J. Acoust. Soc. Jpn. (*E*) 7, 1 (1986)

Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue

Kiyoshi Hashimoto* and Shinsuke Suga**

**Research Institute for Communication Sciences, University of Electro-Communications, 1–5–1, Chofugaoka, Chofu, 182 Japan*
***Doctoral Program in Engineering, University of Tsukuba, Sakura-mura, Niihari, Ibaraki, 305 Japan*

(*Received 7 January 1985*)

This paper is concerned with application of a three-dimensional model of the tongue to the problem of estimating a set of muscular tensions of the human tongue from the given X-ray outline. In order to inspect the motional aspects of the muscular tensions and also in order to obtain the more accurate estimations than the previous single-stage estimation, the present experiment adopts the multistage fitting. From a given tongue outline, the consecutively varied outlines from the neutral shape are produced by interpolation, and a set of muscular tensions of the model tongue is obtained each time the model is fitted best to the interpolated outline. In comparisons with the previous single-stage estimations, the present results are nearer to the actual human EMG data. The present multistage results also reveal the nonlinear time patterns for the muscular tensions, which are presumed due to their participation in pressing the tongue upward to the palate or in pressing it down to the floor.

PACS number: 43. 70. B

1. INTRODUCTION

At least two eminent studies are hitherto known on the physiologically oriented models of the human tongue.[1,2] However, on the thesis that a tongue model should reflect as much truth of articulatory behaviors of the tongue as possible, we have already simulated by a computer a three-dimensional (henceforth 3-D) precise model of the tongue and vocal tract. Since our computer program inherently enables to produce a model shape from a given set of muscular tension, we expanded the program so that the execution of it may find an appropriate set of the muscular tensions which fits the shape of the model best to the given X-ray outline.[3-5]

However, such previous procedure is that of fitting it at a time to a final target shape of the human tongue. This may be called a "single-stage" fitting. On the contrary, the present procedure is "multistage" fitting. In this multistage procedure, the model is controled to fit to each of the several outlines which were produced by linear interpolation between the neutral and the final shapes. At each stage, the initial state of the model is retained from the preceding stage. If the relation between the tension of the particular muscle and a whole shape is linear, then the estimated sequence of the muscular tension may be constant. However, if the relation is not linear due to the interference with other muscles or due to the reaction from the wall, then the estimation may not produce a constant sequence of the muscular tension. In short, the multistage method is expected to reveal such "structural nonlinearity," caused by the reactions to the tongue from the boundaries, or caused by the mutual, or muscle-to-muscle, interactions within the tongue.

2. MODEL OF THE TONGUE AND PALATE

2.1 Finite-element Model of the Tongue

In our process of designing the finite-element tongue model, we referred to an anatomical data for the distributions of the lingual and related muscles.[6,7] The muscles contained in the model are the transversus (T), the verticalis (V), the superior longitudinalis (SL), the inferior longitudinalis (IL), the genioglossus (GG), the hyoglossus (HG), the styloglossus (SG). Since it was reported that the several subparts of the GG appears to work independently,[6] we divided the muscle roughly into six subparts, namely, GG1∼GG6. Although it is not a lingual muscle, we included the mylohyoid (MH) into the model on account of its serious influence on the articulatory behaviors of the tongue. After all, totally thirteen independent muscles are included into the model. Since the tongue body is assumed to be symmetric with respect to the midsagittal plane, it is enough to model only a half of the tongue. Figures 1 (a) and 1 (b) show a perspective and a midsagittal views of the tongue model, respectively. It is composed of 492 elements and 170 nodes.

In order to model the palate and other structures surrounding the tongue, we used several patches of

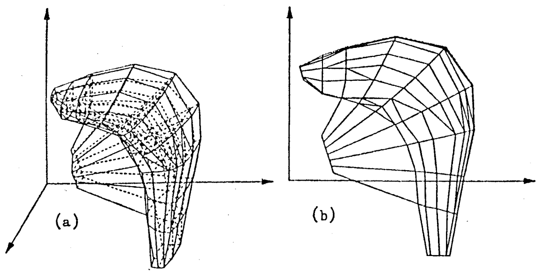

Fig. 1 The neutral shape of the tongue model. (a) perspective view, (b) side view.

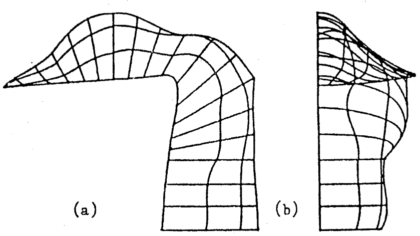

Fig. 2 The P.C. patches for the palate and other structures surrounding the tongue. (a) side view, (b) front view.

curved surfaces expressed by "parametric cubic" (henceforce, P.C.) technique.[8] For instance, the X-coordinates on a P.C. patch is represented by a couple of parameters, u and w, as,

$$X(u, w)=(u^3\ u^2\ u\ 1)S_x(w^3\ w^2\ w\ 1)^{\mathrm{T}}, \qquad (1)$$

where $0\leqq u,\ w\leqq 1$ holds and S_x is 4×4 matrix of constants. Figures 2 (a) and 2 (b) show the side and front views of the palate and other structures surrounding the tongue.

For simplification, we assume that the palate is rigid and that no friction exists between the tongue and palate. Therefore, when the tongue presses the palate, we assume that the tongue receives only a reactive force normal to the surface of palate. Also, we include the mandible as the lower structures on which our tongue model is built. Note that, in this model, the palate is immovable, while the tongue is on the movable coordinates tied to the mandible.

According to the finite element method,[9]

$$\{\delta\}^e=[B]^{\mathrm{T}}[D][B]V\{\delta\}^e=[k]^e\{\delta\}^e \qquad (2)$$

holds for any single tetrahedral element, under the assumption of infinitesimal strain, where $\{F\}^e$ and $\{\delta\}^e$ are the vectors expressing the nodal forces and nodal displacements, respectively, and V is the volume of the tetrahedral element e. The strain-displacement matrix $[B]$ in Eq. (2) is composed of constants, since this is a linear model. The strain $\{\varepsilon\}^e$ for the element e, is expressed in terms of $[B]$ by the equation

$$\{\varepsilon\}^e=[B]\{\delta\}^e\,. \qquad (3)$$

The elasticity matrix $[D]$ for an isotropic material is given by

$$[D]=\frac{E(1-\nu)}{(1+\nu)(1-2\nu)}\cdot\begin{bmatrix} 1 & \frac{\nu}{(1-\nu)} & \frac{\nu}{(1-\nu)} & 0 & 0 & 0 \\ & 1 & \frac{\nu}{(1-\nu)} & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & \frac{1-2\nu}{2(1-\nu)} & 0 & 0 \\ & \text{Sym.} & & & \frac{1-2\nu}{2(1-\nu)} & 0 \\ & & & & & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix}, \qquad (4)$$

where E and ν are the Young's modulus and the Poisson's ratio, respectively. Unfortunately, we could not find any rheological data for the tongue. However, referring to the previously published data on the relation between stress and strain of the contracted muscle,[10] we estimated the Young's modulus for the contracted muscle. The Young's modulus was, though obscured due to the slight nonlinearity, approximately 3.0 kg/cm². Since the Young's modulus for the contracted muscle is thought as being larger than that for the muscleless soft tissues, the Young's modulus for the tongue body is assumed to depend much on the density of the coexisting muscle fibers, or simply on the number of muscles, when all the muscles are equally contracted. Thus, we assigned constant or time-invariant values of Young's moduli to the specified parts of the tongue body, depending on the number of coexisting muscles: 1.0 kg/cm² to the muscleless soft tissues, 1.5 kg/cm² to the part composed of a single muscle, 2.0 kg/cm² to the part composed of two or three muscles, and 4.0 kg/cm² to the part composed of more than three muscles. Furthermore, the maps of the lingual muscles in a microscopic study of the tongue[7] was a good reference on which we could decide the number of coexisting muscles in our tongue model. We also gave the value of 0.49 to the Poisson's ratio in consideration of the nature of the tongue body, that is, hard to compress but slightly plastic. The results of our previous study did not reveal any extraordinary inadequacy due to our assumption of "isotropy" for the model. The equilibrium equation, Eq. (2), can be expanded finally into a set of linear equations

$$\{F\}=[K]\{\delta\}, \tag{5}$$

where $[K]$ is stiffness matrix of the model. By giving the actual values to the unified vector $\{F\}$ for all the nodal forces in Eq. (5), we can solve Eq. (5) with respect to the unknown column vector $\{\delta\}$ which expresses the displacements of all the nodes.

2.2 Model of the Muscular Tensions

The method to compute the vector $\{F\}$ in Eq. (5) is given as follows. First, consider a model of an entire muscle composed of n sections as is shown in Fig. 3. Let $\boldsymbol{a}_i$ be the 3-D vector for the direction of the i-th section of the muscle fiber, and $\boldsymbol{f}_i$ be the force acting on the boundary plane P_i. Furthermore, let α_i $(0\leqq\alpha_i\leqq1)$ be introduced as the specific coefficients that represent the increase or decrease of numbers between the adjacent fibers, and let F_i be the force exerted on the i-th section. Then, F_i is expressed by

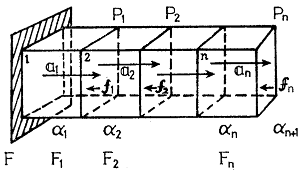

Fig. 3 Schematic illustration of a whole muscle. F_i and 3-D vector $\boldsymbol{a}_i$ indicate the internal tension and the direction of the i-th section of the muscle fiber, respectively. The subscript i runs from 1 to n.

$$F_i=F\prod_{j=1}^{i}\alpha_j, \tag{6}$$

where F is the input tension for each muscle, which will henceforth be simply referred to as the muscular tension. The boundary force $\boldsymbol{f}_i$ is given by the difference between the adjacent muscle forces as

$$\boldsymbol{f}_i=F\left\{\boldsymbol{a}_{i+1}\prod_{j=1}^{i+1}\alpha_j-\boldsymbol{a}_i\prod_{j=1}^{i}\alpha_j\right\}\quad(i<n) \tag{7}$$

and

$$\boldsymbol{f}_n=-F\boldsymbol{a}_n\prod_{j=1}^{n}\alpha_j. \tag{8}$$

Thus, if a whole set of muscular tensions are given, the nodal forces of each tetrahedron are computed for all the constituent tetrahedra, and then a whole shape of the model is obtained.

3. FITTING THE MODEL SHAPE TO THE X-RAY OUTLINE

3.1 Method of Successive Search for the Best Fit

The next step is to fit a shape of the model to the X-ray outline of the real human tongue. This requires a series of executions of computerized, systematic and successive search for an appropriate set of muscular tensions. Prior to this, however, we have to determine the criterion for the goodness of fit. Figure 4 shows the coordinate system for evaluation of the goodness of fit, on which the seven arrows represent an outline of the tongue. Let r_i and r_i' be the lengths of the i-th arrows for the real tongue, and for the model tongue, respectively. Then, the best set of the muscular tensions minimizes the sum of

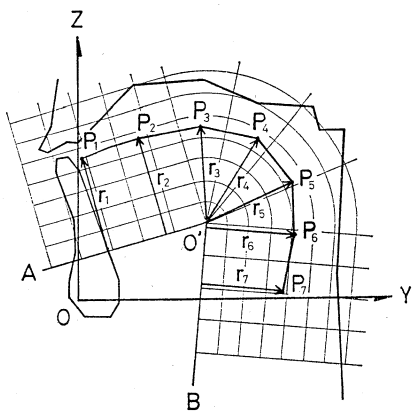

Fig. 4 The coordinate system for evaluation of the degree of fitness of the model shape.

squares of the differences $d_i = r_i' - r_i$. In other words, the best set of the muscular tensions is given as the solution of a minimization problem,

$$D = \min \sum_{i=1}^{7} d_i^2 \,. \tag{9}$$

In order to express the goodness of fit, we adopted, by way of analogy to the familier S/N ratio, a quantity,

$$S/D = 10 \log_{10} \left\{ \sum_{i=1}^{7} r_i^2 \Big/ \sum_{i=1}^{7} d_i^2 \right\} \tag{10}$$

in dB, where S is the squared sum of the lengths of seven arrows, i.e., $\sum_{i=1}^{7} r_i^2$. Henceforth, we refer to this quantity as S/D ratio.

As the minimization algorithm, we adopted the variable metric method.[11] Moreover, in order to accelerate the minimization, we omitted a couple of less important variables, that is, the transversus (T) and the verticalis (V), out of the control variables. This omitting is mainly based on the previous study by Miyawaki *et al.*[13] in which the EMG activity of the transverse-vertical muscle complex is not recognized in vowel-like gestures. Furthermore, since the six subparts of the GG are closely related, we represented it by only the two most dominant principal components, PC1 and PC2. Consequently, we used seven variables for minimization.

The outlines for the actual tongue were obtained originally from X-ray tracings.[12] Out of them, we chose the sustained portions of vowels, /i/,/ɛ/,/a/, /U/,/u/, in utterances /hə'tV/. Regarding each sustained portion as a target shape starting from the neutral shape, we produced a set of five consecutively varying outlines by means of the computerized linear interpolation on the coordinate system shown in Fig. 4. It should be emphasized that each of the interpolated outlines of the tongue provides the target shape in each "stage" in multistage fitting. The linear interpolation, which we adopted in this study, represents the constant velocity of the tongue. In spite of such simplicity of the assumption on the motion of the tongue, we conducted the current multistage experiment, because it is rather the qualitative features of the muscle activities than the exact quantification of the muscle in motion that we intend to know. The muscular tensions are computed with the preceding result retained as the renewed initial condition in the succeeding stage.

3.2 Goodness of Fit of the Model Shape to the Given Outlines

Figure 5 shows the shapes of the model in their final stages, where the given outlines are represented by the dotted lines.

Table 1 shows the S/D ratios as the goodness of fit in each stage. The table also shows the S/D ratios obtained by the previous single-stage fitting. The S/D ratios for the multistage fitting decline as the stage advances. However, the ratios even at the final stages are better than those in the single-stage fitting. The degradation of the S/D ratios at the later stages is remarkable when the model shapes are distorted due to, for instance, pressing of the lateral part of the tongue to the palate.

3.3 Aspects of the Muscular Activities as Functions of Time

In order to simplify the experiment, we neglected the mass of tongue and attempted to express its movement by the succession of the static equilibria of muscular tensions. Furthermore, we assumed a constant velocity of the movement. In such a simple situation, it is highly expected that the tension of any muscle is kept almost constant with respect to time. However, the results we obtained for the tensions are far from being constant, as is shown in Fig. 6. It should be noted that the SL is omitted in Fig. 6, because its tension is found almost zero in the results of the present experiment.

One of the furthest examples from being constant

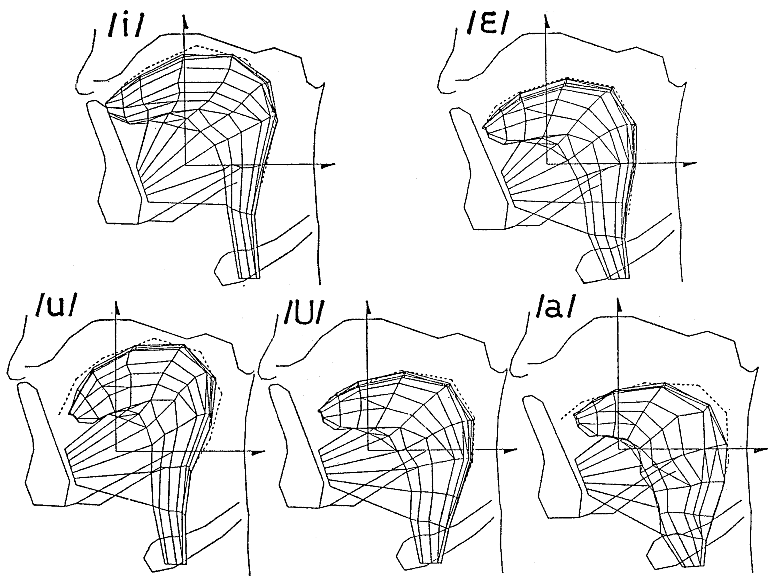

Fig. 5 The tongue shapes best fitted to the given X-ray contour.

Table 1 The S/D ratios in dB expressing the goodness of fit of the model to the outlines of the human tongue, in comparison between the previous single-stage and the present multistage experiments.

	Single-stage fitting (dB)	Multistage fitting (dB)				
		1	2	3	4	5
/i/	24.4	33.5	27.4	29.0	27.0	24.9
/ɛ/	32.1	42.9	36.2	34.3	31.5	29.6
/a/	23.8	42.4	32.3	24.0	29.1	26.6
/U/	24.7	39.5	37.0	34.0	30.6	27.5
/u/	23.3	31.5	35.0	29.1	28.9	24.9

is the result for /i/, where the tension for the frontal part of the GG gradually rises together in accordance with the rise of the MH. The result for /ɛ/ is marked by its relatively high constancy of the sequences of the tensions, except the tension for the SG that rises. The result for /a/ is another example of excessive irregularity. All the subparts of the GG have sudden rises of their tensions at stage 4, and also the HG at stages 3 and 4. On the contrary, /U/ has relatively constant sequences. The result for /u/ is also marked by some irregularities with the GG and the SG.

In summary, since the model assumes initially the neutral shape, if the target shape is almost neutral, the muscular tensions are small and constant. Both vowels /ɛ/ and /u/ are such cases, whereas the remainders have irregular time sequences of the tensions. Such irregularities are presumably evoked primarily by the interactions between the tongue and its surrounding structures such as the palate or the mandible.

3.4 Comparison of the Results with the EMG Data

A strict verification of the adequacy of this model may require a comparison between the results obtained by this model and the EMG activity of the

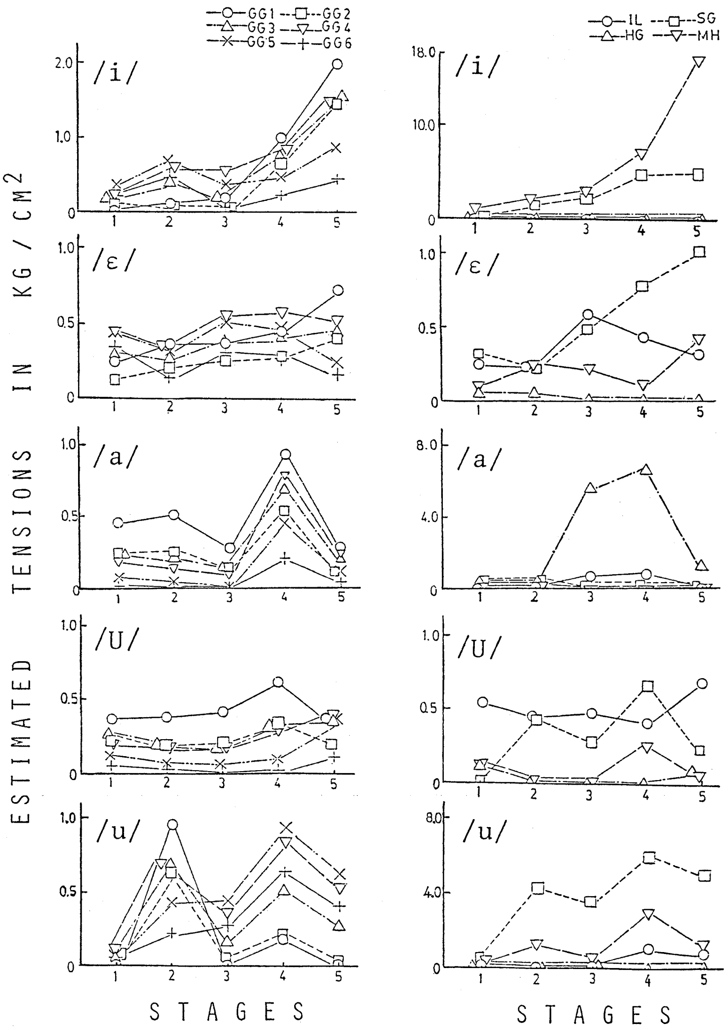

Fig. 6 The estimated muscular tensions for the tongue shapes of the model in stages 1 to 5.

same subject taken simultaneously with the X-ray pictures. However, even a comparison with the only EMG data available for us, which are those for the GG of a male Japanese for sustained vowels /i/, /e/, /a/, /o/, and /u/,[13] may give some overview on the adequacy of the model. Figure 7 shows (a) the EMG activities (top), (b) the previous results obtained by the single stage of computation (middle), and (c) the single-stage equivalences of the multistage muscular tensions obtained by adding each of the five multistage tensions (bottom). The muscular tensions and the EMG activities are normalized by the maximum values and expressed in percentages. A couple of distances between (a) and (b), and between (a) and (c) can be expressed by rms of their differences. They were found to be 37.8% and 31.8%, respectively. However, if the EMG data is linearly compressed upward by 40% around the upmost vertical line of 100%, then these distances decrease to 28.0% and 14.8%, respectively. This means an overwhelmingly smaller distance for the latter. A glance of Fig. 7 indicates that the decrease of the distance for the latter is mainly due to the remarkable decrease of the differences between the muscular tensions in the multistage fitting and the actual EMG activities with respect to the GG1 and GG2. Figure 8 shows the

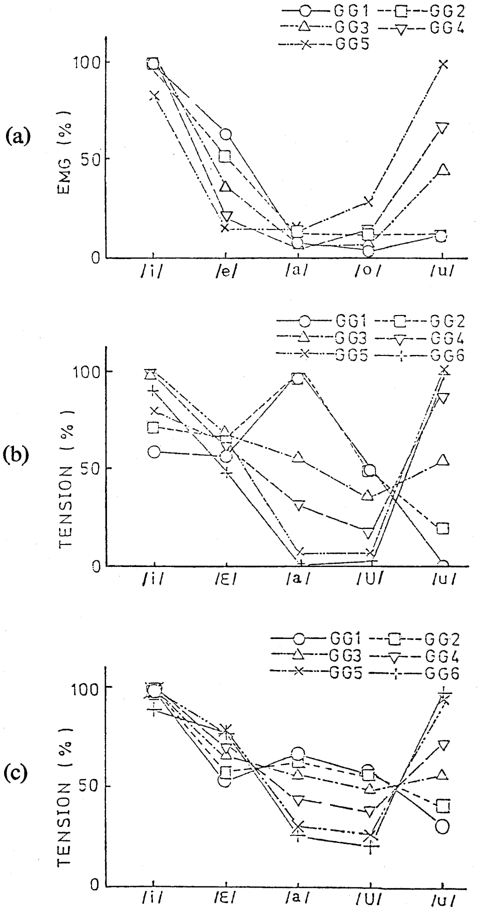

Fig. 7 The relative values for the muscular activities and tensions. (a) EMG (top), (b) previous tensions (middle) and (c) present tensions (bottom).

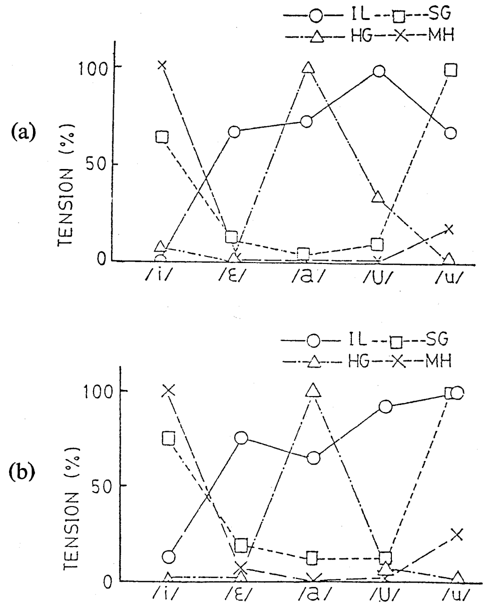

Fig. 8 The relative values for the muscular tensions. (a) previous tensions and (b) present tensions.

tensions of the other muscles than the GG, (a) the previous result (upper) and (b) the present result (lower). However, no significant difference is found between them.

4. DISCUSSION

It is obvious from the above results that the multistage fitting gives better estimates on the muscular tensions than the previous single-stage fitting, where we mean by "better" that it is nearer to the human EMG. The other merit of using the multistage fitting is that it can reveal the motional aspects of the tension of each muscle. And it can be said that, if a sequence of the tension of a certain muscle shows a sudden and remarkable rise, then the muscle must have participated in pressing the tongue to the boundaries. Thus, it can be said from the results, that the MH and the SG participate in pressing of the tongue to the palate, and that the HG and the GG participate in pressing of the tongue to the floor of the mouth.

For future refinements of this model, it will be required for us to rebuild its structures, to renew its parameters, and to consider factors such as the friction between the tongue and its boundaries. However, if we continue to refine this model and continue to develop the estimation technique of the muscular tensions, finally we may reach to one of our ultimate goals, that is, the simulation and the analysis of the disordered articulatory dynamics in patients suffering from apraxia, and related conditions.

ACKNOWLEDGEMENTS

The authors thank Prof. Ishizaka for his adequate comment on the present research. Computations for this experiments were executed on the computer system in the Large Computer Center at Univ. of Tokyo.

REFERENCES

1) J. S. Perkell, "A physiologically-oriented model of tongue activity in speech production," Doctoral Thesis at MIT (1974).
2) S. Kiritani, K. Miyawaki, O. Fujimura, and J. E. Miller, "A computational model of the tongue," Ann. Bull., Res. Inst. Logop. Phoniatr., Univ. Tokyo, **10**, 243–251 (1976).
3) K. Hashimoto and T. Oosakaya, "Muscle-oriented modeling of the tongue," Trans. Comm. Speech Res., Acoust. Soc. Jpn. S77-78 (1978) (in Japanese).
4) K. Hashimoto and Y. Matsushita, "Finite-elements in a three-dimensional model of the tongue," Trans. Comm. Speech Res., Acoust. Soc. Jpn. S88-94 (1981) (in Japanese).
5) K. Hashimoto, Y. Matsushita, and S. Suga, "Simulation of the articulatory behavior of the tongue," in *Applications of Computers in Medicine*, M. D. Schwartz, Ed. (IEEE Press, New York, 1982), pp. 42–48.
6) M. Tsuzaki, "Study on the anatomy of facial muscles and the muscles constructing the tract in the head," J. Otorhinolaryngol. Soc. Jpn. **8**, 1880–1889, (1962) (in Japanese).
7) K. Miyawaki, "A study on the musculature of the human tongue," Ann. Bull., Res. Inst. Logop. Phoniatr., Univ. Tokyo, **8**, 23–50 (1974).
8) G. J. Peters, "Interactive computer graphics application of bi-cubic parametric surface to engineering design problems," AFIPS Conf. Proc. **43**, 491–511 (1974).
9) O. C. Zienkiewicz, *The Finite Element Method* (Mc-Grow-Hill, London, 1977), Chap. 6, p. 96.
10) S. Oka, *Rheology-Biorheology* (Shokabo, Tokyo, 1974), pp. 454–456.
11) R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," Comput. J. **6**, 163–168 (1963).
12) J. S. Perkell, *Physiology of Speech Production: Results and Implication of a Quantitative Cineradiographic Study* (MIT Press, Cambridge, 1975), Appendix B, pp. 71–97.
13) K. Miyawaki, H. Hirose, T. Ushijima, and M. Sawashima, "A preliminary report on the electromyographic study of the activity of lingual muscles," Ann. Bull., Res. Inst. Logop. Phoniatr., Univ. Tokyo, **9**, 91–103 (1975).