simulations suggest that the literal interpretation of spatial targets as actually achieved positions within the vocal tract may be incorrect in the case of continuous speech. Nevertheless, we think it is reasonable to assume that regularities relating speech as a linguistic task to speech at the motor level may be found in terms of the control signals underlying movement. This will, of necessity, entail a comparison of empirical and model data.

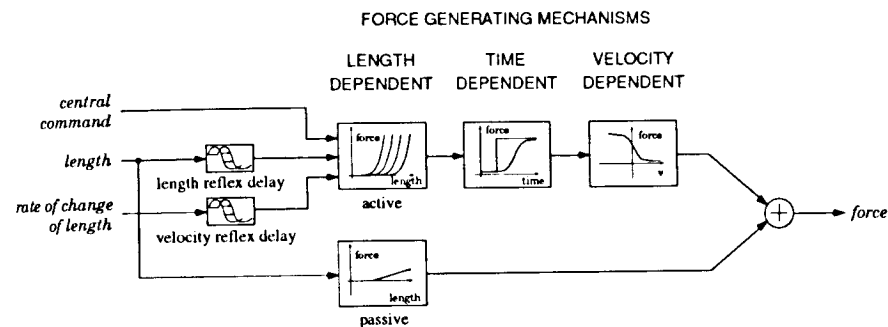


Figure 3: Schematic representation of the muscle model

Regularities relating the units of linguistic description to the control signals of speech motions might be sought in terms of correspondences related to both equilibrium position and rates of equilibrium shift.

Speech Variability Variation in articulatory movement is clearly one of the most pervasive characteristics of speech. Some of the aspects of speech movement variability are almost certainly planned and relate to specific changes in control signals. Others may not be planned but may arise from factors such as muscle mechanics, musculo-skeletal geometry and the dynamics of the physical system. Evidence that inter-articulator variations in speech are planned is supported by the findings of Abry and Lallouache (in press, also see Perkell and Matthies 1992). These authors analyzed anticipatory lip protrusion in [iCy] sequences, in which C represents consonant clusters of 0 to 5 consonants, none of which involved lip protrusion. They showed that the onset time of the protrusion movement increased linearly with the size of the consonant cluster. The fact that lip protrusion to produce the same final vowel begins earlier in some contexts than in others supports the idea that anticipatory patterns are the result of a process which takes account of upcoming phonetic context when planning successive speech movements.

The kinematic patterns of intra-articulator

coarticulation which are readily measurable in empirical studies may also appear to be centrally controlled on the basis of kinematic changes which arise in response to upcoming phonetic segments (e.g., Lindblom 1983). However, without explicit models of speech articulators, measured kinematic effects correctly attributable to central planning cannot be distinguished from the kinematic patterns which are due to dynamics and are not represented in the underlying control. To address this possibility, we will show how kinematic variability may arise even when the underlying control signals related to the specification of articulatory position remain fixed. The main conclusion we will wish to draw is that unplanned effects due to physical sources must be accounted for before drawing conclusions about central control or inferring planning mechanisms.

Using the jaw model, we have studied the predicted kinematic patterns in simulated V_1CV_2 transitions. In these simulations, the equilibrium shifts associated with the V_1C movement remain constant while the equilibrium shifts associated with the CV_2 movement amplitude are systematically varied. Thus, at the level of central control, no account was taken of upcoming phonetic context in the specification of the CV_2 transition. However, when one examines the predicted kinematic patterns (Fig. 5), we see that the V_1C amplitude and duration are systematically affected by the identity of the

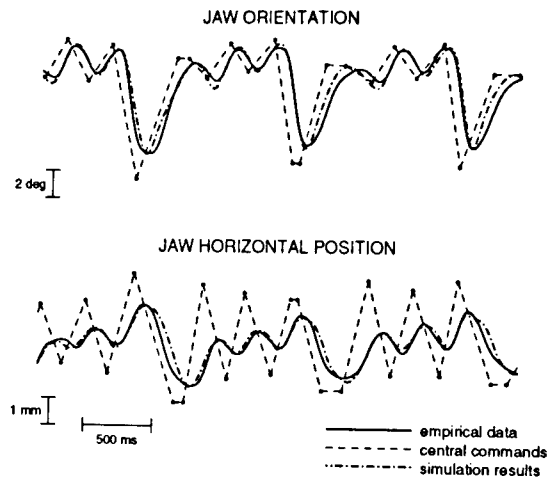


Figure 4: Empirical and model data during repetitions of /isis/. The hyoid bone is at the assumed rest position for occlusion through the simulated movement.

final vowel. As movement amplitude for the final vowel decreases, the simulated amplitude and duration of the initial transition increase. Comparable patterns of intra-articulator coarticulation have been reported in empirical studies of jaw, tongue dorsum, velar, and lower pharyngeal wall coarticulation (Ostry and Gracco 1995; Parush, Ostry, and Munhall 1983; Parush and Ostry 1986; Parush and Ostry 1993). Thus, while on the basis of kinematic evidence alone, it could be concluded that intra-articulator coarticulation is consistent with the notion of planned coarticulation, our present simulations suggest that this possibility should be evaluated with care: unplanned effects due to articulator dynamics must be accounted for before drawing conclusions about the role of central control in intra-articulator coarticulation.

CONCLUSION

In this paper, we have described a model of jaw and hyoid motion based on the EP hypothesis. We have described simulations which examine the form of the central con-

trol signal. We have shown that smoothness in movement may arise from dynamics and need not be planned. We have suggested that regularities relating speech as a linguistic task to speech at the motor level may be found in the control signals underlying movement. We have examined sources of articulatory variability. We have shown that kinematic patterns comparable to those reported in intra-articulator coarticulation may arise as a result of dynamics rather than central planning.

ACKNOWLEDGEMENTS

This research was supported by NIH grant DC-00594 from the National Institute on Deafness and Other Communication Disorders, Advanced Telecommunications Research (ATR), Kyoto, Japan, European Community ESPRIT - BR grant # 6975 Speech Maps, Natural Sciences and Engineering Research Council of Canada, and France-Quebec project # 070192. We also acknowledge the Rhône-Alpes Region for support.

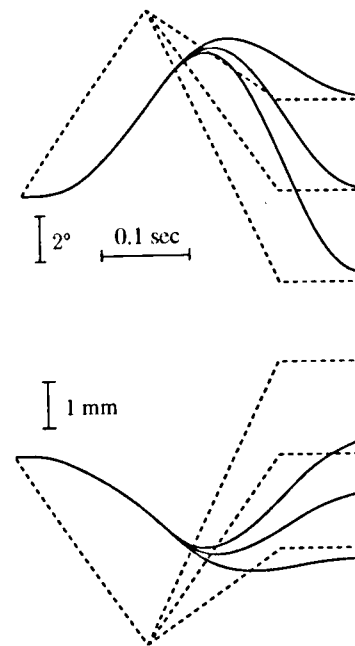


Figure 5: Predicted kinematic patterns of jaw rotation (above) and horizontal translation (below) and the presumed underlying control signals during a V_1CV_2 utterance. Note that the whereas the magnitude and the duration of the equilibrium shifts associated with V_1C transition are fixed, the predicted V_1C duration and movement amplitude vary with V_2 .

REFERENCES

- Abry, C. and T. Lallouache (in press). Propositions pour un modèle d'anticipation dépendant du locuteur. données sur l'arrondissement du français. *Bulletin de la Communication Parlée*. Grenoble, France: Institut de la Communication Parlée, University of Grenoble.
- Bothorel, A. (1975). Positions et mouvements de l'os hyoïde dans la chaîne parlée. *Travaux de l'Institut de Phonétique*

de Strasbourg 7, 80-132.

Hogan, N. (1984). An organizing principle for a class of voluntary movements. *J. Neurosci.* 4(11), 2745-2754.

Laboissière, R., D. Ostry, and A. Feldman (submitted). The control of human jaw and hyoid movement. *J. of Neurophysiol.*

Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage (Ed.), *The Production of Speech*, pp. 217-245. New-York, USA: Springer Verlag.

MacNeilage, P. (1970). Motor control of serial ordering of speech. *Psychol. Rev.* 77, 182-196.

Ostry, D. and V. Gracco (1995, June). Is intra-articulator speech coarticulation planned? *Paper presented at the 129th Meeting of the Acoust. Soc. Am.* Washington, DC.

Ostry, D. J. and K. G. Munhall (1994). Control of jaw orientation and position in mastication and speech. *Journal of Neurophysiology* 71, 1528-1545.

Parush, A. and D. Ostry (1986). Superior lateral pharyngeal wall movements in speech. *J. Acoust. Soc. Am.* 80, 749-756.

Parush, A. and D. Ostry (1993). Lower pharyngeal wall movement in speech. *J. Acoust. Soc. Am.* 94, 715-722.

Parush, A., D. J. Ostry, and K. G. Munhall (1983). A kinematic study of lingual coarticulation in VCV sequences. *J. Acoust. Soc. Am.* 74(4), 1115-1125.

Perkell, J. and M. Matthies (1992). Temporal measures of anticipatory labial coarticulation for the vowel [u]: Within- and cross-subject variability. *J. Acoust. Soc. Am.* 91, 2911-2925.