Biol. Cybern. 77, 11–22 (1997)

Biological
Cybernetics

A control model of human tongue movements in speech

Vittorio Sanguineti[1], **Rafael Laboissière**[2], **Yohan Payan**[2]

[1] Dipartimento di Informatica, Sistemistica e Telematica, Università di Genova, Genova, Italy
[2] Institut de la Communication Parlée, Grenoble, France

Received: 27 August 1996 / Accepted in revised form: 25 February 1997

Abstract. Tongue movements during speech production have been investigated by means of a simple yet realistic biomechanical model, based on a finite elements modeling of soft tissues, in the framework of the equilibrium point hypothesis (λ-model) of motor control. In particular, the model has been applied to the estimation of the "central" control commands issued to the muscles, for a data set of mid-sagittal digitized tracings of vocal tract shape, recorded by means of low-intensity X-ray cineradiographies during speech. In spite of the highly non-linear mapping between the shape of the oral cavity and its acoustic consequences, the organization of control commands preserves the peculiar spatial organization of vowel phonemes in acoustic space. A factor analysis of control commands, which have been decomposed into independent or "orthogonal" muscle groups, has shown that, in spite of the great mobility of the tongue and the highly complex arrangement of tongue muscles, its movements can be explained in terms of the activation of a small number of independent muscle groups, each corresponding to an elementary or "primitive" movement. These results are consistent with the hypothesis that the tongue is controlled by a small number of independent "articulators", for which a precise biomechanical substrate is provided. The influence of the effect of jaw and hyoid movements on tongue equilibrium has also been evaluated, suggesting that the bony structures cannot be considered as a moving frame of reference, but, indeed, there may be a substantial interaction between them and the tongue, that may only be accounted for by a "global" model. The reported results also define a simple control model for the tongue and, in analogy with similar modelling studies, they suggest that, because of the peculiar geometrical arrangement of tongue muscles, the central nervous system (CNS) may not need a detailed representation of tongue mechanics but rather may make use of a relatively small number of muscle synergies, that are invariant over the whole space of tongue configurations.

1 Tongue articulators and muscle synergies

This work aimed to investigate the determinants of position and shape of the human tongue from the observation of speech movements, and to characterize them in terms of the anatomical arrangement of tongue musculature.

The tongue is one of the most important structures involved in speech production. Indeed, it is the main contributor to the shaping of the oral cavities and consequently to the acoustic outcome.

As the human tongue is highly mobile and deformable, with a virtually infinite number of mechanical degrees of freedom and no clearly identifiable anatomic landmarks, it is difficult to specify and to characterize its shape and position in the oral cavity. Empirical observations have suggested that, in spite of its virtually infinite number of mechanical degrees of freedom, its contribution to the vocal tract shape may be determined by a relatively small number of independently controlled components, or articulators.

The proposed tongue articulators have typically been associated with features of the vocal tract profile, which have been characterized in terms of either a simple geometric description (Mermelstein 1973) or some form of factor analysis (Harshman et al. 1977, Maeda 1988). However, the hypothesis that tongue control may be characterized in terms of a small number of articulators should not be seen just as a useful descriptive tool; indeed, it may have important consequences in the organization of the control level.

For instance, based on the observation that the movements related to consonants and vowels tend to be independent of each other, Öhman (1966) argued that there may be at least three independent tongue articulators, namely body (responsible for the front/back and up-down motions observed in vowels), apex (related to tongue tip movement) and dorsum (related to arching-flattening); the last two mostly involved in the generation of consonants. Each articulator may correspond to a different mechanical system, consisting of different muscle groups, perhaps partly overlapping.

Similarly, Perkell (1969) has suggested that separate sets of muscles control the position and shape of the tongue, and, especially, each articulator may be associated with a "small" number of muscles, two or three at most. In particular, based on some differences in their physiology, it was suggested that

---

Correspondence to: V. Sanguineti, Dipartimento di Informatica, Sistemistica e Telematica, Università di Genova, Via Opera Pia 13, I-16145 Genova GE, Italy
(Tel.: +39-10-353-2801, Fax: +39-10-353-2154,
e-mail: sangui@dist.unige.it)

extrinsic muscles are mainly responsible for positioning the tongue in the oral cavity (and, therefore, for vowel gestures) whereas the intrinsic muscles are responsible for controlling tongue shape (and, therefore, for consonants).

Recently, Maeda and Honda (1994) have attempted to relate articulators, defined "a priori" in terms of geometric features of the mid-sagittal profile of the oral cavity (Maeda 1988), to different sets of muscles, based on the electromyographic (EMG) activity measured in selected vocal tract configurations. However, this study does not consider that the specific articulators themselves, i.e. the possibility for the CNS to independently control, for instance, tongue position and shape, are likely to result both from the geometrical arrangement of the muscles and from the organization of their neural control. Therefore, the identification of articulators, as well as that of their associated muscle synergies, has to be carried out by taking into account tongue biomechanics.

In this paper, the problem of tongue articulators is investigated by means of a simple biomechanical model of the tongue, in the framework of the equilibrium point (EP) hypothesis of motor control (λ-model) (Feldman 1966, Feldman et al. 1990, Feldman and Levin 1995). In particular, a model-based analysis of empirical data on tongue movements during speech has allowed us to identify a number of muscle groups or "synergies" that are "maximally independent" from the mechanical point of view, thus identifying articulators on the basis of the geometric arrangement of tongue muscles.

2 The model

In general, the observable effect of central commands on generated movements depends on muscular and neural dynamics, as well as on the interaction with the external environment (for instance, contact forces). Therefore, their identification on the basis of empirical data requires the modeling of tongue mechanics at the muscle level, as well as of its neural control.

2.1 Previous work

The tongue and similar motor organs, like the trunks of elephants and the tentacles of octopuses, are generally referred to as muscular hydrostats (Smith and Kier 1989). In these structures, muscles as well as other soft tissues are used for both support and movement generation, by exploiting the fact that, because of their high content of water, they are virtually incompressible.

A major feature of muscular hydrostats is that expansion in a given dimension implies a contraction in at least one other dimension, so that movements in a given direction may be generated by muscles with an orthogonal direction of action. For instance, in the reptilian tongue, the observed large and fast protrusion movements are generated by the contraction of the transverse muscles (Chiel et al. 1992) that, because of volume conservation, indirectly yield an expansion in the longitudinal direction. In humans, a similar mechanism is believed to be responsible for positioning the tongue in the front/upper part of the oral cavity, like in /i/. EMG studies (Miyawaki et al. 1975, Alfonso et al. 1982; Baer et al. 1988) have suggested that movements for /i/ are achieved indirectly, by contracting the muscles in the tongue floor, mainly the posterior section of the genioglossus.

A first attempt to model the structure of the human tongue at the muscular level, based on anatomical data, was reported by Perkell (1974). In that work, the geometry of the mid-sagittal section of the tongue was approximated by a lattice of mass-bearing points, connected to each other by viscoelastic elements composed of "active" (i.e. controllable) and "passive" components, so that segments corresponding to the same muscle shared the same control command. Volume conservation was accounted for by assuming that variations of the transverse tongue section are neglectable and by hypothesizing that each region of the lattice has an elastic behaviour, counteracting compression as well as expansion around a "rest area" (with a high stiffness).

Most recent models of the tongue are based on the principles of continuum mechanics: the tongue is modelled as a viscoelastic body, whose mechanical behaviour is approximated by means of finite element (FE) techniques. The first 3-dimensional FE model of the tongue, largely based on the dissection studies of Miyawaki (1974), was presented by Kiritani et al. (1976). Improved versions of the same model were reported in Kakita and Fujimura (1977), Fujimura and Kakita (1978), Kakita et al. (1985). A basic feature of this and other similar FE models (Hashimoto and Suga 1986) is that the tongue is treated as an isotropic material with a linear elastic behaviour. Volume conservation is accounted for by an appropriate choice of the Poisson ratio. Muscles are modelled as sets of external pulling forces, applied to selected nodes or, as in Hashimoto and Suga (1986), as continuous, distributed force fields. The above models demonstrated the ability to predict the muscle forces associated with selected postures (corresponding to sustained vowel utterances), which are qualitatively similar to the electromyographic (EMG) activity observed for these same postures. In the case of Hashimoto and Suga (1986), muscle forces were estimated by means of an optimization procedure, aimed at fitting the mid-sagittal shape of the tongue with that observed in X-ray data samples.

These FE models are based on the hypothesis of small and quasi-static deformations, whereas Payan et al. (1995) and Wilhelms-Tricarico (1995) also take into account dynamics. The first model is two-dimensional, and the soft tissue is supposed to obey Hooke's law of linear elasticity; muscles are modelled as spring-like elements in terms of the λ-model (Feldman et al. 1990) of motor control, and dynamics at tissue as well as at muscle level is accounted for. The second is three-dimensional and accounts for the non-linearities inherent in biological soft tissues. Muscles are modelled there as active stress generators, where stress is dependent on fibre elongation and elongation rate, as well as on an activation command. However, complex models must face the lack of available empirical knowledge about the characterization of tongue soft tissue and the force capability of the muscles.

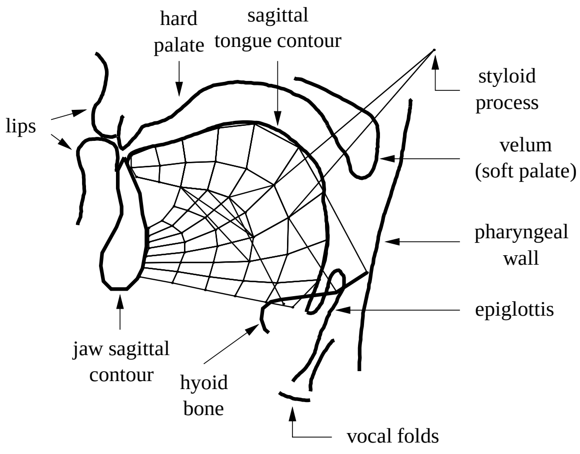


Fig. 1. Finite element (FE) model of the mid-sagittal section of the human tongue. Some of the vocal tract structures shown here (like the lips and the vocal folds) are not taken into account in the present work and are only depicted for reference purposes

2.2 *The "hydrostatic skeleton" of the tongue*

As the focus of this study is on the organization of central commands to muscles, only the very basic aspects of tongue structure have been modelled. The model has been restricted to the mid-sagittal plane for two reasons: first, the changes in the mid-sagittal contour of the tongue have the greatest effect on the acoustical outcome (at least in the lower range of acoustical frequencies); second, most of the available empirical data on tongue position and shape during speech movements are two-dimensional.

In order to keep the model simple while powerful enough to describe the observed tongue deformation in our database, the tongue has been coarsely discretized into 39 elements and 54 nodes; some of them are fixed with respect to the mandible or to the hyoid bone, (see Fig. 1).

The geometrical arrangement of the elements is similar to that of Perkell (1974), reflecting the directions of muscle fibres, with some geometric adaptation to the anatomy of the speaker who produced the empirical data that were used (see Sect. 3). For a preliminary version of the model, see Sanguineti (1995) and Laboissière et al. (1995).

The "hydrostatic skeleton" of the tongue is modelled as an isotropic continuous material, whose elastic behaviour, assumed to be linear, is completely specified by two parameters, namely Young's modulus, E, and Poisson's ratio, ν.

There are no simple ways to account for volume conservation in a planar model, because this would require precise assumptions on the amount of transverse deformation. Following Perkell (1974), as a first approximation we indeed assumed that there is no strain (and therefore deformation) in the transverse direction, and that the X and Y components of strain only depend on position in the XY plane (namely, the plane strain hypothesis). Moreover, we choose for ν a value of $0.49 \approx 1/2$ which corresponds, by definition, to volume conservation at the microscopic level. It should be noted, however, that the above assumptions only approximate area conservation for the whole tongue.

No value for Young's modulus is available in the literature for tongue tissue. Duck (1990) reports a value of 6.2 kPa for the Young's modulus of human muscles at rest, and in the present model a value of E = 10 kPa was empirically found necessary to balance the deformations observed in vowel gestures with the muscles activated at about 20% of their maximum strength, as suggested in Müller et al. 1984.

The passive elastic behaviour of the tongue is summarized by the stiffness matrix K, computed by means of standard FE techniques (Schwarz 1984), by assuming a uniform thickness of 4 cm, which relates a displacement of the nodes from their rest configuration to the corresponding vector restoring forces. The effect of gravity was accounted for by assuming a density of 1040 $\mathrm{kg/m^3}$, reported by Duck (1990) for muscle tissue.

We have focused on how tongue shape and position are affected by the tongue musculature itself and have made the following simplifying assumptions: (i) the effect of the bony parts (i.e. the mandible and the hyoid bone) is only to establish a moving frame of reference with respect to the upper skull; (ii) their position is assumed to be unaffected by the activation of tongue muscles; (iii) the movements of the jaw and the hyoid bone are described by pure translations in the horizontal and vertical directions. We did not take rotation of the jaw and hyoid into account, because they were difficult to estimate accurately from the X-ray data on which the model was based.

The interaction of the tongue with the palatal walls has been also accounted for: contact forces are assumed to be elastic (depending on the level of "penetration" of each node into the palate) and directed normally (corresponding to the hypothesis of zero friction).

2.3 *Muscle geometry*

The inherently distributed geometrical arrangement of tongue muscles has been approximated by a discrete number of polygonal lines, hereafter referred as "macro-fibres", which connect different nodes of the FE lattice, as well as points on the bony structures. Each macro-fibre has its own control command and therefore can be independently activated. In other words, it will be treated as a separate muscle, the length of which is obtained as the sum of the lengths of its segments. Although this may seem to be in conflict with empirical evidence, it will be shown that macro-fibres belonging to the same muscles are functionally grouped at the control level.

The geometry of muscles essentially reflects that reported by Perkell 1974, with some modifications suggested by other sources (Miyawaki 1974, Dickson and Maue-Dickson 1982). Three extrinsic and three intrinsic muscles were included in the model. The extrinsic muscles include: (i) genioglossus. As EMG measurements (Miyawaki et al. 1975) suggest that different portions are activated separately, this muscle has been modelled by seven independent macro-fibres, $\mathrm{GG}_{1\text{–}7}$; see Fig. 2a. (ii) Hyoglossus. Its inherently distributed structure has been accounted for by three different macro-fibres, $\mathrm{HG}_{1\text{–}3}$; see Fig. 2b. (iii) Styloglossus. As in Perkell (1974), it has been modelled by two macro-fibres, SG_1 and SG_2, see Fig. 2c.

Unlike models such as Perkell (1974) and Kakita (1985), we have not included the palatoglossus and the pharyngeal constrictors because, in spite of the fact that they are attached to the tongue, their contribution is mainly to the movement of other structures (velum and pharyngeal walls, respectively) and thus may only marginally affect tongue positioning and shaping. Moreover, we have not considered the mylohyoid and the geniohyoid because, strictly speaking, they are not tongue muscles. Although they are attached only to the hyoid bone and to the jaw, we acknowledge that they may have an effect on tongue lowering/raising due to the strengthening of the tongue floor.

The tongue also includes four intrinsic muscles (superior longitudinalis, inferior longitudinalis, verticalis, and transverse), but only the first three are represented in the present model (transverse is oriented in a transverse direction and is not modelled). (i) superior longitudinalis. Its rather short fibres suggest the possibility for independent contraction of different parts of the muscle (Dickson and Maue-Dickson 1982). Therefore, it has been modelled by seven separate macro-fibres, $SL_{1\text{–}7}$, depicted in Fig. 2d. (ii) inferior longitudinalis. As in Perkell (1974) it has been modelled by two macro-fibres, IL_1 and IL_2, displayed in Fig. 2e. (iii) verticalis. Following Perkell (1974), we have only modelled its most anterior part by two macro-fibres, V_1 and V_2, displayed in Fig. 2f.

In summary, the six muscles in the model are represented by 23 independently controlled macro-fibres.

2.4 Neural control and force generation

As pointed out by Feldman (1966), the degree of muscle activation is not specified centrally but rather results from the interaction of a central command with the afferent inputs provided by muscle spindles and other proprioceptive afferents. Empirical observations have suggested that, in intact limb muscles, the relationship between muscle length and force is described by a family of length-tension curves – the so-called *invariant characteristics* (ICs) – each characterized by a different threshold length for motorneuron recruitment, which can be centrally controlled. This has suggested that muscles and the segmental level, as a whole, behave as non-linear springs with a centrally controlled rest length, which corresponds to the muscle length, λ, at which motorneuron recruitment begins: the so-called λ-model (Feldman 1966, Feldman et al. 1990).

Fundamentally, the λ-model relies on the use of afferent information at the segmental level, typically muscle spindles. In the tongue, spindles are widely distributed in both extrinsic (Walker and Rajagopal 1959) and intrinsic (Cooper 1953) muscles. Low-threshold mechanoreceptors in the tongue may also convey proprioceptive information related to muscle length (Fitzgerald and Sachithanandan 1979).

Afferent information from tongue muscle spindles is mainly carried by the lingual nerve, which projects to the hypoglossal nucleus of the brainstem, with different sensitivities to different stretch directions (Bowman and Combs 1968, Lowe 1978a); hypoglossal motorneurons control the muscles of the tongue (Bowman and Combs 1968, Lowe 1978a).

Animal studies have shown that lingual nerve stimulation excites retractive (R) motorneurons (i.e. those responsible for innervating the styloglossus and the hyoglossus), whereas the effect on protrusive (P) motorneurons (i.e. those basically controlling the genioglossus) is mainly inhibitory (Lowe 1978a). Excitation of P motorneurons (and therefore tongue protrusion) was observed to arise as a result of jaw opening (the jaw-tongue reflex) and was demonstrated to be mainly evoked by the proprioceptive receptors of the temporo-mandibular joint (Lowe 1978b). In the rat, Tolu et al. (1994) found that hypoglossal motorneurons are modulated by the activity of muscle spindles in the masseter, and by mechanical stimulation of periodontal receptors.

The importance of afferent information (mainly due to muscle spindles) in the movements of the tongue involved in speech production has been speculated by Bowman (1971), but comparable stretch reflex responses in tongue muscles have not been reported in humans.

For instance, by using rhythmical, slowly varying stretch stimuli, Neilson et al. 1979 failed to observe a tonic stretch reflex in genioglossus, styloglossus, mylohyoid, geniohyoid, tongue intrinsics, jaw openers and lip muscles; in contrast, the same stimuli evoked tonic stretch responses in jaw closers. It has been suggested (Smith 1992) that the reflex may not have been observed because of the type of stimulus. Moreover, in contrast to the above results, an unloading reflex was indeed observed (Ostry et al., in press) in jaw openers (that have few or no muscle spindles).

In the present model, as in Laboissière et al. (1996), muscle tension is assumed to be a function of both the length and a centrally controllable threshold length, λ, at which motorneuron recruitment begins. The form of length-tension relationship used in the present model is based on empirical observations (Feldman and Orlovsky 1972) that in the intact muscle a linear relationship arises between stiffness and force. As noted by Feldman (1966), this amounts to an exponential length-tension curve. In mathematical terms, if l is the length of the muscle and $A = [l - \lambda]^+$, where $[x]^+ = \max\{x, 0\}$, is a measure of motorneuron activity, muscle force is given by:

$$f_m = \rho\left[\exp(cA) - 1\right] \tag{1}$$

In terms of the more traditional Hill-type muscle models, the above equation summarizes the effects of the contractile and series elastic elements and, in addition, the size principle and the contribution of reflexes. The linear elastic behaviour of the hydrostatic skeleton plays the role of the passive, parallel elastic element. As we only consider static postures, the dynamic effects (dependence of force on shortening speed, calcium-mediated dynamics of muscle contraction, dependence of stretch reflex on speed, reflex delays) are neglected.

As regards the parameters of the muscle model (c and ρ), it has been suggested (Feldman and Orlovsky 1972) that c is constant for each muscle, and in particular $c = 112\ \mathrm{m}^{-1}$ (Laboissière et al. 1996) – a value estimated from the data reported by Feldman and Orlovsky (1972). The value of ρ can be assumed to be proportional to the muscle cross-section area (CSA) and has been taken to be equal to maximum force.

The values of the CSAs for tongue muscles were estimated from anatomic atlases (Gerhardt and Frommhold

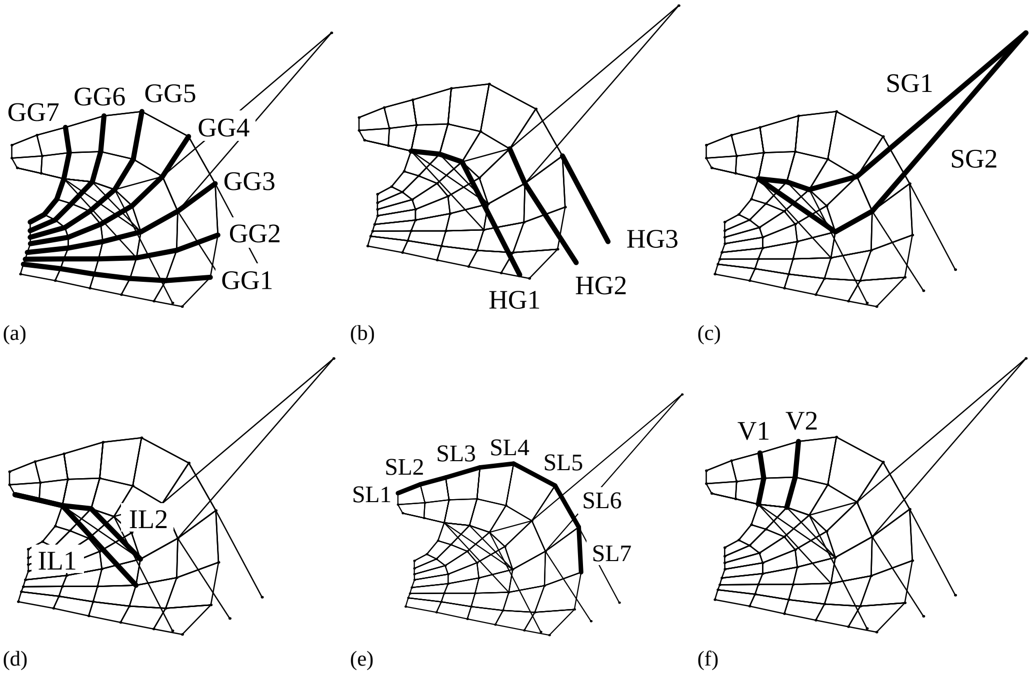


Fig. 2a–f. Muscle macro-fibres (indicated by *bold lines* superimposed on the tongue skeleton). *Top* Extrinsic muscles of the tongue: **a** genioglossus, **b** hyoglossus, **c** styloglossus. *Bottom* Intrinsic muscles: **d** inferior longitudinalis, **e** superior longitudinalis, **f** verticalis

Table 1. Estimated strengths of tongue muscles

Muscle	CSA (mm^2)	${f_m}^{\max}$ (N)
Genioglossus	309	68.0
Hyoglossus	296	65.1
Styloglossus	110	24.2
Superior longitudinalis	88	19.4
Inferior longitudinalis	65	14.3
Verticalis	66	14.5

1988, Gambarelli et al. 1977) and from the detailed drawings of tongue sections provided by Miyawaki (1974). For some intrinsic muscles, namely longitudinalis, it was not possible to obtain a direct estimate of CSA because their fibres interdigitate with those of other muscles; an “effective” CSA was thus roughly estimated by dividing the estimated value by the number of observed intermixed muscles.

Maximum force estimates for each muscle were obtained by multiplying each CSA by the maximal tetanic tension of 22 N/cm^2, reported by Wilhelms-Tricarico (1995) for the geniohyoid muscle. The maximum force estimates for each muscle are summarized in Table 1. Finally, the value of ρ for each macro-fibre was calculated by dividing the above values by the number of macro-fibres (in other words, these were assumed to contribute equally to the total muscle strength).

3 From vocal tract shapes to central commands

The model was used to infer the central commands corresponding to a set of experimentally observed tongue shapes. The available data set is composed of 519 images of the mid-sagittal section of the oral cavity, acquired by means of low-intensity X-rays at a sampling rate of 50 Hz. The subject is a female native French speaker, and the spoken material is composed of 10 sentences that can be considered as a significant sample of the phonemic content of the French language. The X-ray images, acquired at the *Institut de Phonétique de Strasbourg* (IPS), were digitized, and the contours of the vocal tract and other anatomical structures were extracted (Bothorel et al. 1986).

Assuming that the dynamic effects (inertia, viscosity, reflex delays and graded force development in muscles) can be neglected (this corresponds to the assumption that the empirical data correspond to movements that are infinitely slow), for each frame the tongue configuration satisfies the static equilibrium condition

$$J_m(\mathbf{x})^T \mathbf{f}_m + K(\mathbf{x} - \mathbf{x}_0) + \mathbf{f}_g + \mathbf{f}_c = \mathbf{0} \tag{2}$$

where $\mathbf{x}$ and $\mathbf{x}_0$ are, respectively, the present and the rest configuration (i.e. that corresponding to zero passive stress), expressed as the vector containing both the horizontal and the vertical coordinates of each node, i.e.:

$$\mathbf{x} = [x_1\, y_1\, x_2\, y_2 \cdots x_{54}\, y_{54}]^T \tag{3}$$

K is the stiffness matrix, describing the passive elastic behaviour of the tongue; $\mathbf{f}_g$ and $\mathbf{f}_c$ represent, respectively, the contributions of gravity (distributed throughout the tongue) and of contact forces against the palate (if any), for each node; $\mathbf{f}_m = \mathbf{f}_m(\mathbf{l} - \boldsymbol{\lambda})$ is the vector of the forces generated by the muscles, where the vectors $\mathbf{l}$ and $\boldsymbol{\lambda}$ represent, respectively, the actual and threshold lengths; and $J_m(\mathbf{x}) = \partial\mathbf{l}/\partial\mathbf{x}$ is the Jacobian of the transformation $\mathbf{l} = \mathbf{l}(\mathbf{x})$.

If, in addition, the equilibrium configuration is required to fit the observed tongue profile, the nine nodes on the upper surface are only free to move on the observed contour, i.e. they only have a single degree of freedom each.

The problem of finding a vector $\boldsymbol{\lambda}$ of muscle commands and a configuration $\mathbf{x}$ satisfying the equilibrium condition (2) and fitting a given tongue profile has, in general, infinite solutions, but a single vector of commands can be obtained by specifying a cost index to be minimized. We will consider, in particular, the square norm of the vector of muscle forces:

$$\mathscr{J} = \sum_i {f_m}_i^{\,2} = \|\mathbf{f}_m\|^2 \tag{4}$$

that can be interpreted as a measure of the level of muscle co-contraction.

The inversion procedure, consisting of the minimization of $\mathscr{J}$ (4) under the constraints of (i) static equilibrium (2) and (ii) fitting to the observed tongue profile, was applied to the whole data set yielding, for each frame $n = 1, \ldots, 519$, an equilibrium configuration $\mathbf{x}^n$ and a value for the corresponding muscle command, $\boldsymbol{\lambda}^n$.

Moreover, for each frame n of the data set, the reference positions of the jaw, $\mathbf{x}_j^n$ (top of inferior incisor) and the hyoid bone, $\mathbf{x}_h^n$ (inferior aspect of hyoid body) were extracted from the X-ray image, thus determining the positions of the nodes that are assumed to be rigidly connected to them.

The $\boldsymbol{\lambda}^n$'s estimated by this procedure do not correspond, in general, to the commands that are issued by the CNS, because the inversion procedure does not account for dynamics at the skeletal, muscular and neural level (the above might only be true if the movements contained in the data set were infinitely slow). However, they still reflect how muscle commands are varied during tongue movements and, therefore, the topology and the dimensionality of the command space.

4 Results

4.1 Characterization of muscle commands for sustained vowels

The data set contains samples of vowels pronounced in contexts where they can be considered close to sustained ones: for example, stressed vowels at the end of sentences. The corresponding tongue configurations, obtained from the fitting procedure, are depicted in Fig. 3.

In the space of formants (i.e. the acoustic resonant frequencies of the vocal tract), vowel phonemes are approximately arranged to form a triangle (whose vertices are represented, respectively, by /i/, /a/, /u/). The acoustic outcome of the above "canonic" vowels is depicted in Fig. 4 (bottom). The computed $\boldsymbol{\lambda}$ commands corresponding to these same phonemes appear to preserve such a geometrical arrangement: the vowel triangle, projected over the first three principal directions of the $\boldsymbol{\lambda}$ space (i.e. the first three principal components of $\boldsymbol{\lambda}^n$), is shown in Fig. 4 (top).

For static postures and for each muscle, the value of $A = [l - \lambda]^+$ can be associated with the activation of motorneuron pools and therefore with the empirically observed EMG signal (Feldman et al. 1990). In the case of the sustained vowels in the present data, independent estimations of the tonic EMG of tongue muscles are reported in Miyawaki et al. (1975), Alfonso et al. (1982), Baer et al. (1988). The values of A obtained from the fitting procedure are shown in Fig. 5.

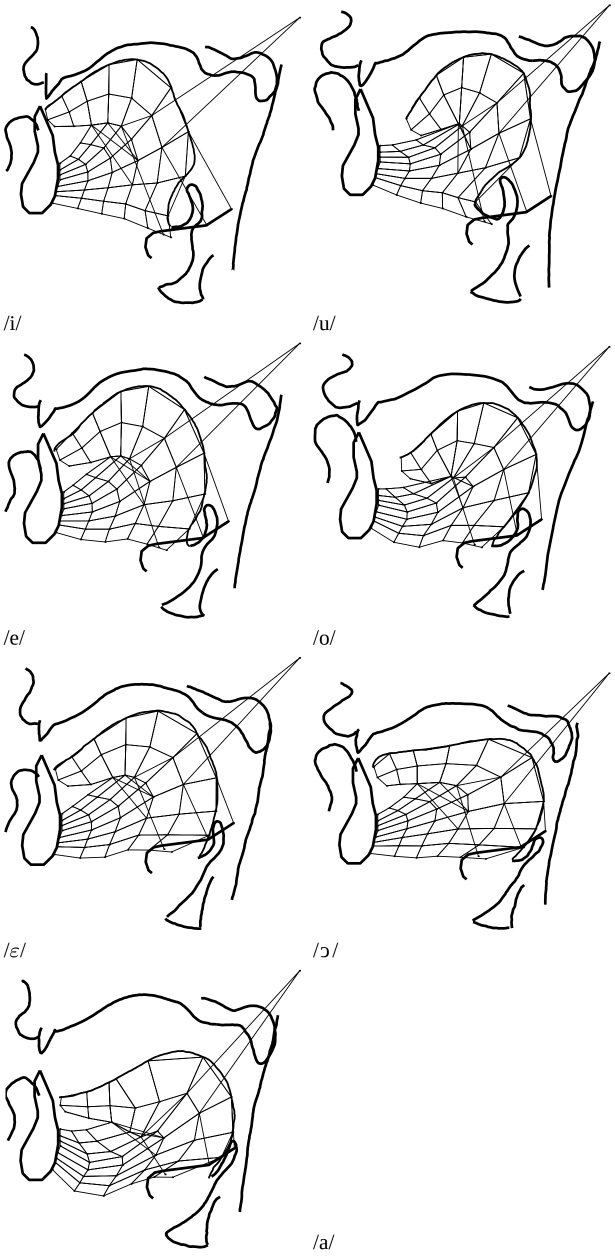


Fig. 3. Equilibrium tongue configurations resulting from the fitting procedure, corresponding to "sustained" French vowels. /i/: Met tes beaux hab**i**ts; /u/: Donne un petit c**ou**p; /e/: Chevalier du gr**é**; /o/: Met tes b**eau**x habits; /ε/: Pr**ê**te lui seize écus; /ɔ/: Une rép**o**nse ambigüe; /a/: Il fume son tab**a**c

Although the model is pretty simple and, in particular, cannot account for effects like the changes in shape along the transverse direction and the action of the muscles that attach to both jaw and hyoid (like the mylohyoid), that are believed to contribute indirectly to tongue elevation (Hashimoto and Suga (1986)), some trends are indeed captured, at least in a qualitative way.

For instance, as argued by Perkell (1969), the model does predict that vowels are mainly controlled by extrinsic muscles. However, for postures requiring bunching of the tongue, the verticalis ($V_{1\text{–}2}$) has to be significantly active;

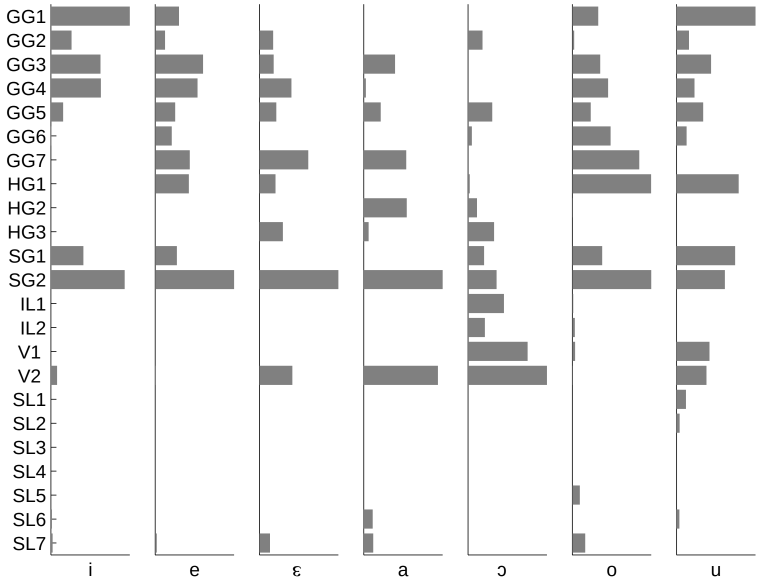


Fig. 5. Electromyographic (EMG) activity corresponding to sustained "canonical" French vowels, as estimated from the model. The bar graph displays the value of A for each muscle (normalized to the maximum A for that phoneme)

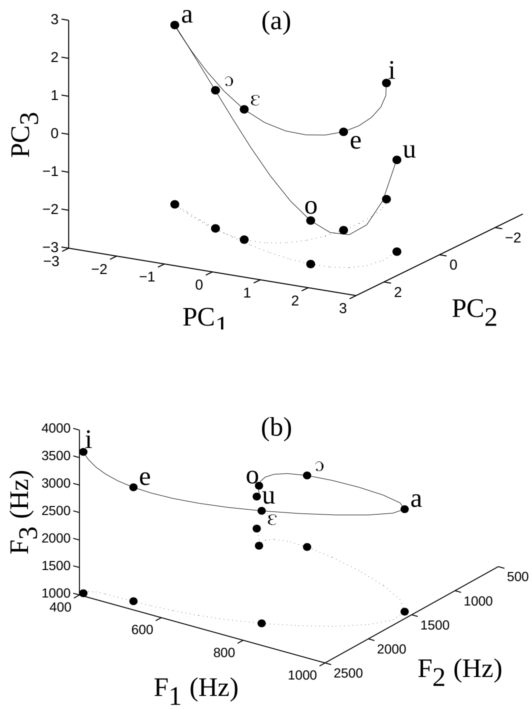


Fig. 4. Vowel triangles: (a) in muscular principal components space (PC_{1-3}) and (b) in formant space (F_{1-3}). The "canonical" vowels are represented by *dots* in both diagrams. Just for presentation purposes, a spline was computed connecting the vowels in the sequence /ieɛaɔou/ and is shown as *solid lines* in 3D or as *dotted lines* when projected to PC_1/PC_2 or F_1/F_2 planes

that happens, in particular, in /a/ and /ɔ/ and, to some extent, in /ε/ and /u/.

In accordance with the observations of Miyawaki (1975), the posterior portion of the genioglossus ($GG_{1–4}$) is active for front vowels (i.e. /i/ and /e/), whereas its anterior portion ($GG_{5–7}$) is more active in /a/ and /o/; the high level of activity in the modelled GG_1 in /u/ can be interpreted as the need for a deconstriction of the pharynx. The predicted activity of the remaining extrinsic muscles also agrees with empirical observations: styloglossus is most active in /u/ whereas hyoglossus is more active in /a/ and /o/.

The main discrepancy between empirical and model results is that SG_2 is very active for all vowels. This may be due to the fact that no passive components in muscle force were considered, or to an inaccurate modeling of the interface between soft and bony parts. For instance, the forces exerted on the tongue floor by the surrounding structure may contribute to tongue elevation; such contributions can be either passive (i.e. reaction forces) or active (i.e. generated by muscles like the mylohyoid and the geniohyoid). This is also suggested by the excess deformation of the tongue floor for vowel /a/ (see Fig. 3). More accurate predictions will probably require including jaw and hyoid muscles, with an accurate modeling of the interface between the tongue and the bony structures.

4.2 Control variables of the tongue

As suggested in Sect. 1, the individual components of the vector of muscle commands, $\boldsymbol{\lambda}$, may not be independently specified by the CNS, but $\boldsymbol{\lambda}$ may be organized in terms of a much smaller number of control variables, each of them independently controlling an individual muscle group or "articulator". Moreover, $\boldsymbol{\lambda}$ also depends on the positions of the jaw, $\mathbf{x}_j$, and the hyoid bone, $\mathbf{x}_h$, which, as stated above, play the role of a moving frame of reference (Smith 1992). In mathematical terms, this means that a mapping

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}(\mathbf{t}, \mathbf{x}_j, \mathbf{x}_h) \tag{5}$$

is defined with $\dim(\mathbf{t}) \ll \dim(\boldsymbol{\lambda})$, where the components of $\mathbf{t}$ are the *control variables* that may be independently specified by the CNS for controlling the different muscle groups that determine tongue position and shape, given a particular position of the jaw, $\mathbf{x}_j$, and of the hyoid bone, $\mathbf{x}_h$.

The problem of identifying the control variables is clearly ill-posed, in the sense that there are infinite ways to define

Table 2. Percentages of variance explanation for each principal component of jaw-hyoid position and its corresponding effect on commands to tongue muscles. Parentheses indicate cumulative values

Component	Effect on jaw-hyoid position	Effect on muscle command
1	53.9 (53.9)	72.3 (72.3)
2	26.3 (80.1)	20.1 (92.4)
3	15.6 (95.7)	6.4 (98.8)
4	4.3 (100.0)	1.2 (100.0)

the mapping of (5). However, as a first approximation, it may be hypothesized that the contributions of each muscle group to $\boldsymbol{\lambda}$ are proportional and additive, i.e. the mapping between **t** and $\boldsymbol{\lambda}$ is *linear*; the components of **t** represent the contribution of each muscle group to the global movement, and a linear shift of **t** yields a linear shift of the commands to the individual muscles.

If, in addition, the dependence of $\boldsymbol{\lambda}$ on jaw and hyoid position is also assumed to be linear (that is, changes of position of the bony parts yield proportional changes in the muscle commands), (5) simplifies to:

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_0 + \Lambda_t \mathbf{t} + \Lambda_b \mathbf{v} \tag{6}$$

where $\mathbf{v} = [\mathbf{x}_j{}^T \, \mathbf{x}_h{}^T]^T$ and $\boldsymbol{\lambda}_0$ is the mean value of $\boldsymbol{\lambda}$ over the whole data set, corresponding to the "average" configuration of the tongue.

Each column of the matrix Λ_t (respectively Λ_b) represents the contribution of a specific tongue control variable (respectively, a specific jaw or hyoid control variable) to the command variables of the individual muscles.

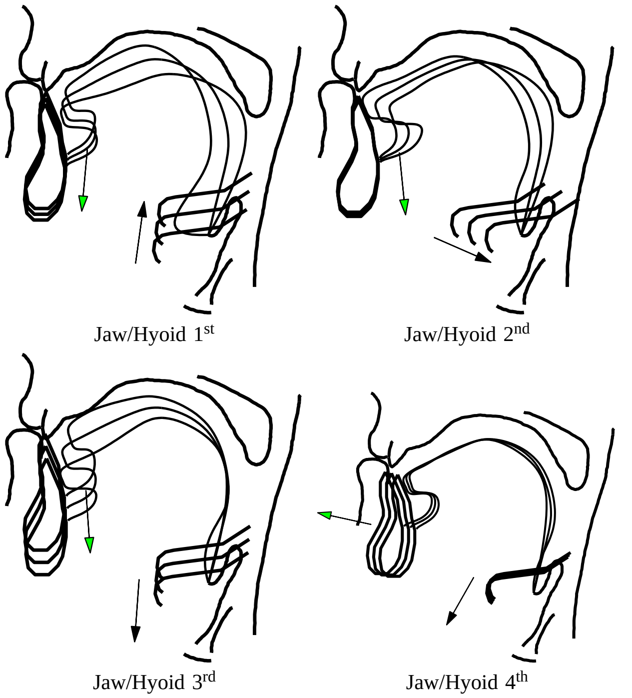
4.2.1 Dependence on jaw and hyoid movements

Let us first examine the dependence of central commands to individual muscles, $\boldsymbol{\lambda}$, on the positions of the bony parts.

During speech movements, the positions of the jaw and the hyoid bone may not be specified independently by the CNS but may result from a form of coordination. In other words, the control level may reduce the dimensionality of the system, and, therefore, the observed jaw and hyoid movements may not need all four variables (for example, the X and Y components of $\mathbf{x}_j$ and $\mathbf{x}_h$) to be completely specified. The amount of correlation between the jaw and hyoid movements can be assessed by performing a principal component analysis (PCA) over the set of vectors $\mathbf{v}^n$, $n = 1, \ldots, 519$; the results are reported in Table 2.

From these values, it may be concluded that only three control variables can account for almost 96% of the variance of the horizontal and vertical positions of the jaw and the hyoid bone, which suggests that the four mechanical degrees of freedom of the jaw-hyoid system are not independently controlled.

However, some jaw-hyoid movements may have a minimal effect on $\boldsymbol{\lambda}$. Suppose, for instance, that the jaw and the hyoid bone are rigidly translated by the same amount: this should not deform the tongue, unless it comes in contact with the surrounding structures, and the only changes in the equilibrium condition of (2) would be the contribution of styloglossus (which originates in the skull) and that of gravity. Therefore, the value of $\boldsymbol{\lambda}$ required for equilibrium is expected to remain almost the same.



Jaw/Hyoid 1st

Jaw/Hyoid 2nd

Jaw/Hyoid 3rd

Jaw/Hyoid 4th

Fig. 6. Effect of the jaw-hyoid components. For each component, *arrows* indicate the directions of movement of the jaw and the hyoid bone. Each panel displays a sequence of tongue contours corresponding, for each component, to ±3 times the standard deviation. Lips and epiglottis do not move and are only displayed for reference

Since we are interested in the control of tongue musculature, the above example suggests that the jaw-hyoid system has to be evaluated in terms of its effect on $\boldsymbol{\lambda}$. This may be obtained by performing a linear regression of **v** over $\boldsymbol{\lambda}$ in the full data set, which results in the following approximation: $\boldsymbol{\lambda} \approx \Lambda_v \mathbf{v} + \boldsymbol{\lambda}_0$. In geometric terms, in the space of muscle commands the matrix Λ_v defines a subspace that accounts for the dependence of the central commands to individual muscles, $\boldsymbol{\lambda}$, on jaw-hyoid movements.

However, in order to investigate the geometric structure of the dependence of $\boldsymbol{\lambda}$ on movements of the jaw-hyoid system, the variables defining jaw-hyoid configuration should be chosen so that their individual contributions to $\boldsymbol{\lambda}$ are uncorrelated, i.e. orthogonal. This can be achieved by determining an orthogonal system of coordinates for the subspace defined by Λ_v, thus obtaining an orthogonal matrix, $\Lambda_v^\perp$, and a new description of jaw-hyoid configuration, $\mathbf{v}^\perp$, so that $\Lambda_v \mathbf{v} = \Lambda_v^\perp \mathbf{v}^\perp$. The new variable describing jaw-hyoid configuration, $\mathbf{v}^\perp$, is thus related to **v** by the following equation: $\mathbf{v}^\perp = {\Lambda_v^\perp}^{-1} \Lambda_v \mathbf{v}$.

The results of this factor analysis (see Table 2) show, in fact, that the first two components of $\mathbf{v}^\perp$ can explain 92.4% of the total variance of the subspace defined by $\Lambda_v \mathbf{v}$. Moreover, the effect of each component of $\mathbf{v}^\perp$ on tongue shape has been evaluated by computing the equilibrium tongue position, i.e. a configuration **x** that satisfies (2), specified by the command vector $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0 + \Lambda_v^\perp \mathbf{v}^\perp$. The effect of each individual component of $\mathbf{v}^\perp$ is depicted in Fig. 6.

The first component yields an elevation of the tongue and elicits a coordinated jaw lowering and hyoid elevation; the effect of the second is a horizontal translation of the

Table 3. Percentages of variance explanation for each jaw/hyoid and tongue components. The jaw-hyoid values are those already reported in the right column of Table 2. Parentheses indicate cumulative values

Component	Percentage of variance
Jaw/hyoid 1	16.8 (16.8)
Jaw/hyoid 2	4.7 (21.4)
Jaw/hyoid 3	1.6 (23.0)
Jaw/hyoid 4	0.3 (23.3)
Tongue 1	35.0 (58.3)
Tongue 2	15.3 (73.5)
Tongue 3	6.6 (80.1)
Tongue 4	4.2 (84.3)
Tongue 5	3.6 (88.0)
Tongue 6	2.9 (90.8)

tongue, due to the combination of jaw lowering and hyoid retraction.

As expected, the other two less significant components have little effect on the control commands of tongue muscles: the third basically corresponds to a vertical translation of both the jaw and the hyoid bone, and therefore of the tongue, whereas the fourth corresponds to jaw protrusion and hyoid depression (and almost no effect on tongue shape).

In conclusion, the first three components of $\mathbf{v}^{\perp}$ (hereafter indicated as b_1, b_2 and b_3) are an appropriate description of the effect of the jaw-hyoid system on tongue configuration. Their corresponding vectors of commands are the first three columns of $\Lambda_v^{\perp}$, indicated hereafter by $\boldsymbol{\lambda}_b^1$, $\boldsymbol{\lambda}_b^2$ and $\boldsymbol{\lambda}_b^3$.

4.2.2 Control variables of the tongue

As regards the pure "tongue" components, $\mathbf{t}$, the optimal value of Λ_t (in terms of percentage of explained variance) is provided by the PCA of the residual part of the set of $\boldsymbol{\lambda}^n$.

The percentage of variance explanation over the whole data set for each "jaw-hyoid" and "tongue" component, relative to that of the whole set of $\boldsymbol{\lambda}^n$'s, is summarized in Table 3. Notice that the variances in the table are relative to the total variance of $\boldsymbol{\lambda}$ in the data set, in contrast to those reported in Table 2 that are relative to the variance of the projection onto the subspace defined by Λ_v.

In analogy with jaw-hyoid variables, the effect of each of the above commands on tongue position and shape can be assessed by looking at the corresponding effects on tongue posture, depicted in Fig. 7. It can be observed that the first two components correspond to elementary movements and have a direct counterpart in the articulators defined by Maeda (1988) on the basis of a statistical analysis of vocal tract profiles, using the same database as ours. In particular, by using the terminology introduced in Sect. 1 to qualitatively describe elementary tongue gestures, t_1 elicits an arching/flattening movement of tongue dorsum whereas t_2 yields a front/back movement of tongue body.

The third component, t_3, is mainly responsible for protrusion/retraction of tongue tip, whereas the other components, i.e. t_4, t_5 and t_6, have smaller and more localized effects (essentially changes in tongue shape) which are not easily interpretable.

The above procedure also allows us to identify the individual muscles associated with each articulator. For instance, the articulator t_i affects the muscles whose corresponding component in the ith column of $\boldsymbol{\lambda}_t$, i.e. $\boldsymbol{\lambda}_{ti}$, is non-zero. In particular, "agonist" muscles are defined as those that increase their activation as t_i increases (i.e. for which λ decreases, and thus the corresponding component of $\boldsymbol{\lambda}_{ti}$ is negative), and vice versa.

The agonist-antagonist patterns corresponding to the above components are depicted in Fig. 8. The predicted main agonist-antagonist pairs for the first two articulators (respectively, SG$_{1–2}$ and GG$_{5–7}$ – the anterior portion of genioglossus – for t_1; HG$_{1–3}$ GG$_{1–4}$ – the posterior portion of genioglossus – for t_2) are in good agreement with the available EMG data and with previous studies (Maeda and Honda 1994).

As regards t_3, it may be observed that the agonist-antagonist pairs can be identified, respectively, with transverse and longitudinal muscles, in accordance with the model of protrusion-retraction proposed by Chiel et al. (1992) for a reptilian tongue.

4.2.3 Control of tongue co-contraction

The values of $\boldsymbol{\lambda}^n$ that result from the fitting procedure, as well as the above set of control variables, minimize the level of co-contraction (see Sect. 3). In fact, as there is neurophysiological evidence (Humphrey and Reed 1983) suggesting that the CNS may separately control movement and postural behaviour (i.e. the mechanical impedance), it is appropriate to wonder if the geometrical arrangement of tongue muscles allows for an independent control of the level of co-contraction.

Co-contraction of tongue muscles may play an important role in mastication, and it has been suggested (Perrier et al. 1996) that voluntary changes in the level of co-contraction of tongue muscles may explain the variability observed in vowel reduction experiments; in particular, Perrier et al. (1996) propose that co-contraction may be varied in order to attain different levels of phonetic stress.

A hypothetical, invariant co-contraction command for the tongue can be defined as a direction in the $\boldsymbol{\lambda}$ space corresponding to little or no movement of the tongue. Moreover, along this direction, muscle force should increase monotonically (and so the global stiffness). It can be simply determined by exploring the subspace of $\boldsymbol{\lambda}$ defined by the 23 (macro-fibres) $-$ 6 (commands) = 17 principal components that were not chosen as tongue commands because they have almost no effect on tongue posture; let the columns of the matrix Λ_n describe these less significant principal directions. The subspace spawned by those 17 components should correspond to muscle commands that produce few tongue movements, and we can refer to it as the no-motion manifold, in accordance with the terminology introduced by Laboissière et al. (1996).

Let $\boldsymbol{\lambda}_C = \Lambda_n \mathbf{u}_n$, for an arbitrary vector $\mathbf{u}_n$, be a vector of the no-motion manifold. The question now is how to determine appropriately the components of the vector $\mathbf{u}_n$. One of the requirements is that the components of $\boldsymbol{\lambda}_C$ should be negative, such that when multiplying this vector by a positive constant, the resulting effect will be to increase the activation [A, see (1)] of each individual muscle, and therefore to increase the force exerted by each muscle.

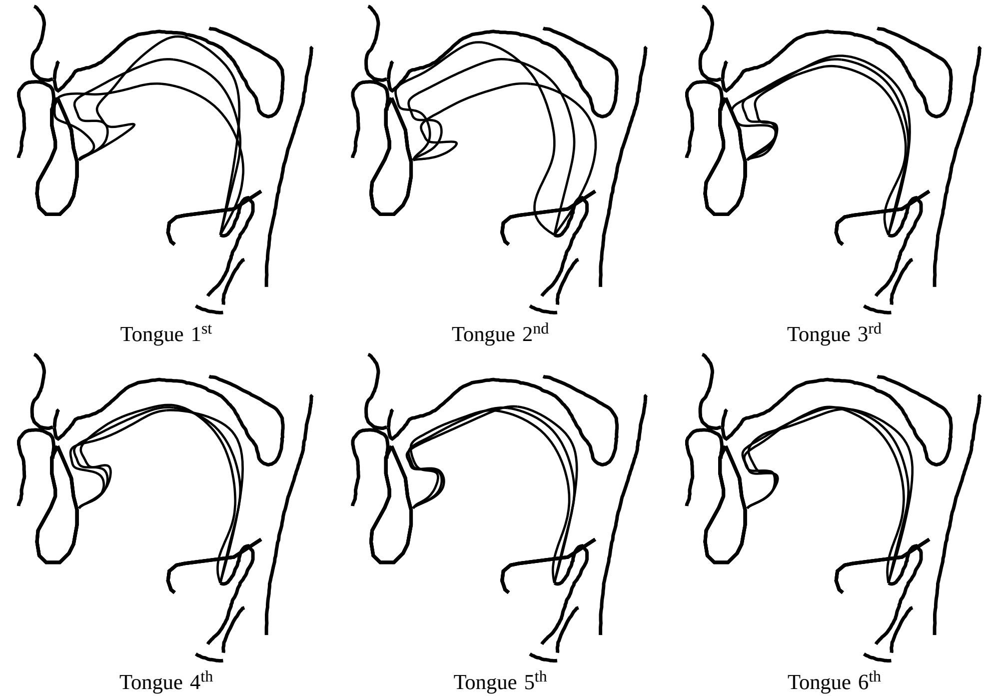


Fig. 7. Effect of the first six tongue components. Each panel displays a sequence of tongue contours corresponding, for each component of **t**, to ±3 times the standard deviation. Lips and epiglottis do not move and are only displayed for reference

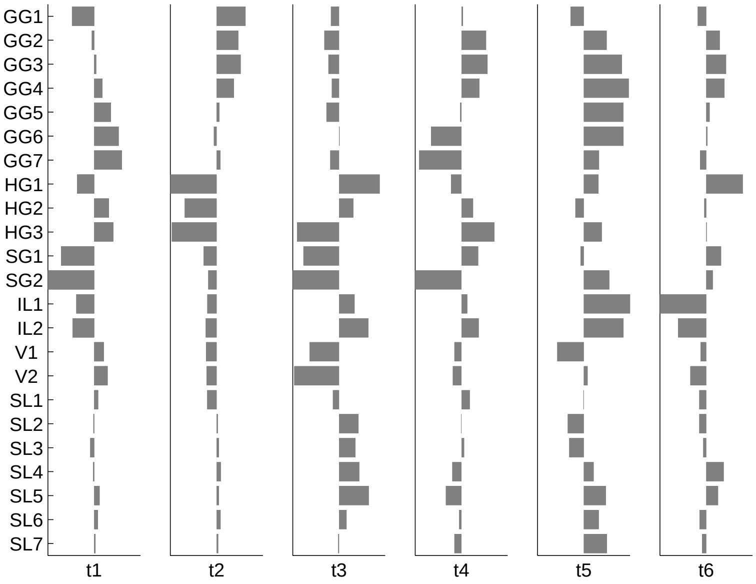


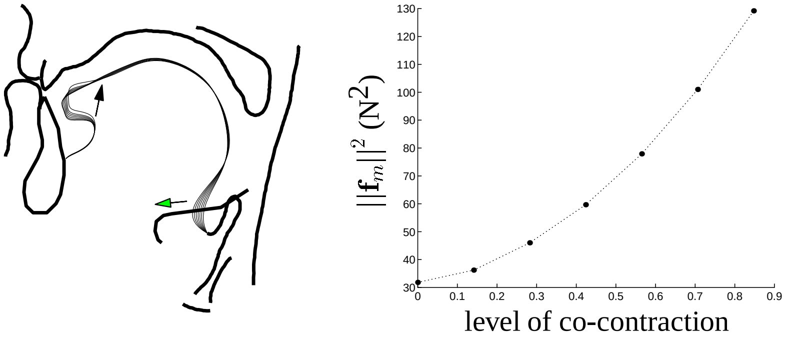
Fig. 8. Agonist and antagonist muscles for each of the tongue components. For each component t_i, $i = 1, \ldots, 6$, the bar graph displays the value of $\boldsymbol{\lambda}_{ti}$, normalized with respect to the maximum; *negative* (i.e. *left*) and *positive* (i.e. *right*) *bars* indicate, respectively, agonists and antagonists

With the additional constraint that the norm of $\boldsymbol{\lambda}_C$ should be equal to 1, and applying an optimization procedure favouring the principal directions associated with a smaller variance, it was possible to find an appropriate value for $\mathbf{u}_n$, and hence for $\boldsymbol{\lambda}_C$.

Figure 4.2.3 shows the effect of scaling the co-contraction command computed by the above procedure for the tongue posture corresponding to $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$ (a similar effect is observed in different postures). As expected, changes in the co-contraction level have almost no effect on tongue posture, whereas the norm of the vector of muscle forces increases monotonically. In conclusion, the geometrical arrangement of tongue muscles allows, in fact the separate control of position and stiffness.

5 Discussion

A biomechanical model of the tongue has been used to analyse tongue movements in speech, in terms of their control at the muscle level. In spite of the observed large variability of tongue shapes and the highly complicated arrangement of tongue musculature, it has been shown that tongue movements can be described in terms of a relatively small number of factors, or control variables.



(a)
(b)

Fig. 9. Effect of co-contraction command on tongue shape (**a**) and on the square of the total level of muscular force **b**. *Arrows* indicate the direction of deformation while increasing the level of co-contraction

Such a factor analysis has a precise biomechanical interpretation. In fact, a PCA in muscle space has the effect of decomposing the space of muscle commands into a set of orthogonal components, that correspond to the directions of maximum variance. This identifies a set of "maximally independent" muscle groups and, indirectly, a set of elementary or "primitive" movements that are specified by the directions of action of the muscles involved in each group. The latter also define a set of "functional" degrees of freedom, which depend on the structure of the system as it can be inferred from the data set. Moreover, as a particular muscle may contribute to movements in different directions, it may be eventually shared by different muscle groups.

As a consequence, provided that the data set is representative of the multiplicity of configurations assumed by the tongue during speech movements, the above procedure defines a somewhat "optimal" (in the sense of maximal independence) set of articulators that, different from those defined in terms of the geometry of the vocal tract (Maeda 1988), may be directly related to tongue musculature.

In fact, because of the requirement of minimum co-contraction in the fitting procedure, the statistics of $\boldsymbol{\lambda}$ is dominated by that of the vector of muscle lengths, $\mathbf{l}$ (i.e. a geometric quantity). Therefore, the resulting set of articulators should be basically regarded as a consequence of the geometrical arrangement of muscle fibres in the tongue. However, the elementary gestures associated with the control variables are not an artifact of the particular model implementation, as demonstrated by the qualitatively similar results that have been derived from preliminary versions of the model (Sanguineti 1995, Laboissière et al. 1995).

Figure 8 suggests that the predicted structure of the muscle synergies associated with each articulator may be somehow more complicated than previously argued (Perkell 1969). Although intrinsic muscles only provide a small contribution to tongue positioning (t_1 and t_2), the inverse is certainly not true for extrinsic muscles in $t_{3\text{–}6}$; therefore, the dichotomy between extrinsic muscles (affecting tongue position) and the intrinsic ones (affecting tongue shape) may be somewhat simplistic. Indeed, in accordance with Öhman (1967), the present results suggest that individual muscles may participate in the control of several articulators – virtually, all; see also Smith (1992).

The whole study is based on the assumption that the bony parts behave as a moving frame of reference, whose position is not affected by the activation of tongue muscles. However, the dependence of $\boldsymbol{\lambda}$ on the positions of the jaw and the hyoid bone (23.3% of total variance) clearly suggests that the tongue and the bony parts are tightly coupled from the mechanical point of view; moreover, the activation of tongue muscles may indeed affect the equilibrium at least of the hyoid bone; see also Honda (1994). Therefore, from the mechanical point of view, the jaw-hyoid-tongue system should be treated as a whole, and the identification of "global" articulators and muscle synergies, also accounting for the anatomy and musculature of the jaw and the hyoid bone, is worth exploring.

From the control point of view, the above results suggest that the movements of the tongue may be accounted for by a simple model, in which the muscle commands are obtained by central commands through a linear relationship:

$$\begin{aligned}\boldsymbol{\lambda} &= \boldsymbol{\lambda}_0 + \boldsymbol{\lambda}_{b1}b_1 + \boldsymbol{\lambda}_{b2}b_2 + \boldsymbol{\lambda}_{b3}b_3 + \boldsymbol{\lambda}_{t1}t_1 \\ &\quad + \boldsymbol{\lambda}_{t2}t_2 + \boldsymbol{\lambda}_{t3}t_3 + \boldsymbol{\lambda}_c c \qquad (7)\end{aligned}$$

where b_1, b_2 and b_3 specify the position of the jaw and the hyoid bone; t_1, t_2 and t_3 specify a tongue posture; and the scalar $c > 0$ accounts for the level of co-contraction.

This is consistent with the suggestion (Lacquaniti 1989) that the CNS does not need to maintain a detailed internal representation of body geometry but rather may make use of simple approximations. For instance, Laboissière et al. (1996) have found that the movements of the jaw-hyoid system may be controlled by a few commands or muscle synergies (corresponding to pure translations or rotations) that are approximately invariant over the whole workspace.

The present tongue model (plus the above control scheme) may be considered as a step toward a new generation of models of the vocal tract for speech production where, different from geometric or statistical models, the control variables or articulators are defined at the level of muscles.

Acknowledgements. This research was supported by the European Community ESPRIT - BR grant # 6975 SPEECH-MAPS, by a grant from the Communauté de Travail des Alpes Occidentales and by the France-Italy cooperation project GALILEO. We are indebted to D. Ostry, whose comments and suggestions helped to improve the manuscript considerably.

References

Alfonso PE, Honda K, Baer T, Harris KS (1982) Multichannel study of tongue EMG during vowel production. J Acoust Soc Am 71: S54(A)

Baer T, Alfonso PJ, Honda K (1988) Electromyography of the tongue muscles during vowels in / əpVp/ environment. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 22: 7–19

Bothorel A, Simon P, Wioland F, Zerling JP (1986) Cinéradiographie des voyelles et consonnes du français. Technical report, Travaux de l'Institut de Phonétique de Strasbourg, France

Bowman, JP (1971) The muscle spindle and neural control of the tongue: implications for speech. Charles C Thomas, Springfield

Bowman JP, Combs CM (1968) Discharge patterns of lingual spindle afferent fibers in the hypoglossal nerve of the Rhesus monkey. Exp Neurol 21: 105–119

Chiel HL, Crago P, Mansour JM, Hathi K (1992) Biomechanics of a muscular hydrostat: a model of lapping by a reptilian tongue. Biol Cybern 67: 403-415

Cooper S (1953) Muscle spindles in the intrinsic muscles of the human tongue. J Physiol 122: 193–202

Dickson DR, Maue-Dickson W (1982) Anatomical and physiological bases of speech. Little, Brown, Boston

Duck FA (1990) Physical properties of tissue: a comprehensive reference book. Academic Press, London

Feldman AG (1966) Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II. Controllable parameters of the muscle. Biophysics 11: 565–578

Feldman AG, Levin MF (1995) The origin and use of positional frames of reference in motor control. Behav Brain Sci 18: 723–744

Feldman AG, Orlovsky GN (1972) The influence of different descending systems on the tonic reflex in the cat. Exp Neurol 37: 481–494

Feldman AG, Adamovich SV, Ostry JD, Flanagan JR (1990) The origins of electromyograms – explanations based on the equilibrium point hypothesis. In: Winters JM, Woo SL-Y (Eds) Multiple muscle systems – biomechanics and movement organization. Springer-Verlag, Berlin Heidelberg New York

Fitzgerald MJT, Sachithanandan SR (1979) The structure and source of lingual proprioceptors in the monkey. J Anat 128: 523–552

Fujimura O, Kakita Y (1978) Remarks on quantitative description of the lingual articulation. In: Lindblom B, Ohman S (Eds) Frontiers of speech communication research. Academic Press, London, pp 17–24

Gambarelli J, Guerinel G, Chevrot L, Mattei M (1977) Coupes sériées du corps humain: anatomie, radiologie, scanner. Masson, Paris

Gerhardt P, Frommhold W (1988) Atlas de corrélations anatomiques en tomodensitometrie et imagerie par résonnance magnétique [Atlas of anatomic correlations in CT and MRI]. Flammarion Medecine-Sciences, Paris

Harshman R, Ladefoged P, Goldstein L (1977) Factor analysis of tongue shapes. J Acoust Soc America 62: 693–707

Hashimoto K, Suga S (1986) Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue. J Acoust Soc Japan (E)7: 39–46

Honda K, Hirai H, 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

Humphrey DR, Reed DJ (1983) Separate cortical systems for control of joint movement and joint stiffness: reciprocal activation and coactivation of antagonist muscles. In: Desmedt JE (Ed) Motor control mechanisms in health and disease. Raven Press, New York, pp 347–372

Kakita Y, Fujimura O (1977) Computational model of the tongue: a revised version. J Acoust Soc Am 62: S15(A)

Kakita Y, Fujimura O, Honda K (1985) Computation of mapping from muscular contraction patterns to formants patterns in vowel space. In: Fromkin VA (Ed) Phonetic linguistics. Academic Press, London, pp 133–144

Kiritani S, Miyawaki K, Fujimura O, Miller JE (1976) A computational model of the tongue. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 10: 243–251

Laboissière R, Sanguineti V, Payan Y (1995) On the biomechanical control variables of tongue during speech movements. In: Proceedings of the 4th European Conference on Speech Communication and Technology, Madrid, Spain, pp 1289–1292

Laboissière R, Ostry DJ, Feldman AG (1996) Control of multi-muscle systems: human jaw and hyoid movements. Biol Cybern 74: 373–384

Lacquaniti F (1989) Central representations of human limb movement as revealed by studies of drawing and handwriting. Trends Neurosci 12: 287–291

Lowe AA (1978a) Excitatory and inhibitory inputs to hypoglossal motoneurons and adjacent reticular formation neurons in cats. Exp Neurol 62: 30–47

Lowe AA (1978) Mandibular joint control of genioglossus muscle activity in the cat and monkey (*Macaca irus*). Arch Oral Biol 23: 787–793

Maeda S (1988) Improved articulatory model. J Acoust Soc Am 81: S146

Maeda S, Honda K (1994) From EMG to formant patterns of vowels: the implication of vowel spaces. Phonetica 51: 17–29

Mermelstein P (1973) Articulatory model for the study of speech production. J Acoust Soc Am 53: 1070–1082

Miyawaki K (1974) A study on the musculature of the human tongue. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 8: 23–50

Miyawaki K, Hirose H, Ushijima T, Sawashima M (1975) A preliminary report on the electromyographic study of the activity of lingual muscles. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo 9: 91–106

Müller E, Milencovic P, MacLeod G (1984) Perioral tissue mechanics during speech production. In: Eisenfeld J, DeLisi C (Eds) Mathematics and computers in biomedical applications. Elsevier-North Holland, Amsterdam, pp 363–371

Neilson PD, Andrews G, Guitar BE, Quinn PTG (1979) Tonic stretch reflexes in lip, tongue and jaw muscles. Brain Res 178: 311–327

Öhman SEG (1966) Coarticulation in VCV utterances: spectrographic measurements. J Acoust Soc Am 39: 151–168

Öhman SEG (1967) Numerical model of coarticulation. J Acoust Soc Am 41: 310-320

Ostry DJ, Gribble PL, Levin MF, Feldman AG (1997) Phasic and tonic stretch reflexes in muscles with few muscle spindles: human jaw opener muscles. Exp Brain Res (in press)

Payan Y, Perrier P, Laboissière R (1995) Simulation of tongue shape variations in the sagittal plane based on a control by the equilibrium point Hypothesis. In: Proceedings of the XIII International Conference on Phonetic Sciences. Vol 2, Stockholm, Sweden, pp 474–477

Perkell J (1969) Physiology of speech production. MIT Press, Cambridge, Mass.

Perkell J (1974) A physiologically-oriented model of tongue activity in speech production. Unpublished doctoral dissertation, Department of Electrical Engineering, MIT

Perrier P, Loevenbruck H, Payan Y (1996) Control of tongue movements in speech: the equilibrium point hypothesis perspective. J Phonetics 24: 53–75

Sanguineti V (1995) On the control variables of the jaw-tongue system in speech. In: Proceedings of the XV Congress of the International Society of Biomechanics, Jyväskylä, Finland

Schwarz HR (1984) Finite element methods. Academic Press, London

Smith A (1992) The control of orofacial movements in speech. Crit Rev Oral Biol Med 3: 233–267

Smith K, Kier WM (1989) Trunks, tongues, and tentacles: moving with skeletons of muscle. Am Sci 77: 29–35

Tolu E, Caria MA, Simula ME, Lacana P (1994) Muscle spindle and periodontal trigeminal afferents modulate the hypoglossal motoneuronal activity. Arch Ital Biol 132: 93–104

Walker LB, Rajagopal MD (1959) Neuromuscular spindles in the human tongue. Anat Rec 133: 438

Wilhelms-Tricarico R (1995) Physiological modeling of speech production: methods for modelling of soft-tissue articulators. J Acoust Soc Am 97: 3085–3098