

A dynamic biomechanical model for neural control of speech production

Vittorio Sanguineti^{a)}

*Dipartimento di Informatica, Sistemistica e Telematica, Università di Genova, Via Opera Pia 13,
16145 Genova, Italy*

Rafael Laboissière

Institut de la Communication Parlée, 46, av. Felix Viallet, F-38031 Grenoble, France

David J. Ostry

McGill University, 1205, Dr. Penfield Avenue, Montréal, Quebec H3A 1B1, Canada

(Received 23 April 1997; accepted for publication 2 October 1997)

A model of the midsagittal plane motion of the tongue, jaw, hyoid bone, and larynx is presented, based on the λ version of equilibrium point hypothesis. The model includes muscle properties and realistic geometrical arrangement of muscles, modeled neural inputs and reflexes, and dynamics of soft tissue and bony structures. The focus is on the organization of control signals underlying vocal tract motions and on the dynamic behavior of articulators. A number of muscle synergies or “basic motions” of the system are identified. In particular, it is shown that systematic sources of variation in an x-ray data base of midsagittal vocal tract motions can be accounted for, at the muscle level, with six independent commands, each corresponding to a direction of articulator motion. There are two commands for the jaw (corresponding to sagittal plane jaw rotation and jaw protrusion), one command controlling larynx height, and three commands for the tongue (corresponding to forward and backward motion of the tongue body, arching and flattening of the tongue dorsum, and motion of the tongue tip). It is suggested that all movements of the system can be approximated as linear combinations of such basic motions. In other words, individual movements and sequences of movements can be accounted for by a simple additive control model. The dynamics of individual commands are also assessed. It is shown that the dynamic effects are not neglectable in speechlike movements because of the different dynamic behaviors of soft and bony structures. © 1998 Acoustical Society of America. [S0001-4966(98)04801-2]

PACS numbers: 43.70.Aj, 43.70.Bk [AL]

INTRODUCTION

In this paper, we report on a midsagittal plane model of the motion of the tongue, jaw, hyoid bone, and larynx. We describe both the development of the biomechanical model and, in the context of the model, we consider the way control signals to muscles are organized to produce multi-articulator motion. We consider in addition the effects of articulator dynamics on the motions of the tongue and jaw. We will report analyses which suggest that the control of speech movements can be accounted for by a small set of independent commands. We will also suggest that to understand the control of orofacial motions, realistic physical and biomechanical models as well as modeled control signals are needed. Initial versions of the jaw and hyoid model (Laboissière *et al.*, 1996) and the tongue model have been reported previously (Sanguineti *et al.*, 1997).

Most modeling work to date has focused on individual orofacial structures, and models of the face and lips (Müller *et al.*, 1984; Terzopoulos and Waters, 1990; Lee *et al.*, 1995), tongue (Perkell, 1974; Kiritani *et al.*, 1976; Hashimoto and Suga, 1986; Wilhelms-Tricarico, 1995; Sanguineti

et al., 1997; Payan and Perrier, 1997), jaw (Baragar and Osborn, 1984; Throckmorton and Throckmorton, 1985a, 1985b; Otten, 1987; van Eijden *et al.*, 1988; Laboissière *et al.*, 1996), hyoid bone, and larynx have been proposed. The models include graphical animations, models of muscle mechanical properties, and static force estimation. More complete models, which include muscle properties, dynamics, and simulated neural mechanisms have been reported (Otten, 1987; Laboissière *et al.*, 1996; Sanguineti *et al.*, 1997; Payan and Perrier, 1997).

Although the majority of modeling studies deal with individual articulators, many problems in orofacial research can only be addressed in terms of the combined action of multiple articulators. These problems include the basis of coordination in speech and mastication, the complexity of interarticulator coupling at the level of the control signals, and the determinants of interarticulator coarticulation (Ostry *et al.*, 1996). The presence of vocal tract mechanical interactions underscores the need for multi-articulator models. It is essential to account for the interactions between soft tissue and bony structures in order to have accurate prediction of vocal tract motion (Honda *et al.*, 1994; Sanguineti *et al.*, 1997).

The model presented here is based on the λ version of the equilibrium point hypothesis of motor control. The model

^{a)}Mailing address: Department of Physiology, Northwestern University Medical School, 303 East Chicago Ave., Chicago, IL 60611. Electronic mail: sangui@parker.physio.nwu.edu

includes muscle properties and realistic muscle geometry, modeled neural inputs and reflexes, and articulator dynamics. As in our jaw and hyoid simulations (Laboissière *et al.*, 1996) and our recent model of the tongue (Sanguineti *et al.*, 1997), we have assumed that control signals to individual muscles are coordinated to enable the nervous system to produce a number of independent motions.

In the present paper, we consider a number of problems relating to the identification of the control signals underlying orofacial movements and how they are coordinated to produce multi-articulator motion. We first perform a model-based factor analysis of the Strasbourg x-ray data base (Bothorel *et al.*, 1986) in order to identify the basic motions of the system and to infer their associated commands. This relates control in the model to empirical data, addresses the extent to which control signals are organized in a low-dimensional control space, and provides a basis for comparison with purely geometric articulatory models.

We examine the extent to which the effects of commands which we derive are additive. Additivity removes the need for context specificity in central commands. Purely geometric articulatory models such as that proposed by Maeda (1990) assume that the effect of the individual articulators on vocal tract shape is additive. However, this may not necessarily be the case given the complex mechanics of the orofacial system. To the extent that additivity in control signals can be demonstrated in the present model, it suggests that the predictions related to additivity in geometric models may still hold.

We focus as well on the predicted dynamics of motions associated with the individual commands. This is motivated by the observation that speech involves relatively synchronous articulator motions and therefore complex command patterns may be necessary if the dynamics of individual articulators differ.

I. THE MODEL

Jaw motions in the model have two degrees of freedom—orientation in the sagittal plane and translation along the articular surface of the temporal bone; also see Laboissière *et al.* (1996); the hyoid has three degrees of freedom, horizontal and vertical position and sagittal plane orientation. The larynx is modeled as a point mass with a single degree of freedom—vertical position, which has the largest kinematic effect. Midsagittal plane tongue movements are modeled, as described below, using finite element techniques (Schwarz, 1984).

Previous studies have suggested that interactions between the individual vocal tract structures—between hard and soft tissues (Sanguineti *et al.*, 1997) and between the larynx and the tongue or the hyoid bone (Honda *et al.*, 1994)—are significant in determining the global mechanical behavior of the system. For these reasons, care has been taken in modeling the interaction between individual structures, by deriving the global equations of motion for this system (see the Appendix). This ensures that mechanical interactions including reaction forces and velocity-dependent forces are accounted for.

The model geometry (see Fig. 1) is that of a young fe-

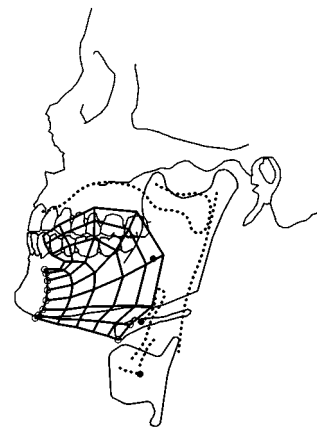


FIG. 1. Biomechanical model of the mid-sagittal section of the oral cavity. Circles indicate the nodes of the tongue mesh which are fixed with respect to either the jaw or the hyoid bone. Dots indicate the centers of mass of the jaw, the hyoid bone, and the larynx. Dotted lines indicate the approximate boundaries of the oral cavity. Thick lines correspond to the tongue mesh.

male speaker, for whom an x-ray data set of midsagittal plane vocal tract images is available (Bothorel *et al.*, 1986). The modeled jaw position and orientation at occlusion were estimated by superimposing a normative model of the jaw (Scheideman *et al.*, 1980) on the x-ray data. The tongue surface contour, the hyoid position and orientation, and the larynx height were likewise obtained from the x-ray data set.

A. The tongue

The tongue is the main determinant of vocal tract shape. It has been modeled as a viscoelastic continuum whose behavior has been assumed, as a first approximation, to be linear and isotropic. The coefficient of elasticity, or Young's Modulus, E , is that of passive muscle tissue (Duck, 1990)— $E=6.2$ kPa. A Poisson's ratio of $\nu=0.49$ is used (Hashimoto and Suga, 1986). This approximates conservation of volume at a microscopic level. (By microscopic level, we mean here the limit behavior for an infinitely small discretization of the continuous tongue tissue. As our FE (finite element) discretization is quite coarse, volume conservation is not completely guaranteed in our model.) We assumed that there is no deformation in the transverse direction, and that the X and Y components of deformation only depend on position on the XY plane (this corresponds to the hypothesis of plane strain).

Tongue mass has been assumed to be $m_t=0.1$ kg. Its density has been taken to be that of muscle tissue, which is $d_t=1040$ kg/m³ (Duck, 1990). This is slightly greater than the density of water.

The interaction of the tongue and the palate has also been accounted for. Contact forces are assumed to be elastic (depending of the level of "penetration" of each node into the palate), and directed normally. It is thus assumed that there is zero friction.

By applying standard finite element (FE) techniques (Schwarz, 1984), the tongue configuration has been approximated by a discrete mesh (Fig. 1), whose configuration is completely specified by the vector \mathbf{x} that includes the X and Y coordinates of each of the nodes in the mesh. We used a

6×8 mesh to describe the tongue, thus yielding a 96-dimensional configuration vector \mathbf{x} . However, some of the nodes are fixed with respect to either the jaw or the hyoid bone (see the Appendix and Fig. 1 for details).

B. Jaw, hyoid bone, and larynx

The geometrical arrangement of jaw and hyoid bone is based on Laboissière *et al.* (1996). The jaw is represented as a rigid body that can rotate about the temporomandibular joint and translate along the articular surface of the temporal bone. The shape of this surface has been described by a third-order polynomial, $y = a_1x^3 + a_2x^2$ (see Laboissière *et al.*, 1996). Thus, if x_0 is the X coordinate of the center of rotation of the jaw, the corresponding Y coordinate is $y_0 = y(x_0)$. Accordingly, jaw configuration \mathbf{q}_j is a vector with two components: x_0 , and the orientation angle α_j , relative to the occlusal plane. As in Laboissière *et al.* (1996), jaw mass and inertia have been estimated to be $m_j = 1$ kg and $I_j = 0.0042$ kg m².

The hyoid bone has been modeled as a free rigid body, characterized by its position and orientation. In particular, hyoid configuration is described by the vector $\mathbf{q}_h = [\mathbf{x}_{Gh}^T \alpha_h]^T$, where \mathbf{x}_{Gh} represents the X and Y coordinates of the center of mass, and α_h is the orientation. Hyoid mass has been assumed to be $m_h = 0.1$ kg; its corresponding moment of inertia (relative to the center of mass) has been calculated to be $I_h = 2.8 \times 10^{-5}$ kg m². This value has been estimated by approximating the hyoid bone as a U-shaped object of midsagittal length of 3 cm and radius 1.5 cm, with uniformly distributed mass.

The larynx is a complex musculo-cartilagenous structure whose main function is to control vocal fold configuration. Only the thyroid cartilage is attached to the bony structures of our model, namely the hyoid bone and the sternum. As we assume that the muscles originating on these bony structures insert on the thyroid at a single point, the larynx is modeled as a point mass with $m_l = 0.1$ kg. We assume also that it can only translate vertically, which is a good approximation for our x-ray data. In summary, the height of the larynx is assumed to correspond to the observed height of the vocal folds as determined from the x-ray tracings (see Sec. II).

Other degrees of freedom of the laryngeal system, related to the motion of the vocal folds and to the relative motions of the cricoid and the thyroid cartilage, have not been modelled. Although horizontal thyroid motion is important acoustically (Honda *et al.*, 1994), its amplitude is small. It has been omitted since the primary focus here is on biomechanics rather than acoustics.

C. Muscle properties and neural control

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 (Feldman *et al.*, 1990). By changing the values of λ 's over time, the musculoskeletal system may be caused to move to a new equilibrium position.

This mechanism is modeled by assuming that muscle activation (A) develops in proportion to the difference between λ and a reflex component, depending on actual muscle length (l) and its rate of change:

$$A(t) = [l(t-d) - \lambda(t) + \mu \dot{l}(t-d)]^+, \quad (1)$$

where $[x]^+ = \max[x, 0]$ and d is reflex delay. The parameter μ characterizes the dependence of the muscle's threshold length on velocity and provides damping due to proprioceptive feedback. Damping due to muscle intrinsic properties is also included (see below). For simplicity, we have assumed that μ is the same for all muscles and constant (0.07 s). The value for μ was set on the basis of simulation studies carried out with a multi-joint arm model (Gribble *et al.*, 1998)—the value of μ was adjusted so that simulated joint viscosity in statics matched empirically obtained estimates for this variable (Tsuji *et al.*, 1995, see Gribble *et al.*, 1998 for details). We have used a reflex delay, d , of 15 ms for all muscles. The value was based on observed delays in human jaw openers and closer muscles (Lamarre and Lund, 1975; Ostry *et al.*, 1997b).

It should be noted that the model assumes that afferent input associated with muscle length and velocity is combined with descending input to α MNs to yield muscle activation. Position- and velocity-dependent afferent input in limb muscles arises from muscle spindle receptors. However, several orofacial muscles including the jaw opener, anterior digastric, and the jaw protruder, lateral pterygoid, have few if any muscle spindles. We have nevertheless recently demonstrated both phasic and tonic stretch reflexes in human jaw opener muscles (Ostry *et al.*, 1997). This suggests that these reflexes are not necessarily mediated exclusively by muscle spindle afferents. In tongue muscles, stretch responses have also been reported (see Sanguineti *et al.*, 1997 for review).

Increases in muscle activation due to changes in λ are associated with MN recruitment and increases in firing rate and muscle force. Active force, M , has been modeled as an exponential function of the form

$$M = \rho [\exp(cA) - 1], \quad (2)$$

which has been suggested by the experimental studies of Feldman and Orlovsky (1972) and accounts for both the intrinsic and reflex components of active force. The parameter ρ is assumed to vary with muscle force generating ability, and may be estimated from each muscle's maximum force capability. In particular, a value of ρ equal to 25% of maximum muscle force has been found (see Gribble *et al.*, 1997 for details) to be consistent with the static stiffness observed in the human arm. Here c is a form parameter, related to the MN recruitment gradient, and is assumed to be equal for all muscles (see Laboissière *et al.*, 1996). The exponential relationship between force and muscle length is consistent with the size principle (Henneman *et al.*, 1965), that is, as the difference between the actual and threshold muscle length increases, progressively larger motor units are recruited and larger increments in force are obtained.

We also included in the model (see Fig. 2 for a schematic diagram) the dependence of muscle force on muscle lengthening or shortening velocity (Joyce and Rack, 1969),

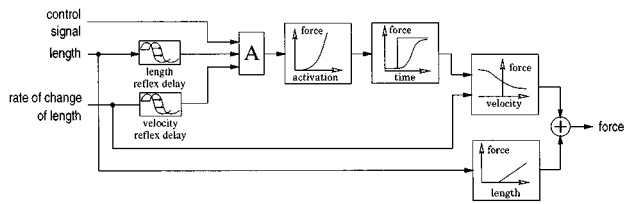


FIG. 2. Block diagram of the muscle model (see text for details).

the graded development of force over time (Huxley, 1957), and the passive elastic stiffness of muscle (Feldman and Orlovsky, 1972); see Laboissière *et al.* (1996) for details. The force–velocity relation was modeled with a sigmoidal function (Laboissière *et al.*, 1996) which was obtained by fitting data for cat soleus muscle. Separate parameter estimates were obtained for tongue muscles and for other orofacial muscles. The parameters were selected to match empirically reported force-velocity functions for fast (tibialis) and slow (soleus) muscle, respectively (Wells, 1965). The gradual development of muscle force was modeled using a second-order low pass filtering of active muscle force, M . The filter was critically damped and had a time constant of 15 ms which led to an asymptotic response to a step input in about 90 ms (Miller, 1991). Passive muscle stiffness was assumed to vary with physiological cross-section area.

D. The muscle system

The muscles of the oral cavity have a complex geometrical arrangement. The individual fibers within a muscle may have very different lines of action and their paths may be curved. Thus, the directions of muscle action may not be simply approximated as a straight line.

In the present model, we have assumed that some muscles are made of a discrete number of “macro-fibers” that are formed by division of a distributed muscle into a number of spatially segregated compartments. The number of compartments used for each muscle depends upon the shape of the muscle and in particular upon its directions of action. The geometrical arrangement of each macro-fiber is approximated by a series of line segments that connect the nodes of the tongue mesh or connect the tongue mesh to specific points on the bony structures. Each macro-fiber is treated as a single entity. Its length and velocity are defined as the sum of the lengths and velocities of the individual segments.

The geometrical arrangement of modeled muscles is based on anatomical descriptions (Miyawaki, 1974; Dickson and Maue-Dickson, 1982; McDevitt, 1989) and on previous modeling work (Laboissière *et al.*, 1996; Sanguineti *et al.*, 1997). In the tongue model, we have included three extrinsic muscles, genioglossus (GG), hyoglossus (HG), and styloglossus (SG), and three intrinsic muscles, superior longitudinalis (SL), inferior longitudinalis (IL), and verticalis (VE). Muscles acting on the jaw include a jaw opener (OP), which models the effects of geniohyoid and the anterior belly of digastric, a jaw closer (CL), which represents the effects of the masseter and medial pterygoid, anterior and posterior temporalis (AT and PT), and superior and inferior lateral

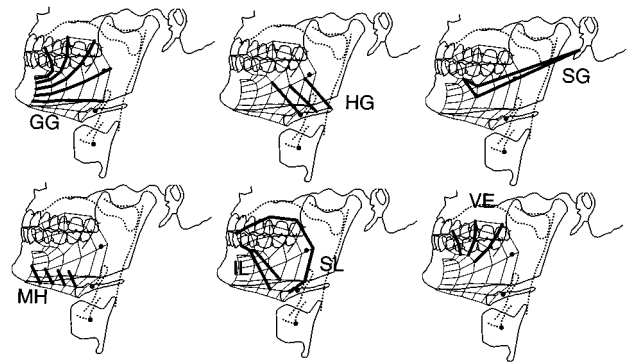


FIG. 3. The geometric arrangement of tongue muscles. Top, from left to right: genioglossus (GG, 5 macro-fibers); hyoglossus (HG, 3 macro-fibers); styloglossus (SG, 2 macro-fibers). Bottom: mylohyoid (MH, 4 macro-fibers), superior (SL, 6 macro-fibers) and inferior (IL, 2 macro-fibers) longitudinalis, verticalis (VE, 3 macro-fibers). Thick lines represent the macro-fibers that were used to model each muscle.

pterygoid (SP and IP). The mylohyoid (MH) originates on the jaw and inserts onto the tendinous median raphe and on the hyoid bone. In the present model, we focused on its role in forming the tongue floor, and accordingly its attachments to the hyoid bone are not included. However, its effect on the hyoid bone is taken into account by the finite element modeling of the tongue floor. In other words, despite the fact that no connection is explicitly modeled between MH and the hyoid, the model accounts correctly for the effect of MH contraction, i.e., raising and protrusion of the hyoid bone.

Additional muscles that act on the hyoid bone and larynx include a hyoid retractor (RE), which models the effects of the posterior belly of digastric and the stylohyoid, the thyrohyoid (TH), the sternohyoid (SH), and the sternothyroid (ST). It should be noted that by modeling the larynx as a point mass, the attachments of muscles to the larynx had to be restricted to this point. This results in some inaccuracy in the lines of action of TH and ST. The musculo-skeletal geometry of the model is depicted in Figs. 3 and 4.

Individual λ 's are associated with each muscle and each macro-fiber in the model. The latter point requires comment. Relatively little is known about the neural organization of

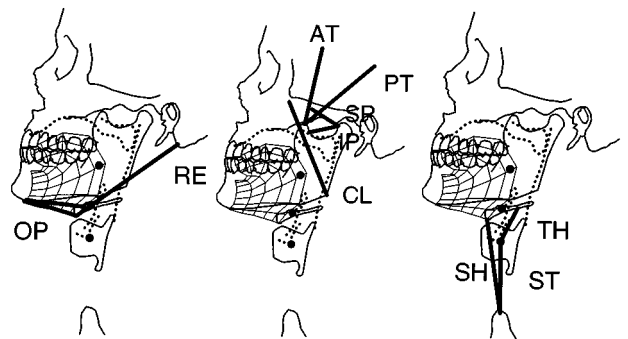


FIG. 4. The geometric arrangement of jaw, hyoid, and laryngeal muscles. Left: opener (OP, 2 macro-fibers) and retractor (RE). Middle: closer (CL), anterior (AT), and posterior (PT) temporalis, superior (SP) and inferior (IP) pterygoid. Right: thyrohyoid (TH), sternohyoid (SH), and sternothyroid (ST). Thick lines represent the macro-fibers that were used to model each muscle.

TABLE I. The estimated muscle parameters.

Muscle	CSA (mm ²)	f_m^{\max} (N)	ρ (N)	K_p (N/m)
Genioglossus (GG)	309	67.8	13.6	0
Hyoglossus (HG)	296	65.1	13.0	0
Styloglossus (SG)	110	24.2	4.84	0
Mylohyoid (MH)	186	40.9	8.18	32.2
Superior longitudinalis (SL)	65	14.3	2.86	0
Inferior longitudinalis (IL)	88	19.4	3.88	0
Verticalis (VE)	66	14.5	2.90	0
Jaw opener (OP)		115	23.0	34.7
Jaw closer (CL)		639	128	192
Retractor (RE)		86.3	17.6	23.1
Superior pterygoid (SP)		126	25.2	38.0
Inferior pterygoid (IP)		252	50.4	76.0
Anterior temporalis (AT)		362	72.6	109
Posterior temporalis (PT)		197	39.4	59.4
Thyrohyoid (TH)		28.7	5.74	8.65
Sternohyoid (SH)		28.7	5.74	8.65
Sternothyroid (ST)		28.7	5.74	8.65

control signals in distributed muscle structures such as the tongue. There is some evidence in the context of empirical electromyographic data of functional partitioning of the genioglossus muscle (Baer *et al.*, 1988) and some suggestion that the superior longitudinalis may not behave as a single muscle (Dickson and Maue-Dickson, 1982). However, our decision, in the present context to associate a separate λ with each macro-fiber, arises as a compromise. Since the dimensionality of control to individual muscles is essentially unknown, by providing individual λ 's to each macro-fiber we are able to analyze the dimensionality of control on the basis of the patterns of covariation of λ 's which arise in fitting the tongue model to the x-ray data base (see Sec. II for details).

The maximum forces for tongue muscles, f_m^{\max} , have been determined from estimates of their cross-sectional areas on the basis of anatomic atlases (see Sanguineti *et al.*, 1997 for details), and by assuming a maximum specific tension of 22 N/mm², reported in Wilhelms-Tricarico (1995) for the geniohyoid muscle. In the case of jaw and hyoid muscles, the values of maximum force and passive stiffness, K_p , reported by Laboissière *et al.* (1996) were used (see Table I).

E. Organization of control signals

A number of additional assumptions may be made concerning the organization of control signals to individual muscles. 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 springlike behavior, the set of λ 's associated with individual muscles (or muscle compartments) specify an equilibrium configuration for the system. This does not mean that the individual λ 's are independently controlled. Indeed, control is presumably organized into a relatively small number of different combinations of λ changes, which we will refer to as "commands." Commands in effect define muscle synergies that correspond to elementary or primitive motor behaviors. 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? A first possible criterion is that of independent motions. In the case of jaw motion, the observation of a variety of different patterns of coordination between jaw protrusion and rotation has suggested (Ostry and Munhall, 1994) that in speech its mechanical degrees of freedom can be controlled independently. Moreover, the data of Westbury (1988) suggest that the observed patterns of motion of the hyoid bone are largely uncorrelated with jaw movements. On the other hand, simulation studies (Honda *et al.*, 1994) have demonstrated a close mechanical coupling between the hyoid bone, the larynx and the tongue. These findings suggest that, although the hyoid-larynx system and the jaw are not mechanically independent, these structures are controlled by different muscle synergies.

A second criterion is that of independent muscle groups. Öhman (1967) and Perkell (1969) 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 (Maeda, 1990; Sanguineti *et al.*, 1997).

Our approach to this identification problem is essentially data driven, as will be described in Sec. II. Central commands (i.e., synergies of muscle λ 's) are inferred from an empirical data set using a numerical optimization technique and a factor analysis. The obtained results reflect the variability of vocal tract configurations present in the corpus.

II. RESULTS

In this section, we focus on the organization of control and its relation to the mechanical properties of the system and the anatomical arrangement of muscles. In Sec. II A, we identify the basic motions of the system and their associated commands. In Sec. II B, we examine the related issue of whether the individual compartments of the spatially distributed muscles of the tongue are independently controlled. In

Sec. II C, we explore the extent to which summation applies to the system's basic motions. Finally, in Sec. II D we present simulations of the dynamic behavior of the system in response to simple rhythmic commands that act on individual articulators. These results provide initial predictions concerning the dynamic behavior of the tongue/jaw system.

A. Determination of independent commands

Based on empirical evidence which shows that the behavior of orofacial structures may each be characterized by a small number of independent motions (see Sec. II E), we have used the Strasbourg x-ray data set (Bothorel *et al.*, 1986) in conjunction with our model to identify these basic motions, and to infer the mapping between their associated commands and the control signals to individual muscles.

Consistent with experimental evidence and with related simulation studies (Laboissière *et al.*, 1996; Sanguineti *et al.*, 1997), each of these motions can be represented as a different combination of changes to muscle λ 's. One combination of λ 's results in an increase of the global stiffness of the system, without accompanying motion.

We will show that systematic sources of variation in the x-ray data can be accounted for with six independent commands, each corresponding to a direction of articulator motion, or more specifically, to a linear combination of control signals to individual muscles (λ 's). There are two commands for the jaw corresponding to sagittal plane jaw rotation and jaw protrusion, one command controlling larynx height, and three commands for the tongue corresponding to forward and backward motion of the tongue body, arching and flattening of the tongue dorsum, and motion of the tongue tip.

The Strasbourg data base consists of 519 frames of mid-sagittal plane x-ray images of a single female speaker (subject number 3, PB) during the continuous production of ten short sentences, pronounced in a normal-to-fast rate. The sentences were chosen to be representative of the phonetic variation of French. The sampling frequency for x-ray images is 50 Hz. The midsagittal tongue contours and those of the bony structures were estimated by hand tracing from the x-ray lateral views.

For each x-ray image, we ran a constrained optimization procedure, in order to determine the set of individual muscle λ 's and the corresponding model configuration. The constraints were the requirements that the system be in mechanical equilibrium, and that the nodes on the upper side of the tongue mesh lay on the empirically observed tongue contour. The observed positions and orientations of jaw, hyoid, and larynx were also extracted from the x-ray image, thus determining the positions of the associated model articulators. The criterion to be minimized was the level of cocontraction, which was defined as the squared distance between actual muscle lengths and λ 's:

$$C(\mathbf{Q}, \boldsymbol{\lambda}) = [\mathbf{I}(\mathbf{Q}) - \boldsymbol{\lambda}]^T \cdot [\mathbf{I}(\mathbf{Q}) - \boldsymbol{\lambda}], \quad (3)$$

where \mathbf{Q} is the system configuration (see the Appendix). The quantities $\mathbf{I}(\mathbf{Q})$ and $\boldsymbol{\lambda}$ are, respectively, the vectors of lengths and λ 's for each of the muscles and macro-fibers in the model.

The procedure resulted in a set of λ 's that can be interpreted as the representation, in the space of muscle control signals, of the variety of configurations that the system can assume during speech movements.

Commands associated with the motion of individual articulators were derived from the above set of muscle λ 's by means of a two-step factor analysis [see Maeda (1990) for a similar approach]. First, the contributions of jaw rotation, jaw protrusion, and larynx elevation to muscle λ 's were derived by linear regression. This step was motivated by empirical observations which suggest that the nervous system controls jaw and larynx motion in terms of their mechanical degrees of freedom (Ostry and Munhall, 1994; Laboissière *et al.*, 1996). In total, 15.8% of variance in the set of muscle λ 's (derived from the x-ray data base) was attributable to the motion of the jaw and larynx.

The contributions of tongue motion to muscle λ 's were derived by carrying out a principal components analysis in the subspace of muscle λ 's that were not correlated with jaw and larynx motions. In fact, principal components in λ space define a number of muscle groups that act independently and have orthogonal directions of action (Sanguineti *et al.*, 1997). This is a property of the geometric arrangement of tongue muscles that is implied in the conjecture (Öhman, 1967; Perkell, 1969) that tongue motions are determined by a small number of independently controlled components, or articulators.

Each of the factors or regression coefficients described above is a vector specifying a direction of change in λ space. The application of the vector corresponds to a shift of the equilibrium configuration of the system. Movements of different amplitude can be obtained by varying the magnitude of the vector. This vector thus defines a "command" for the system.

The factor analysis led to the identification of three commands for tongue motion: tongue dorsum arching/flattening, tongue tip raising/lowering, and tongue body front/back. The tongue movement commands accounted for 40.5% of the total variance in the set of muscle λ 's derived from the x-ray data. Note that while the tongue, jaw, and laryngeal commands taken together account for only 57% of the total variance in λ space, the residual factors have almost no observable effects on the posture or configuration of the system.

In addition to the commands which result in tongue, jaw, and larynx motion, a command controlling the global level of muscle cocontraction could be defined. The cocontraction command was determined by finding, in the space of the residual factors (the factors not already included in the set of jaw, tongue, and larynx commands), a direction of λ change which resulted in an increase in force in each of the modeled muscles. By changing the magnitude of the cocontraction command, global stiffness may be increased without movement.

Figure 5 shows the effects of the above commands. Each panel shows three tracings corresponding to the effects of a single command on the configuration of the system. All panels show a neutral configuration (corresponding to the statistical mean of all λ 's) plus two additional tracings representing ± 4 times the standard deviation of the factor associated

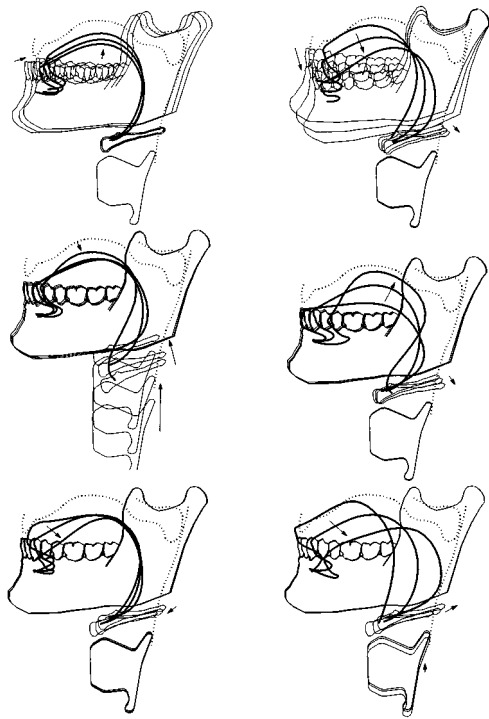


FIG. 5. Effect of individual commands on vocal tract configuration. Top, from left to right: jaw protrusion, jaw rotation. Middle: larynx height, tongue dorsum. Bottom: tongue tip, tongue body. Arrows indicate the motion of each structure; arrow lengths reflect actual movement magnitude. The penetration of the palate in the lower right-hand panel is a consequence of modeling contacts with elastic forces.

with that command. These outer tracings are the extreme positions for each command.

The top two panels of Fig. 5 show the results of the jaw commands. The protrusion command produces, in addition to jaw protrusion and retraction, small but noticeable changes in tongue elevation (in particular during retraction) but little movement of the hyoid bone. The command for jaw rotation affects both jaw orientation and tongue position. Both lowering and raising appear to have an active effect on tongue postures. In both cases greater movements of the tongue blade are observed than would be expected if the tongue simply moved passively with the jaw. Such a synergistic action of jaw and tongue—when the jaw opens, the tongue is actively lowered—can be observed in this particular x-ray data set, and is therefore reflected in our factor analysis.

The middle panel of Fig. 5 (left-hand side) shows the effect of the larynx height command. Changes in larynx height are observed to affect hyoid vertical position but have little effect on hyoid orientation or upon the positions of the jaw and tongue.

The three tongue commands affect the tongue and hyoid, and, in one case, the elevation of the larynx. The tongue dorsum command produces arching and flattening of the tongue, the tongue tip commands produce raising and lowering of the tip with almost no effect on the posterior profile of the tongue, and the tongue body command advances and retracts the tongue.

It should be noted that the set of the commands derived

in this fashion are primarily dependent upon the geometrical arrangement of muscles and not upon the cost function used to do the optimization. This was shown by repeating the procedure using a different cost function. One cost function, shown in Eq. (3), determines the set of λ 's which minimize the average squared distances between actual muscle lengths and λ 's. A second cost function determines λ 's which minimize the average squared muscle force (normalized for muscle cross-sectional area). While both cost functions can be interpreted as measures of cocontraction, they are not linearly related and therefore they should yield different values of the optimal λ 's.

Differences between the resulting commands and their effects were assessed quantitatively by computing the angles between each of the six tongue, jaw, and larynx commands, obtained by using the two different cost functions. The angles ranged from 14 to 42 deg, with an average 29.5 deg (in the space of λ changes). Moreover, the commands derived with the second cost function resulted in vocal tract configurations that were comparable to those shown in Fig. 5. This was assessed quantitatively in terms of the positions of three selected nodes on the tongue surface, namely (from anterior to posterior) tongue tip (TT), tongue blade (TB), and tongue dorsum (TD) (see also Sec. II C). The directions of motion of these points as a result of the application of each of the commands (ranging from -4 to 4 as in Fig. 5) were compared for the two different cost functions. The average angle between the directions of motion was found to be 18 deg.

In summary, the muscle λ 's for tongue, jaw, hyoid, and larynx muscles were derived by fitting the model to the Strasbourg x-ray data base. Commands corresponding to basic motions of the tongue, jaw, and larynx were obtained using factor analysis. The commands are associated with maximally independent sources of λ variation and may be interpreted as corresponding to the muscle synergies which underlie motions of this system.

B. Functional independence of muscle compartments

No direct empirical evidence exists on which to identify the organization of control signals to spatially distributed muscles. In the present section, we attempt to infer this organization in the context of the tongue on the basis of pattern of variation of tongue muscle λ 's.

The spatially distributed nature of tongue muscles was represented by a number of macro-fibers which were treated as if they were independently controlled. This leads to an increase in the number of degrees of freedom of the tongue. However, systematic patterns of correlation were found among the set of λ 's for the macro-fibers associated with individual muscles. To assess the dependence among the λ 's of the macro-fibers of each muscle, we carried out, for each muscle separately, a principal components analysis on the set of λ 's associated with all macro-fibers for that muscle (for the entire data set). Figure 6 gives the cumulated proportion of variance accounted for by the principal components of each muscle. The figure shows that 75% or more of the variance in λ 's in extrinsic tongue muscles can be accounted for by two factors for genioglossus, one each for styloglossus

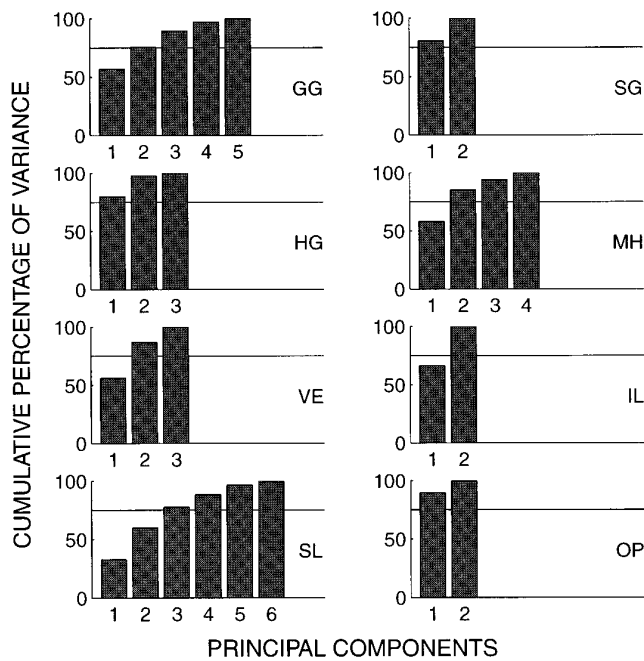


FIG. 6. Cumulative percentage of variance accounted for by the principal components associated with the macro-fibers of each tongue muscle (75% of total variance is shown with solid lines). See text for muscle labels.

and hyoglossus and two for mylohyoid. For the intrinsic tongue muscles, two factors are required for verticalis and inferior longitudinalis and three for superior longitudinalis. One factor is sufficient to account for control signals to the jaw opener muscles (anterior digastric and geniohyoid). Thus, the analysis suggests that control signals to the spatially distributed fibers which comprise each of the muscles of the tongue may themselves be grouped into a small number of independent commands.

C. Additivity of commands

Articulatory models sometimes assume that the effects of commands are additive (for example, Maeda, 1990). However, it is unclear whether this assumption holds in systems that have complex geometry and mechanics.

In the case of the present model, 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.

System behavior was characterized in terms of the positions of three selected nodes on the tongue surface, namely tongue tip (TT), tongue blade (TB), and tongue dorsum (TD), and of the tip of the mandibular incisor (MN). Changes of their positions were assessed as a result of the application of each of the commands. The procedure was repeated for a wide range of initial vocal tract configurations.

Figure 7 shows, for each command, the displacement of nodes as arrows connecting the initial to final positions. In some cases, the arrows overlap and thus the number of lines may appear to differ.

The top two panels show the jaw protrusion and jaw rotation commands, the middle panels are for larynx height and for the tongue dorsum command, and the bottom panels give the tongue tip and tongue body commands. The critical aspect of each figure is the behavior of the node most closely associated with a particular command. Hence, with the exception of the larynx command which results in little movement of the tongue and jaw, each of the other commands produces movements of its associated node that change little in direction as a result of changes in the configuration of the tongue and jaw.

We have assessed the extent to which the individual commands (except for the larynx command) produce movements of comparable direction in their associated nodes (MN for jaw commands, TD for tongue dorsum, TT for tongue tip, TB for tongue body) when initiated from different vocal tract configurations. For each of the tongue and jaw commands, the standard deviation of the direction of node movements, shown in Fig. 7, was computed about their respective population means. The resulting standard deviations of command directions with changes in vocal tract configuration were protrusion command, 3 deg, jaw rotation command 0.8 deg, tongue dorsum command, 6.5 deg, tongue tip command, 1.9 deg, and tongue body command 3.8 deg.

These findings are consistent with the idea that the effects of different commands are additive, in terms of positioning of points on the tongue surface inside the oral cavity. Since a given command has essentially the same effect in terms of postural change for any workspace configuration, 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.

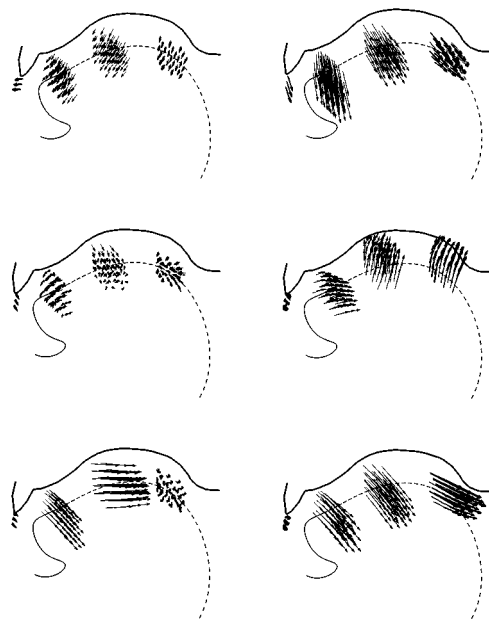


FIG. 7. Effect of each individual command in different initial configurations. Top, from left to right: jaw protrusion, jaw rotation. Middle: larynx height, tongue dorsum. Bottom: tongue tip, tongue body. In each panel, the displacement of selected nodes (from left to right, MN, TT, TB, and TD) is represented by arrows connecting their initial and final positions.

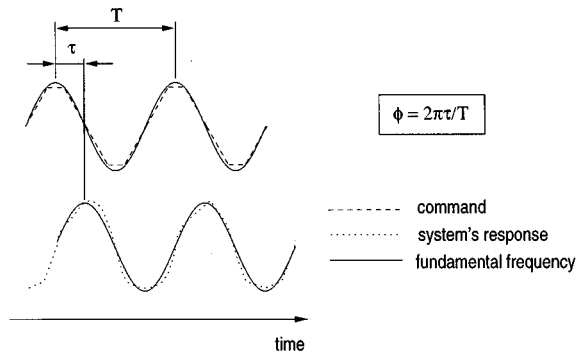


FIG. 8. The phase lag of the fundamental Fourier component of the response with respect to the control signal.

D. Dynamics of articulators

Aspects of the dynamic behavior of the system have been assessed by examining simulated movements which result from simple periodic commands. Simulations were carried out for each of the jaw, larynx, and tongue commands individually. The tests used cyclic (stepwise linear) control signals at different frequencies (namely, 1, 3, and 5 Hz) and, to test system nonlinearity, at different amplitudes (two and three times the standard deviation; see Fig. 5). The simulations were repeated at different levels of cocontraction: 0%, i.e., no cocontraction, and 25% (a level of 100% corresponding to the situation in which at least one of the muscles has reached its maximum force capability). The above commands and levels of cocontraction are assumed to be representative of the working conditions of the system during speech.

System behavior was characterized in terms of the time-varying trajectories of MN, TT, TB, and TD. For both the X and Y coordinates of these nodes, we estimated amplitude and phase lag of the fundamental Fourier component (corresponding to the frequency of the command); see Fig. 8. In the case of jaw rotation and protrusion, tongue motion can be decomposed into a “passive” component due to jaw movement, and an “active” tongue deformation. In this situation, the estimation procedure was carried out for tongue movements relative to the jaw.

This procedure allows us not only to assess the general dynamic behavior of the system, but also to identify differences in the dynamics of individual structures (e.g., jaw and tongue, i.e., bony and soft structures), and also of different portions of the tongue. It may be predicted, for instance, that the tongue can move faster than the jaw due to its smaller mass, and also to the larger proportions of “fast” fiber types that are found in tongue muscles. Figure 9 summarizes the phenomena observed in the simulations, averaged across command amplitudes and levels of cocontraction. As expected, in jaw rotation movements the vertical motion of MN displays a phase lag that is much larger than that observed in tongue nodes; see Fig. 9 (top).

The vertical motion of tongue nodes results from the combined effect of a “slow” component, due to jaw rotation, and a “fast” component, due to active tongue lowering. The relative contributions of these “slow” and “fast” com-

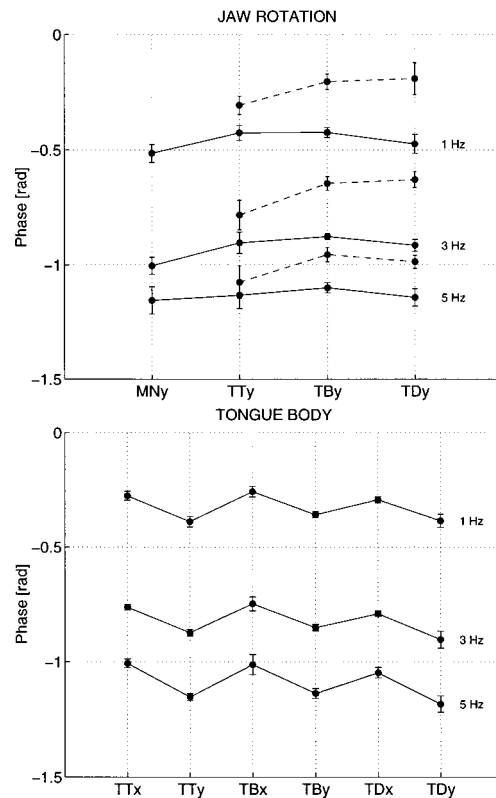


FIG. 9. Averaged phase lags observed in motions of jaw and tongue nodes, during jaw rotation (top) and tongue body (bottom) movements. Continuous lines represent absolute motions, dashed lines (top panel) indicate motion of tongue nodes relative to the jaw. Error bars reflect variability across command amplitudes and levels of cocontraction. The three families of lines correspond to the three different command frequencies; from top to bottom, 1, 3, and 5 Hz.

ponents to the overall tongue motion can be assessed by estimating the motion of each tongue node with respect to the jaw, and its corresponding phase lag. The phase lag of this “fast” component of tongue motion is comparable to those observed in pure tongue movements.

The simulations also show that, regardless of the particular command, horizontal tongue movements tend to be faster than the vertical ones. This effect can be observed in Fig. 9 (bottom), which shows movements resulting from the tongue body command. The pattern observed here may be due to mechanics of muscular hydrostats, for which movements along the main dimension (the long axis) are larger and faster than the transverse ones (Chiel *et al.*, 1992). Changes in the level of cocontraction result in small but observable modifications of the phase lags of the observed movements. Figure 10 shows this effect in the typical case of jaw rotation. In particular, it can be observed that increases in the level of cocontraction result in a smaller phase lag, thus suggesting a decrease in the apparent “damping” of the system. This is consistent with the notion that cocontraction controls the stiffness of the entire system.

III. DISCUSSION

We have presented a physiological model of the motions of the tongue, jaw, larynx, and hyoid bone, based on the λ version of the equilibrium point hypothesis.

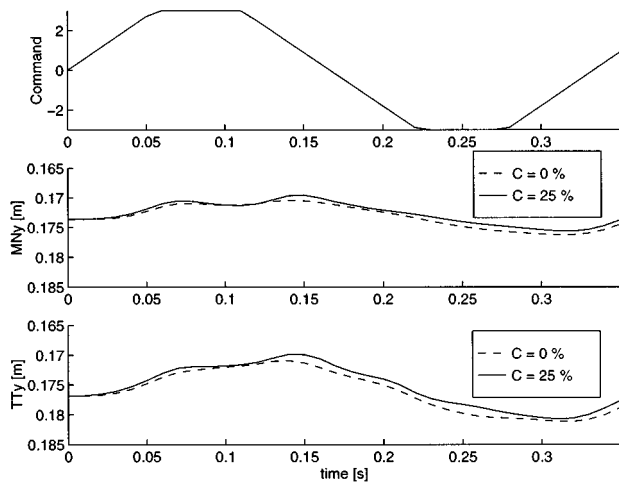


FIG. 10. Effect of cocontraction on vertical motion of mandibular incisor and tongue tip, in the case of jaw rotation commands of amplitude 3 S.D. and frequency of 3 Hz.

Using empirical x-ray data in conjunction with the model, we have shown that the motions of these articulators can be accounted for by a small number of independent commands in λ space. For the jaw and larynx, we identified commands that are related to the mechanical degrees of freedom of these articulators. For the tongue, the commands correspond to different combinations of control signals to individual muscles which produce maximally independent motions. We have shown that these commands have essentially similar effects regardless of the vocal tract configuration and therefore any movement can be expressed as the composition of such independent commands or basic motions.

The idea that control is organized in terms of coordinated commands or muscle synergies derives in part from observations that, in the jaw, independent motions may be produced in the jaw's mechanical degrees of freedom (Ostry and Munhall, 1994; Ostry *et al.*, 1997a). Data on tongue motions are also consistent with the idea that muscles act synergistically to produce the basic or elementary vocal tract motions. Ohman (1967), for example, suggested that there was independent motion of different parts of the tongue depending on phonetic context. Harshman *et al.* (1977) and Maeda (1990) have shown that tongue shapes and tongue motions can be partitioned into statistically independent components.

Accordingly, in the context of the model, we explored a hypothesis of organization of control signals that is based on the following assumptions: (i) for the jaw, rotation and protrusion/retraction movements are separately controlled; (ii) vertical motions of the larynx-hyoid complex can be carried out independently of jaw movements; (iii) the tongue can move independently of the bony parts; and (iv) the basic motions of the tongue reflect the geometric arrangement of tongue muscles.

In particular, consistent with related approaches (Maeda, 1990), we have found that tongue movements can be accounted for by three independent commands. While the effects of the commands are similar to those reported by

Maeda and Honda (1994), the effects arise in different ways. Whereas in previous work functional subdivisions in tongue motion have been associated with disjoint subsets of tongue muscles (Maeda and Honda, 1994; Perkell, 1969), in the present model, the control signals to all tongue muscles contribute to the production of each of the basic motions (also see Smith, 1992).

Moreover, consistent with previous findings, our correlation analyses suggest that the tongue muscles which contribute to these synergies may not act as unitary structures. Thus, we have been able to observe a functional subdivision, based on groupings of λ change, of genioglossus and mylohyoid, each of which contributes to tongue movements in two different ways. It should be noted, however, that the subdivisions observed in the pattern of control signals to tongue muscles does not appear to have direct parallels to anatomical subdivisions (see, for example, Baer *et al.*, 1988).

A particularly interesting and indeed somewhat surprising finding was that individual tongue and jaw command changes result in similar changes in tongue and jaw position regardless of the initial configuration of the system. The result is surprising since with changes to vocal tract configurations, the distribution of forces associated with commands defined in terms of λ shifts might have varied considerably and led to significant variation in the resulting movements. The observed invariance in the effects of these commands presumably occurs as a result of compensations which arise at the geometrical level. A consequence of invariance is that it permits an exceedingly simple organization of commands (namely, an additive model) in which computations are unneeded to account for changes in workspace geometry (Laboissière *et al.*, 1996; Ostry *et al.*, 1996).

Few models of orofacial motion have explicitly included dynamics—see Laboissière *et al.* (1996) for the jaw-hyoid system, and Wilhelms-Tricarico (1995) and Payan and Perrier (1997) for the tongue. Those which have assume that the individual structures are mechanically independent. In the present paper, we show that consideration of the dynamic interaction of vocal tract bony and soft structures is needed to correctly account for orofacial dynamics. Failure to account for these structures results in a model which does not account for forces acting on the tongue and as a result leads to incorrect predictions of tongue movements and achieved tongue positions in behaviors such as speech. Thus, the jaw is not simply a moving frame of reference for the tongue but jaw motion imparts force to the tongue which changes its shape. In the case of the larynx, changes in tongue shape have been shown to be capable of producing changes in pitch due to the mechanical coupling of the tongue and the larynx (Honda *et al.*, 1994).

The “simplicity” of the central commands that the nervous system must provide is a major problem in understanding how speech production is planned. In the case of human arm movements, it has been suggested (Gribble *et al.*, 1997) that muscle viscoelastic properties and the peripheral neural circuitry have a built-in capability to “compensate” for dynamic effects such as inertia and Coriolis/centrifugal forces, a feature that is captured by the λ model. Such a compensatory ability of muscles and reflexes suggests that the nervous

system may not need precise information about the dynamics of the body and of the external world in order to plan movements. Hence, “central” control commands may be “simple” in the sense that they only need to account for the kinematic aspects of the desired movement.

In the neural control of speech production, this feature would be particularly desirable because of the mechanical complexity of the system and of the performance requirements of speech movements (speed, precision, need for an accurate synchronization of motions of different articulators). However, the differences in jaw and tongue dynamics, which emerge from our simulations, appear to challenge the hypothesis of command “simplicity,” in particular with regards to their coordination.

In fact, orofacial movements in speech must satisfy precise timing constraints. For instance, what makes voiced and voiceless stops perceptually distinguishable is the difference in the timing between glottal opening and release of occlusion. If the individual structures differ in their dynamic behavior, such a synchronization can only be achieved at the planning level, by means of some form of prediction of system dynamics. For instance, it has been suggested (Perrier *et al.*, 1996) that the amplitude of equilibrium shifts for the different commands may be varied to maintain the basic temporal synchrony of movement. More specifically, the commands related with slower structures may overshoot the spatial endpoint of the movement. Or, the different structures could well be controlled by commands that have a com-

pletely different temporal structure, and synchrony of movements could emerge as the result of mechanical interactions. It must be noted, however, that these and other hypotheses can only be addressed by interpreting empirical observations by means of a realistic physiological model that accounts for the mechanical interaction among the different structures.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grant No. DC-00594 from the National Institute on Deafness and Other Communication Disorders—USA, NSERC—Canada, FCAR—Québec, and by a postdoctoral fellowship to VS from the University of Genova (Italy). The authors thank the Institut de Phonétique de Strasbourg (IPS) for providing the data set, and Arturo Galvan for the port of the data to MATLAB. We are also indebted to the reviewers (K. Honda, E. Saltzman, and an anonymous reviewer) for their detailed comments and suggestions.

APPENDIX: DERIVATION OF THE EQUATIONS OF MOTION

The dynamic behavior of the whole system is completely specified by its global kinetic and potential energy function.

In the case of the jaw, kinetic energy is given by $T_j = \frac{1}{2} \dot{\mathbf{q}}_j^T \cdot M_j(\mathbf{q}_j) \cdot \dot{\mathbf{q}}_j$, where

$$M_j(\mathbf{q}_j) = \begin{bmatrix} m_j[1 + y'(x_0)^2] & m_j[1y'(x_0) \cdot R(\alpha_j - \theta + \pi/2) \cdot \hat{\mathbf{x}}_{Gj}] \\ m_j[1y'(x_0) \cdot R(\alpha_j - \theta + \pi/2) \cdot \hat{\mathbf{x}}_{Gj}] & I_j \end{bmatrix} \quad (\text{A1})$$

is the matrix of inertia of the jaw; m_j and I_j are, respectively, jaw mass and its moment of inertia with respect to the center of the condyle, whereas $\hat{\mathbf{x}}_{Gj}$ is the position of the center of mass relative to the center of the condyle, θ is jaw orientation at occlusion (constant, estimated from the x-ray data), and $y'(x_0) = dy(x)/dx|_{x_0}$.

The matrix $R(\alpha)$ represents a rotation of an angle α . The only contribution to potential energy is that of gravity and, therefore, supposing that gravity is directed vertically:

$$V_j = m_j g [x_0 \cos \theta + y(x_0) \sin \theta + \hat{x}_{Gj} \cos \alpha_j + \hat{y}_{Gj} \sin \alpha_j],$$

where \hat{x}_{Gj} and \hat{y}_{Gj} are the components of $\hat{\mathbf{x}}_{Gj}$.

In the case of the hyoid bone, kinetic energy is defined as $T_h = \frac{1}{2} \dot{\mathbf{q}}_h^T \cdot M_h \cdot \dot{\mathbf{q}}_h$, where the matrix of inertia, diagonal, is

$$M_h = \begin{bmatrix} m_h & 0 & 0 \\ 0 & m_h & 0 \\ 0 & 0 & I_h \end{bmatrix}, \quad (\text{A2})$$

whereas the potential energy is given by $V_h = m_h g y_{Gh}$.

Finally, the kinetic and potential energy functions for the larynx are simply expressed, respectively, by $T_l = 1/2 m_l \dot{y}_l^2$ and $V_l = m_l g y_l$.

As regards the tongue, kinetic energy is $T_t = 1/2 d_t \int \dot{\mathbf{u}}^T \cdot \dot{\mathbf{u}} dV$, where \mathbf{u} is the deformation field and d_t is tongue density, assumed uniform; potential energy is $V_t = 1/2 \int \boldsymbol{\sigma}^T \cdot \boldsymbol{\varepsilon} dV - d_t \mathbf{g}^T \cdot \int \mathbf{u} dV$, where $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ are, respectively, the stress and the strain fields; \mathbf{g} is the constant gravity acceleration.

Strain is the gradient of the deformation field, i.e.,

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix} \mathbf{u}. \quad (\text{A3})$$

Stress is related to strain through the Hooke's law that, in the case of the plane strain hypothesis, is defined as

$$\boldsymbol{\sigma} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & (1-2\nu)/2 \end{bmatrix} \boldsymbol{\varepsilon}. \quad (\text{A4})$$

After application of FE techniques, the above expressions for T_t and V_t can be rewritten as $T_t = 1/2 \dot{\mathbf{x}}^T \cdot M_t \cdot \dot{\mathbf{x}}$ and $V_t = 1/2 (\mathbf{x} - \mathbf{x}_0)^T \cdot K_t \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{f}_g^T \cdot \mathbf{x}$, where \mathbf{x}_0 is the rest, undeformed configuration and the constant matrices M_t and K_t are, respectively, tongue inertia and stiffness. The constant vector \mathbf{f}_g represents the distribution of tongue weight on the nodes of the mesh.

Some of the nodes of the tongue mesh are fixed with respect to the jaw or the hyoid bone; we will suppose that nodes are ordered so that we can write: $\mathbf{x} = [\mathbf{x}_t^T \mathbf{x}_j(\mathbf{q}_j)^T \mathbf{x}_h(\mathbf{q}_h)^T]^T$.

The equations of motion that describe the dynamics of the jaw-hyoid-tongue-larynx system can be obtained from the global Lagrangian function, $L = T - V$, where $T = T_t + T_j + T_h + T_l$ and $V = V_t + V_j + V_h + V_l$ are, respectively, the total kinetic and potential energy. We can define a global configuration vector for the jaw-hyoid-tongue-larynx system: $\mathbf{Q} = [\mathbf{x}_t^T \mathbf{q}_j^T \mathbf{q}_h^T y_l]^T$.

The global kinetic energy can be rewritten as $T = 1/2 \dot{\mathbf{Q}}^T \cdot M(\mathbf{Q}) \cdot \dot{\mathbf{Q}}$, where the "global" matrix of inertia, $M(\mathbf{Q})$, is now defined as

$$M(\mathbf{Q}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & M_j(\mathbf{q}_j) & 0 & 0 \\ 0 & 0 & M_h & 0 \\ 0 & 0 & 0 & m_l \end{bmatrix} + J^T(\mathbf{Q}) M_t J(\mathbf{Q}) \quad (\text{A5})$$

and the "global" jacobian matrix, $J(\mathbf{Q})$, is given by

$$J(\mathbf{Q}) = \begin{bmatrix} I_t & 0 & 0 & 0 \\ 0 & J_j(\mathbf{q}_j) & 0 & 0 \\ 0 & 0 & J_h(\mathbf{q}_h) & 0 \end{bmatrix}, \quad (\text{A6})$$

where I_t is an unit matrix, $J_j(\mathbf{q}_j) = \partial \mathbf{x}_j / \partial \mathbf{q}_j$, and $J_h(\mathbf{q}_h) = \partial \mathbf{x}_h / \partial \mathbf{q}_h$. Similarly, the "global" potential energy can be rewritten as $V = V(\mathbf{Q})$.

It is now possible to derive the global equation of motion for the jaw-hyoid-tongue-larynx system, which has the form

$$M(\mathbf{Q}) \cdot \ddot{\mathbf{Q}} + C(\mathbf{Q}, \dot{\mathbf{Q}}) \cdot \dot{\mathbf{Q}} = \mathbf{G}(\mathbf{Q}) + J_t(\mathbf{Q})^T \cdot \mathbf{f}_m(\mathbf{l}, \lambda), \quad (\text{A7})$$

where

$$C_{ij}(\mathbf{Q}, \dot{\mathbf{Q}}) = \sum_k \partial M_{ij} / \partial Q_k \cdot \dot{Q}_k - 1/2 \sum_k \partial M_{kj} / \partial Q_i \cdot \dot{Q}_k$$

defines a velocity-dependent interaction term, and $J_L(\mathbf{Q}) = \partial \mathbf{l} / \partial \mathbf{Q}$ is the Jacobian of the transformation between the configuration \mathbf{Q} and the vector $\mathbf{l} = \mathbf{l}(\mathbf{Q})$ of muscle lengths. The term $\mathbf{G}(\mathbf{Q}) = -\partial V / \partial \mathbf{Q}$ accounts for gravity and for the passive elastic properties of the tongue, whereas $\mathbf{f}_m(\mathbf{l}, \lambda)$ is the vector of muscle forces.

The above equation completely describes the forward dynamics of the whole system, relating the forces generated by muscles to motion generated in the mechanical degrees of freedom of the system. The equation was numerically integrated by using Gear's algorithm for stiff systems. Jaw, hyoid, larynx, and tongue motions were therefore obtained simultaneously.

- Baer, T., Alfonso, P. J., and Honda, K. (1988). "Electromyography of the tongue muscles during vowels in /pVp/ environment," *Annu. Bull., Res. Inst. Logoped. Phoniatr. Univ. Tokyo* **22**, 7-19.
- Baragar, F. A., and Osborn, J. W. (1984). "A model relating patterns of human jaw movement to biomechanical constraints," *J. Biomech.* **17**, 757-767.
- Bothorel, A., Simon, P., Wioland, F., and Zerling, J. P. (1986). "Cin rudiographie des voyelles et consonnes du fran ais," Technical report, Travaux de l'Institut de Phon tique de Strasbourg, France.
- Chiel, H. L., Crago, P., Mansour, J. M., and Hathi, K. (1992). "Biomechanics of a muscular hydrostat: a model of lapping by a reptilian tongue," *Biol. Cybern.* **67**, 403-415.
- Dickson, D. R., and Maue-Dickson, W. (1982). *Anatomical and Physiological Bases of Speech* (Little, Brown and Co., Boston).
- Duck, F. A. (1990). *Physical properties of tissue: A comprehensive reference book* (Academic, London).
- Feldman, A. G., and Orlovsky, G. N. (1972). "The influence of different descending systems on the tonic reflex in the cat," *Exp. Neurol.* **37**, 481-494.
- Feldman, A. G., Adamovich, S. V., Ostry, D. J., and Flanagan, J. R. (1990). "The origins of electromyograms—Explanations based on the equilibrium point hypothesis," in *Multiple Muscle Systems—Biomechanics and Movement Organization*, edited by J. M. Winters and S. L.-Y. Woo (Springer-Verlag, New York).
- Gribble, P. L., Ostry, D. J., Sanguinetti, V., and Laboissiere, R. (1998). "Are complex computations required for the control of arm movement?," *J. Neurophysiol.* (in press).
- Harshman, R., Ladefoged, P., and Goldstein, L. (1977). "Factor analysis of tongue shapes," *J. Acoust. Soc. Am.* **62**, 693-707.
- Hashimoto, K., and Suga, S. (1986). "Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue," *J. Acoust. Soc. Jpn. (E)* **7(1)**, 39-46.
- Henneman, E., Somjen, G., and Carpenter, D. O. (1965). "Functional significance of cell size in spinal motoneurons," *J. Neurophysiol.* **28**, 560-580.
- Honda, K., Hirai, H., and Dang, J. (1994). "A physiological model of speech production and the implication of tongue-larynx interaction, in *International Congress on Spoken Language Processing*, Yokohama, pp. 175-178.
- Huxley, A. F. (1957). "Muscle structure and theories of contraction," *Prog. Biophys. Chem.* **7**, 255-318.
- Joyce, G. C., and Rack, P. M. H. (1969). "Isotonic lengthening and shortening movements of cat soleus muscle," *J. Physiol. (London)* **204**, 475-491.
- Kiritani, S., Miyawaki, K., and Fujimura, O. (1976). "A computational model of the tongue," *Res. Instit. Logoped. Phoniatr. Annu. Bull.* **10**, 243-252.
- Laboissiere, R., Ostry, D. J., and Feldman, A. G. (1996). "Control of multi-muscle systems: Human jaw and hyoid movements," *Biol. Cybern.* **74(3)**, 373-384.
- Lamarre, Y., and Lund, J. P. (1975). "Load compensation in human masseter muscles," *J. Physiol. (London)* **253**, 31-35.
- Lee, Y., Terzopoulos, D., and Waters, K. (1995). "Realistic modeling for facial animation," in *Computer Graphics Proceedings, Annual Conference Series*, pp. 55-62.
- Maeda, S. (1990). "Compensatory articulation during speech: evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht), pp. 131-149.
- Maeda, S., and Honda, K. (1994). "From EMG to formant patterns of vowels: The implication of vowel spaces," *Phonetica* **51(1-3)**, 17-29.
- McDevitt, W. E. (1989). *Functional Anatomy of the Masticatory System* (Wright, London, UK).
- Miller, A. (1991). *Craniomandibular Muscles: Their Role in Function and Form* (CRC, Boca Raton, FL).
- Miyawaki, K. (1974). "A study on the musculature of the human tongue," *Annu. Bull. Res. Instit. Logoped. Phoniatr., Univ. Tokyo* **8**, 23-50.
- M ller, E. M., Milenkovic, P. H., and MacLeod, G. E. (1984). "Perioral tissue mechanics during speech production," in *Proceedings of the Second IMAC International Symposium on Biomedical Systems Modeling*, edited by C. DeLisi and J. Eisenfeld (North-Holland, Amsterdam).
-  hman, S. E. G. (1967). "Numerical model of coarticulation," *J. Acoust. Soc. Am.* **41**, 310-320.

- Ostry, D. J., and Munhall, K. G. (1994). "Control of jaw orientation and position in mastication and speech," *J. Neurophysiol.* **71**, 1515–1532.
- Ostry, D., Gribble, P., and Gracco, V. (1996). "Coarticulation of jaw movements in speech production: Is context sensitivity in speech kinematics centrally planned?," *J. Neurosci.* **16**(4), 1570–1579.
- Ostry, D. J., Vatikiotis-Bateson, E., and Gribble, P. L. (1997a). An examination of the degrees of freedom of human jaw motion in speech and mastication, *J. Speech Lang. Hear. Res.* **40**(6), 1341–1352.
- Ostry, D. J., Gribble, P. L., Levin, M. F., and Feldman, A. G. (1997b). "Phasic and tonic stretch reflexes in muscles with few muscle spindles: Human jaw opener muscles," *Exp. Brain Res.* **116**(2), 299–308.
- Otten, E. (1987). "A myocybernetic model of the jaw system of the rat," *J. Neurosci. Meth.* **21**, 287–302.
- Payan, Y., and Perrier, P. (1997). "Synthesis of V–V Sequences with a 2D Biomechanical Tongue Model Controlled by the Equilibrium Point Hypothesis," *Speech Commun.* (in press).
- Perkell, J. (1969). *Physiology of Speech Production* (MIT, Cambridge, MA).
- Perkell, J. (1974). "A Physiologically-Oriented Model of Tongue Activity in Speech Production," unpublished doctoral dissertation, Department of Electrical Engineering, MIT.
- Perrier, P., Ostry, D. J., and Laboissière, R. (1996). "The Equilibrium Point Hypothesis and Its Application to Speech Motor Control," *J. Speech Hear. Res.* **39**, 365–378.
- Sanguineti, V., Laboissière, R., and Payan, Y. (1997). "A control model of human tongue movements in speech," *Biol. Cybern.* **77**(1), 11–22.
- Scheideman, G. B., Bell, W. H., Legan, H. L., Finn, R. A., and Reisch, J. S. (1980). "Cephalometric analysis of dentofacial normals," *Am. J. Orthod.* **78**(4), 404–420.
- Schwarz, H. R. (1984). *Finite Element Methods* (Academic, London).
- Smith, A. (1992). "The Control of Orofacial Movements in Speech," *Crit. Rev. Oral Biol. Med.* **3**(3), 233–267.
- Terzopoulos, D., and Waters, K. (1990). "Physically-based facial modeling, analysis and animation," *J. Visual Comput. Animat.* **1**, 73–80.
- Throckmorton, G. S., and Throckmorton, L. S. (1985a). "Quantitative calculations of temporomandibular joint reaction forces—I. The importance of the magnitude of the jaw muscle forces," *J. Biomech.* **18**, 445–451.
- Throckmorton, G. S., and Throckmorton, L. S. (1985b). "Quantitative calculations of temporomandibular joint reaction forces—II. The importance of the direction of the jaw muscle forces," *J. Biomech.* **18**, 453–461.
- Tsuji, T., Morasso, P. G., Goto, K., and Ito, K. (1995). "Human hand impedance characteristics during arm posture," *Biol. Cybern.* **72**, 475–485.
- van Eijden, T. M. G., Klok, E. M., Weijs, W. A., and Koolstra, J. H. (1988). "Mechanical capabilities of the human jaw muscles studied with a mathematical model," *Arch. Oral Biol.* **33**, 819–826.
- Wells, J. B. (1965). "Comparison of mechanical properties between slow and fast mammalian muscles," *J. Physiol. (London)* **178**, 252–269.
- Westbury, J. R. (1988). "Mandible and hyoid bone movements during speech," *J. Speech Hear. Res.* **31**, 405–416.
- Wilhelms-Tricarico, R. (1995). "Physiological modeling of speech production: Methods for modelling of soft-tissue articulators," *J. Acoust. Soc. Am.* **97**, 3085–3098.