Journal of Phonetics (2002) **30,** 261–279
doi:10.1006/jpho.2002.0173
Available online at http://www.idealibrary.com on IDEAL®

Compensatory articulation during bilabial fricative production by regulating muscle stiffness

Hiroaki Gomi* and Masaaki Honda

NTT Communication Science Laboratories, NTT Corporation, Japan and CREST, Japan Science and Technology Corporation, Japan

Takayuki Ito

CREST, Japan Science and Technology Corporation, Japan

Emi Z. Murano

CREST, Japan Science and Technology Corporation, Japan and The University of Tokyo, Japan

Received 9th November 2001, and accepted 3rd January 2002

The cooperative mechanisms in articulatory movements were examined by using mechanical perturbations during bilabial phonemic tasks. The first experiment compares the differences in compensatory responses during sustained productions of the bilabial fricative /Φ/ for which lip constriction is required, and /a/, for which the lips and jaw are relatively relaxed. In the second experiment, we perturbed jaw movement with different load-onsets in the sentence "kono /aΦaΦa/ mitai". In both experiments, labial distances were recovered partly or fully by the downward shifts of the upper lip. The upper lip response was frequently prior to the EMG response observed in the sustained task. Additionally, initial downward displacement of the upper lip was frequently larger when the load was supplied during /Φ/ than when it was supplied during /a/ in the sustained and sentence tasks, respectively. The stiffness variation estimated by using a muscle linkage model indicates that the stiffness increases for the bilabial phonemic task in order to robustly configure a labial constriction. The results suggest that the change in passive stiffness regulated by the muscle activation level is important in generating quick cooperative articulation.

*Address correspondence to H. Gomi, NTT Communication Science Laboratories, Wakamiya 3-1, Morinosato, Atsugi, Kanagawa 243-0198, Japan. E-mail: gomi@idea.brl.ntt.co.jp

0095–4470/02/$ - see front matter

1. Introduction

Although articulatory gestures vary depending on the speech context or constraining conditions, such as smoking a pipe or supporting the jaw with the arm, the articulatory organs cooperate to achieve the desired utterances. Cooperating articulatory movements can be seen not only under such planned or anticipated conditions, but also for unpredictable disturbances. When an unanticipated mechanical perturbation is inflicted on the lower lip during the bilabial explosive consonant /p/ or /b/, the closure between the upper and lower lips is accomplished by a downward shift of the upper lip (Abbs & Gracco, 1983, 1984; Gracco & Abbs, 1985). Similar compensatory movements were observed in bilabial phonemic tasks by applying an electrical perturbation to the lower lip (Folkins & Zimmermann, 1982) and in bilabial or linguo-dental phonemic tasks by applying a mechanical perturbation to the jaw (Folkins & Abbs, 1975; Kelso, Tuller, Bateson & Fowler, 1984; Shaiman, 1989).

These studies have also demonstrated functional changes in the coordination of compensatory movements for different kinds of unanticipated perturbations. For example, Gracco & Abbs (1985) studied the changes in the coordination of upper and lower lips for different onsets of perturbations: delayed onset of perturbation caused a large nonautogenic compensation of the upper lip because there was insufficient time for autogenic compensation of the lower lip. Rapid changes in the electromyogram (EMG) (Abbs & Gracco, 1984; Kelso *et al.*, 1984; Gracco & Abbs, 1985) and their task-dependencies imply a quick regulation mechanism of neural-linkage by suprabulbar pathways.

In spite of these remarkable findings, compensatory responses without any change in the EMG have been reported (Kelso *et al.*, 1984). Due to mechanical linkages between the lips, jaw, and tongue, as mentioned in Kelso *et al.* (1984), some passive dynamics (inertia of organs and/or muscle stiffness) could contribute to the cooperative (or task-conformable) behavior of these articulators. It has been suggested, on the other hand, that regulating the passive dynamics (mechanical impedance) is important when the arm is in interaction with manipulated objects or with external environments (Hogan, 1984; Mussa-Ivaldi Hogan & Bizzi, 1985), and that the mechanical impedance of the arm is governed by arm-kinematics and coordination of multiple-muscle activation (Mussa-Ivaldi *et al.*, 1985; Gomi, 1998; Gomi & Osu, 1998).

For articulatory movements, which involve mechanical interaction and coupling among lips, jaw, palate, and tongue, however, few attempts have so far been made at regulating the passive dynamics of articulators by changing muscle activity. Thus, the effect of the passive dynamics has not been well examined. In addition to the functional neural-linkages previously mentioned, we focus, in this study, on compensatory responses to a perturbation of the jaw for bilabial fricative consonants to clarify how the passive dynamics vary and how passive characteristics are exploited during coordinated articulatory movements.

We carried out two experiments. In the first, by using sustained productions of /Φ/ and /a/, we focused on the behavioral and EMG responses of the articulators associated with jaw perturbation to examine passive and active (or neuronally driven) components. In the second, we observed compensatory responses for jaw perturbations with different onsets during sentence productions. Additionally, by using these observed behaviors and a mechanical linkage model, the stiffness variation of the linkage between the upper lip and jaw were characterized. A part of this work has been presented elsewhere (Ito, Gomi & Honda, 2000*a*).

2. Method

2.1. *Experimental setup*

Fig. 1(a) shows the experimental setup. The subject sat on a chair with his shoulders fixed to the back of the chair by straps, and his head was strapped to a head-support device. The jaw of the subject was tightly held between a formed metal/plastic teeth-splint and a chin-plate, which were connected to the bar beneath the jaw by a piano wire. This mechanism enables us to perturb jaw motion in the jaw open/close directions by moving a bar connected to a torque motor (Shinmaywa DD-B09 with 1296k pulse/rev encoder) placed at the side of the head (roughly coinciding with jaw rotational center) (see Fig. 1(b)). Using a wire-driven

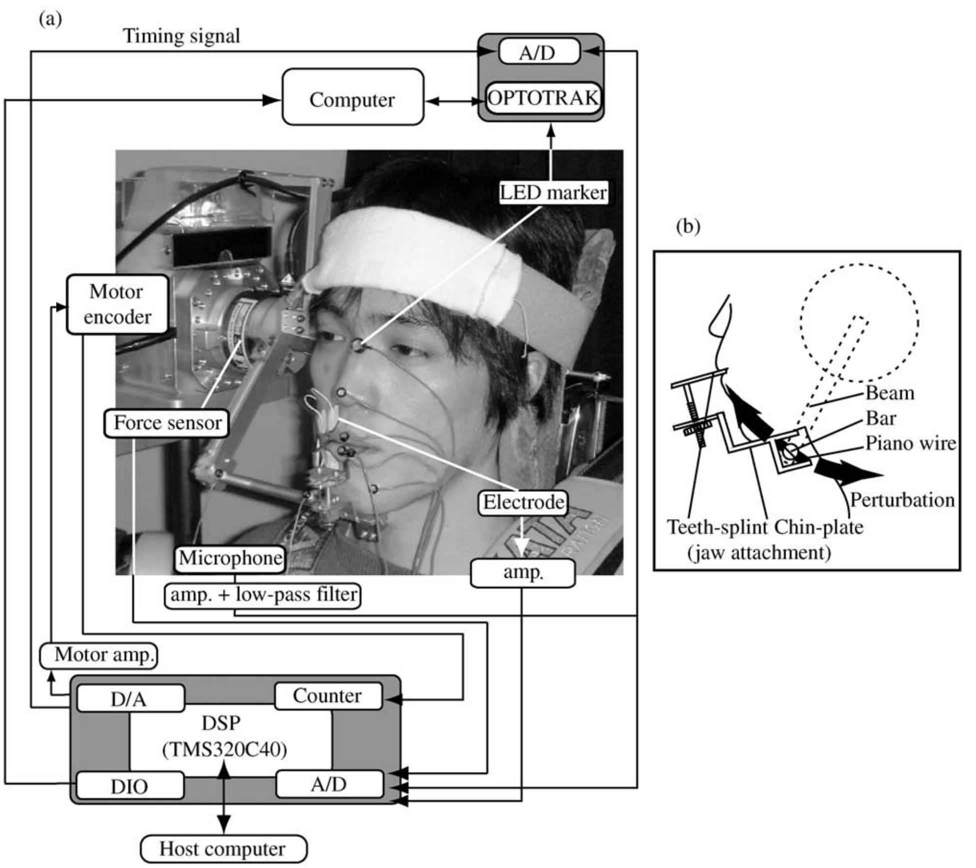

Figure 1. (a) Jaw perturbation system and experimental setup. (b) Side view of the configuration of experimental setup and the movement direction. A screw between a formed metal/plastic teeth-splint and a chin plate tightly fixes the jaw.

mechanism, we achieve low constraints in the roll, pitch, and yaw of the jaw with respect to the wire-connection point to the bar, and can easily fix the head to the headrest. Previous methods have delivered forces to the end of the jaw splint protruding from the mouth thus producing rotational forces that may dislodge the splint and/or strain the teeth (Folkins & Abbs, 1975; Kelso *et al.*, 1984; Shaiman, 1989). Our cantilever system eliminates these forces by applying a largely translational force (relative to the teeth) just under the chin.

The torque motor was controlled by a Digital Signal Processer (Texas Instruments TMS320C40). A force sensor (Nitta UFS-3012A15) was installed between the motor and the beam. As a result of active load compensations by the controller, the gravitational forces on the jaw attachment, the bar, and the beam were cancelled, and their inertial forces were reduced to ensure natural movements as much as possible, resulting in an inertial load of 101 g.

The positions of the upper lip, lower lip, jaw (beam of teeth-splint), and two nose markers were measured with a 3D-optical position sensor (OPTOTRAK 3020) at 250 Hz. The nose markers were used for off-line calibration. We confirmed that head movement caused by the jaw perturbation was too small to affect the markers' positions (and their derivatives) relative to the nose. The acoustic signal was simultaneously recorded at 16 kHz after low-pass filtering (8 kHz).

In the perturbation experiments of the sustained and sentence production tasks (see the Experimental procedure section), the EMG of orbicularis oris superior (OOS) and inferior (OOI) were recorded using bipolar surface electrodes (Ag/AgCl; MEDICOTEST NEUROLINE700) placed 18–20 mm apart center to center. These signals were digitally recorded (2 kHz) after filtering (50 Hz–1.5 kHz band pass), and then rectified and smoothed (temporal averaging of 5 ms for the sustained phonemic task to maintain time accuracy and 10 ms for the sentence production task). The electrodes for EMG of OOS were placed just above the vermilion border of the upper lip on the right side, and those for OOI were placed just below the vermilion border of the lower lip on the same side.

To investigate the activities of the perioral muscles, in the distinct experiment (see Fig. 8), the EMG of the upper lip elevation muscles (ULE) and the depressor anguli oris (DAO) were recorded using a pair of insulated hooked wires (Alloy P91, 0.05 mm diameter) with their tips disinsulated 0.5 mm and bent. To insert the electrodes in the levator labii superioris (one of ULE), the needle (30 gauge, 19 mm) was inserted 10 mm deep at a point about 13 mm lateral and 5 mm above the inferior border of the nasal alae, and the tip of the needle was directed 45° upwards to the maxillary process. After the procedure, the subject was asked to snarl to activate ULE for verification of the electrode placement as in O'Dwyer, Quinn, Guiter, Andrews & Neilson (1981). For the DAO, two single hooked wires were inserted separately 3 mm apart and with different depths (5–8 mm approx.). The needles were inserted at a point approximately 25 mm lateral from the mouth angle and half way to the inferior edge of the mandible, and the tip of the needle was directed towards the mouth angle. The placement was verified by having the subject lower the corners of his mouth. In this experiment, the EMG of the OOS was also recorded using bipolar surface electrodes. All EMG signals were stored at 24 kHz in a digital signal recorder (SONY PC216AX) after low-pass filtering (3 kHz), and then they were filtered (high pass with 30 Hz), rectified, and smoothed (10 ms) digitally by software.

2.2. *Experimental procedure*

In the first experiment, the subjects were asked to say "kono /aΦ/" or "kono /aΦa/" and then sustain the last phoneme /Φ/ or /a/ with static posture (3.0 s) repetitively. A step perturbation (4.0 N, 1.0 s, jaw open direction) was applied to the jaw during the sustain phase of /Φ/ or /a/ in 20% (20/100) of the trials randomly selected. Two male adults (Japanese natives: A, B) participated in this experiment. After this experiment, EMG signals during strong protrusion were recorded as references to maximum values to check the range of EMG variation.

In the second experiment, the subject was asked to say a carrier sentence "kono /aΦaΦa/ mitai" with the assistance of beeping sounds (fundamental freq.: 2320 Hz; duration: 20 ms; tempo: 1.67 Hz), repetitively. A step perturbation (4.0 N, jaw open direction) was triggered 0, 30, 60, 90, or 120 ms after the start of jaw elevation from the first /a/ (see Fig. 4), which was detected by the zero crossing of the velocity of the motor encoder, during 10% (50/500) of the randomly selected trials. All perturbations started between the beginning of the first /Φ/ and the beginning of the second /a/, and ended around the last phase of the third /a/ indicated by "Release" in Fig. 4. Four subjects (all male adults, Japanese natives: A, C, D, E) participated in this experiment.

2.3. *Data analysis*

To examine the perturbation effects, we must extract in the data analysis the positional changes of the upper lip and the jaw caused by the perturbations. Because of the trial fluctuations in the sentence production task, large errors could occur if we were to extract the positional change by taking the simple difference between the averaged control and perturbed trajectories. To reduce the influence of trial variation on the extraction of positional change, we used the following method. We excluded failed trials in which speech speeds of sentence production were found to be clearly different from others by the visual inspection of the response pattern. For each perturbed trajectory, we also chose the top 10 control trajectories having the highest correlation between the control and perturbed lip-trajectories during the time between the onsets of acoustic signal and perturbation. The mean of the correlation values of the selected trials for all subjects was 0.97, whereas that of all trials was 0.80. The mean trajectory of the selected control trials for each perturbed trial was subtracted from the corresponding perturbed trajectory after canceling the positional offset at the start of the jaw shift by perturbation. Note that, in the analysis of sustained phonemic tasks, the procedure for control trial selection was not required because the position before perturbation was almost constant during the task, but the procedure for offset elimination was applied to reduce the error caused by trial variation in extracting positional change.

3. Results

3.1. *Sustained productions of /Φ/ and /a/*

Fig. 2 shows the responses of applied force detected by the force sensor, the motions of the upper lip (UL), lower lip+jaw (LL+J), jaw (J), and EMG of OOS and OOI

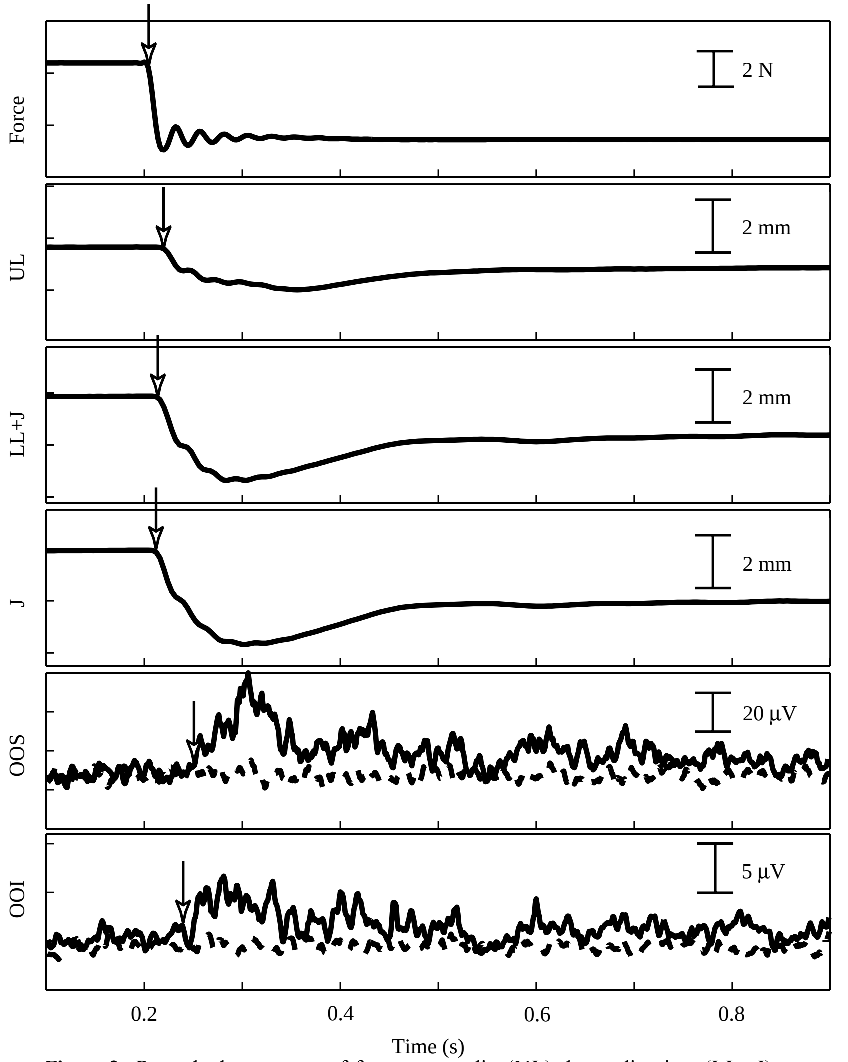

Figure 2. Perturbed responses of force, upper lip (UL), lower lip+jaw (LL+J), jaw (J), and EMG of OOS and OOI during a sustained production /Φ/ (Subject A). EMG was rectified and smoothed with 5 ms temporal averaging. The arrows indicate the onsets of responses to the perturbation detected by the method described in the text.

TABLE I. Downward displacements of jaw and upper lip by jaw perturbation and EMG activity of OOS during sustained productions of /Φ/ and /a/. The values in the parentheses denote standard deviations of trial variation

	Subject A		Subject B	
Articulator (mm)	/Φ/	/a/	/Φ/	/a/
Jaw (x_j)	2.21 (0.15)	2.74 (0.28)	2.17 (0.24)	1.42 (0.22)
Upper lip (x_u)	0.87 (0.10)	0.34 (0.03)	1.41 (0.14)	0.40 (0.08)
EMG (OOS) (µV)	28.7 (4.6)	18.7 (3.5)	34.9 (7.5)	16.2 (4.6)

during the sustained production /Φ/ (Subject A, mean of 18 perturbed trials). When the jaw was perturbed downward (fourth panel from the top), the upper lip moved downward (second panel from the top). This downward shift of the upper lip can be regarded as a compensatory movement to maintain a constriction between the lips for the /Φ/ productions. The EMGs of OOS and OOI (solid line in the bottom two panels) increased gradually after the application of the jaw perturbation. After the initial transient phases (0.2–0.4 s), the EMGs of OOS and OOI were sustained over the corresponding control responses (dashed line in each panel). The percent change (Weber & Smith, 1987) of the averaged EMG between 0.5 and 0.8 s was 40% during the /Φ/ productions in this case.

Table I summarizes the displacements of the jaw and upper lip 40 ms after the load onset and the EMG activity (averaged 100 ms) just before the load onset during sustained productions /Φ/ and /a/ for the two subjects. As shown in this table, for subject A, jaw displacement caused by the perturbation for /Φ/ was significantly smaller ($p<0.05$ for the null hypothesis by *t*-test) than that for /a/, but was significantly larger for subject B. For both subjects, the upper lip downward shift for /Φ/, on the other hand, was significantly larger than that for /a/: this displacement helps to keep the bilabial constriction for sustaining the production of /Φ/ even for unanticipated perturbations.

As for the muscle activity shown in Table I, the EMG of OOS for /Φ/ was approximately two times higher than that for /a/ phonemic tasks for both subjects. This EMG increase occurs in order to form a lip configuration for bilabial constriction. Note that the EMGs of OOS for /Φ/ before the load onsets were 21.1 and 26.2% of the maximum for subjects A and B, respectively.

To investigate the causal relationship between the behavioral and EMG responses, we quantified the latencies of these responses to the perturbation. The onset of the perturbation was detected by the peak of the second-order derivative of force signal (see the arrow in the top panel of Fig. 2). The onsets of positional responses of the upper and lower lips and jaw were detected using a particular threshold ($1500\,mm/s^2$) of the corresponding acceleration (see the arrows in Fig. 2). This threshold was set to avoid mis-detection of the perturbed movement because the threshold of the standard deviation of acceleration before perturbation for two subjects was $<1403\,mm/s^2$. The onsets of EMG of the OOS and OOI were determined by detecting the difference between the averaged control trial response ($\pm$3S.D. of 30 trials) and a single perturbed trial within a particular time window (until 100 ms after the load onset). The arrow in each panel of Fig. 2 indicates the detected onset of the response associated with the perturbation.

Fig. 3 shows the mean latencies (trimmed minimum and maximum) of those responses during /Φ/ of the two subjects. As shown in this figure, for both subjects, the jaw moved first, followed by the lower lip+jaw and the upper lip, sequentially. These onset differences support the idea that the perturbation force propagates in a step-by-step manner through mechanically connected articulators. On the other hand, the latencies of the EMG of OOS were 47.5 ms ($\pm$27.5 ms) for subject A and 49.0 ms ($\pm$47 ms) for subject B. The downward shift of the upper lip preceded the increase in the EMG of OOS associated with the perturbation, suggesting that the initial downward shift of the upper lip was not induced by the EMG increase.

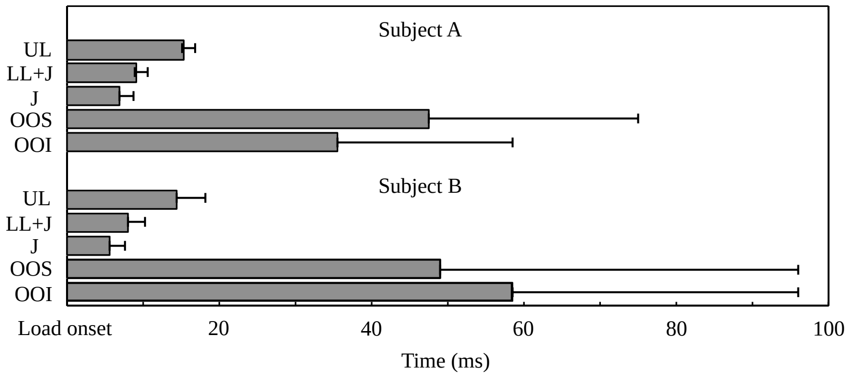

Figure 3. Latencies of the displacements of UL, LL+J, J, and EMG changes of OOS and OOI during sustained production /Φ/ of two subjects. The error bars denote the standard deviations of the corresponding values.

3.2. *Sentence productions*

Fig. 4 illustrates audio signal, articulatory movements, and EMG activities of OOS during the unperturbed utterance of "kono /aΦaΦa/ mitai" (Subject A, mean of seven control trials). The second to fifth panels show upper lip, labial distance, lower lip plus jaw, and jaw movements, respectively. As shown in the second panel, the upper lip movement has three downward dips, which correspond to the first and second /Φ/ and to /m/ in the sentence. At the three phonemic tasks of /a/, the upper lip moves upwards. Conversely, the trajectories of the jaw and lower lip+jaw have three upward peaks for the two /Φ/ and /m/, which are close to being mirror images of the upper lip movement. The EMG of OOS was roughly synchronized, but slightly shifted forward in time, with the downward movement of the upper lip.

Fig. 5 demonstrates two perturbed behaviors with a control behavior between 0.4 and 0.8 s. The solid and dash-dot lines denote the trajectories (mean of seven trials) perturbed 30 ms and 90 ms after the start of jaw elevation of first /a/ (these perturbation triggers are indicated by the second and fourth arrows in Fig. 4), and the dashed line denotes the unperturbed trajectory (mean of seven trials). Note that to show the details of the sudden response to the perturbation in the 30 ms condition, the response is plotted by a solid line in each panel. As shown by the solid line in the bottom panel, the jaw was suddenly moved downward by the perturbation during the upward movement for the first /Φ/, and it did not recover to its position in the control trial during the perturbation. The lower lip+jaw trajectory (third panel) also shifted downward just after the perturbation, but went back to its position in the control trial around the second /a/, indicating that the lower lip compensated the jaw depression with a certain delay. Due to the downward shift of the upper lip (solid line in the top panel), the distance between the upper and lower lips (solid line in the second panel) was close to that of the control trial.

When the jaw was perturbed in the mouth-opening phase of preparing for /a/ (dash-dot line in Fig. 5, loaded 90 ms), the jaw was more largely shifted downward by the load than it was by the load applied during the upward movement of preparing for the /Φ/ (solid line). As in the above case, the jaw depression was

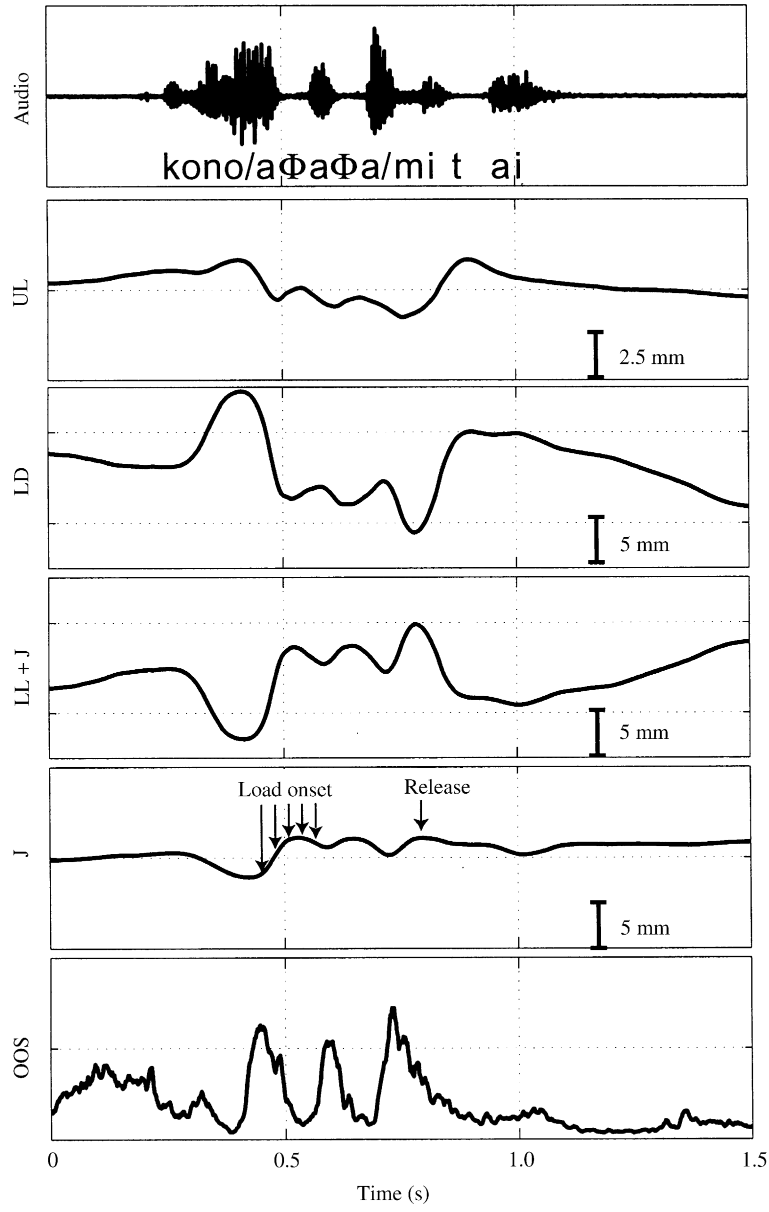

Figure 4. Articulatory responses during sentence production. From top to bottom, audio signal, upper lip (UL), labial distance (LD), lower lip (LL)+jaw (J), and jaw (J) responses, and the EMG (rectified and smoothed with 10 ms temporal averaging) of OOS during the unperturbed utterance of "kono /aΦaΦa/ mitai" (mean of seven control trials, Subject A). The arrows in the fifth graph from the top denote the load onsets at five different times and the load release time.

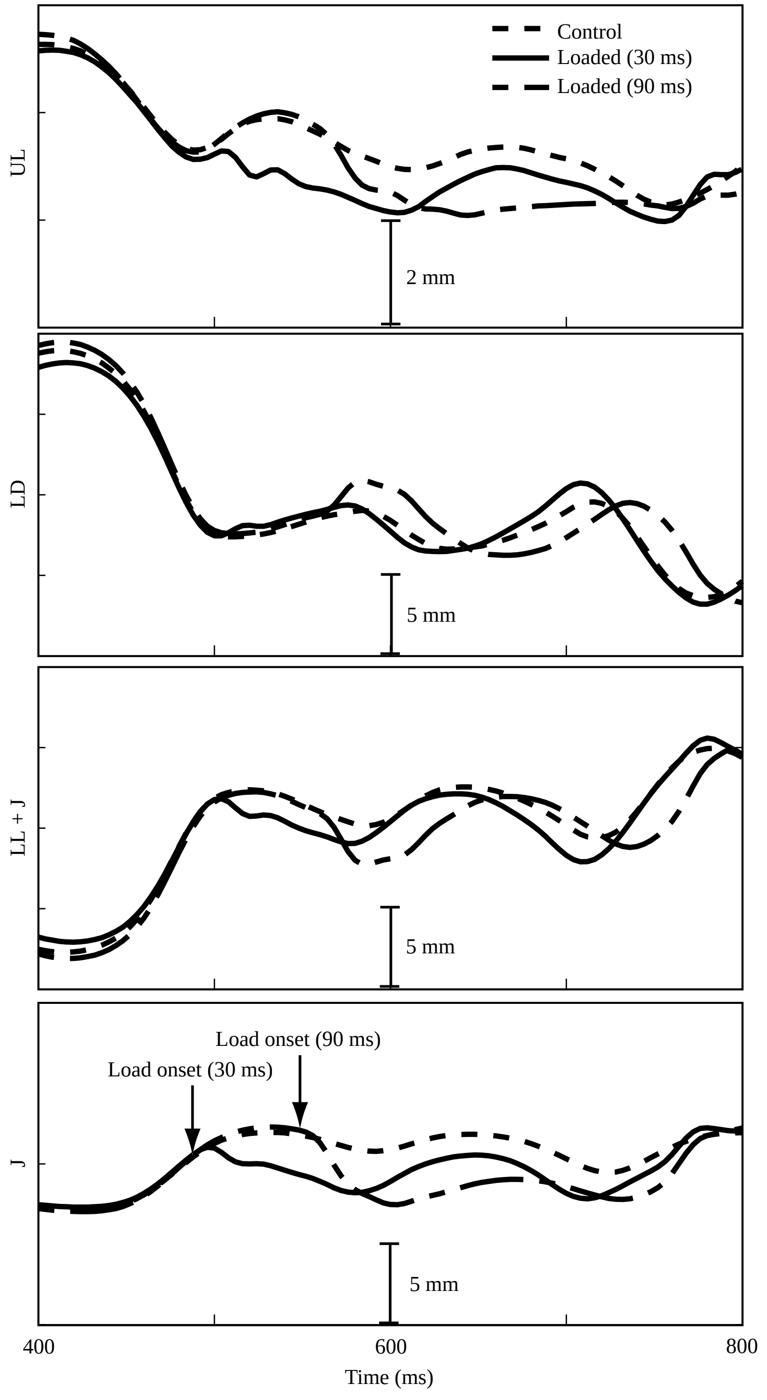

Figure 5. Articulatory trajectories during perturbed (solid line: load onset 30 ms after the jaw elevation; dashed-dot line: 90 ms after the jaw elevation) and unperturbed trials (dashed line) for subject A. Top to bottom: upper lip (UL), labial distance (LD), lower lip (LL)+jaw (J), and jaw (J) responses.

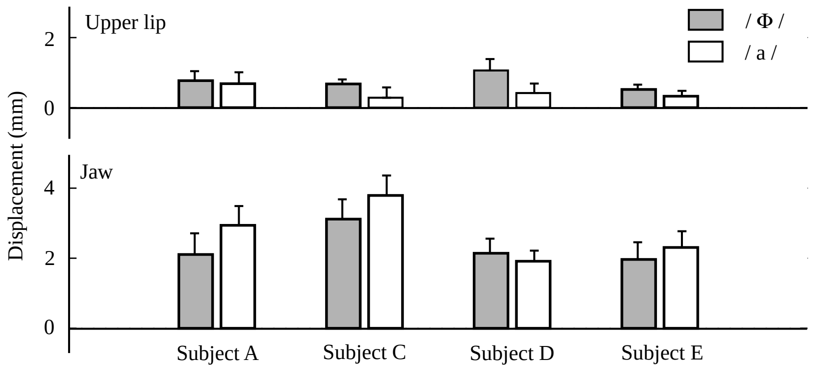

Figure 6. Downward displacements of the upper lip and the jaw 40 ms after the load onsets at the first /Φ/ and the second /a/ in the sentence production of four subjects. These load onsets for each phoneme production correspond to the second and fourth ones for Subject A, and the third and fifth ones for Subjects C, D, and E. Each error bar denotes the standard deviation of trial variation of perturbed trials. The displacement extraction procedure is placed in the Method section.

compensated by the lower lip moving upward before the second /Φ/ production (dash-dot line in the third panel), and the bilabial constriction for the second /Φ/ (dash-dot line in the second panel) was achieved with the assistance of the downward movement of the upper lip (top panel). The downward shift of the upper lip was slightly slower for this perturbation than for the perturbation in the 30 ms condition (solid line in the top panel), whereas the jaw downward speed was faster for this perturbation than for the 30 ms condition. Although the phase of the labial distance was slightly retarded, its temporal pattern was almost the same as that of the control trial after the second /Φ/ production.

Fig. 6 summarizes the displacements (40 ms after the load onset) of the upper lip and jaw caused by the perturbation to /Φ/ and /a/ productions for all subjects. The upper lip shift was significantly larger ($p < 0.05$ for the null hypothesis by *t*-test) for /Φ/ than for /a/ for Subjects C, D, and E. For Subject A, the upper lip shifts for these two phonemic tasks were not statistically different, but the jaw shift was significantly smaller for /Φ/ than for /a/, whereas, for Subjects D and E, the jaw shifts for these two phonemic tasks were not statistically different from each other.

3.3. *A model of interaction between upper lip and jaw*

To characterize the interaction between the upper lip and jaw using the above observations, here we will derive a dynamical model of the upper lip movement. Due to the muscle inherent properties, muscle stiffness increases as muscle activity increases under isometric conditions (Gottlieb & Agarwal, 1988; Kearney & Hunter, 1990; Osu & Gomi, 1999). Additionally, it has been reported that the mechanical impedance of the musculoskeletal system is governed by the coordination of multiple-muscle activations (Hogan, 1984; Gomi & Osu, 1998). From these studies it is inferred that, even in articulatory muscle coordination, the stiffness of perioral muscles also varies with different combinations of muscle activations.

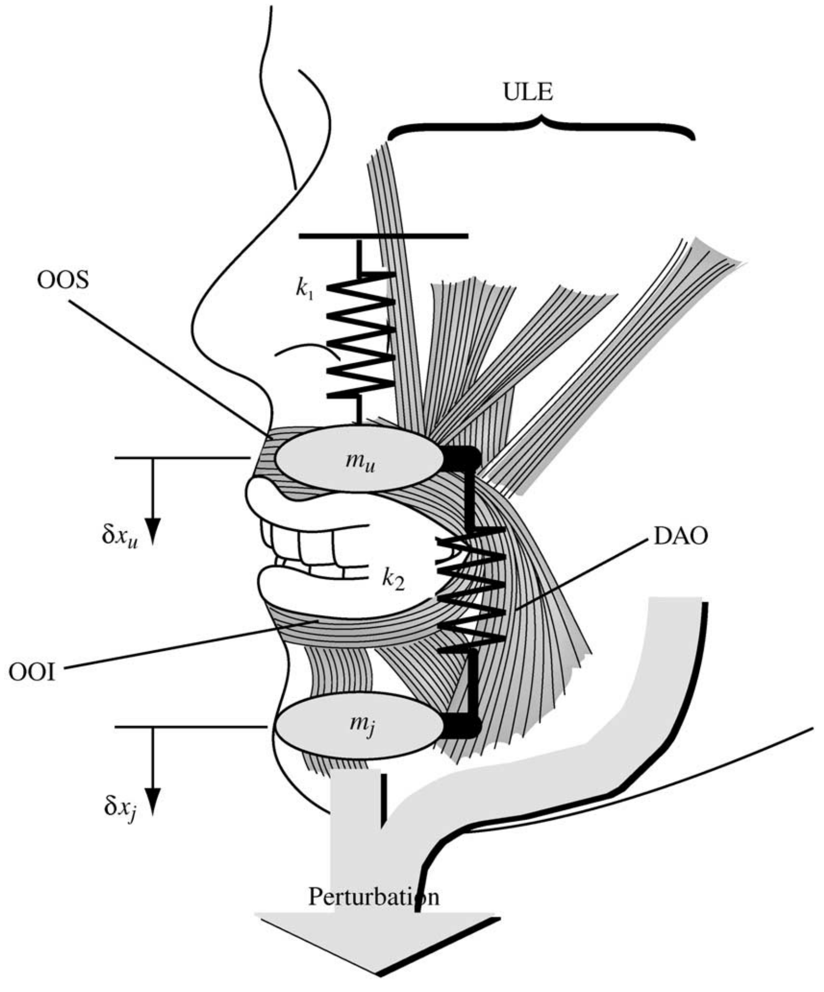

Figure 7. A serial connected spring model that represents the mechanical linkage of the perioral muscles and tissues between the upper lip and jaw. OOS, OOI, DAO, and ULE denote orbicularis oris superior, orbicularis oris inferior, depressor anguli oris, upper lip elevation muscles, respectively.

Fig. 7 shows the configuration of the perioral muscles, which are regulated to form labial postures. Folkins (1978) demonstrated that the electrical stimulation of DAO causes an inferior movement at the corner of the mouth. It suggests that the simultaneous activations of OOS and DAO induce the upper lip downward movement due to the muscle connections. The upper lip elevation muscles (levator labii superior, levator anguli oris, zygomaticus major and zygomaticus minor) indicated by ULE in the figure are activated for pulling up the upper lip. We modeled this muscle linkage by a mass–spring connection as shown in this figure. The jaw (m_j), upper lip (m_u), and head (or upper perioral matter) are serially connected by two springs. Because the upper lip is driven by forces generated by the upper-lip elevation muscles (f_1) and of the upper-lip depressor muscles (f_2), the following dynamical equation can be stated:

$$m_u \ddot{x}_u = f_1 + f_2 \tag{1}$$

where $\ddot{x}_u$ is the acceleration of the upper lip. As a first-order approximation, the muscle forces, f_1 and f_2, can be represented as

$$f_1 = -k_1(\hat{x}_u - x_u) - c_1\dot{x}_u, \quad f_2 = k_2((\hat{x}_j - x_j) - (\hat{x}_u - x_u)) + c_2(\dot{x}_j - \dot{x}_u) \tag{2}$$

Here, x_u and x_j are the positions of the upper lip and jaw, $\dot{x}_u$ and $\dot{x}_j$ are their velocities, k_1 and k_2 are the stiffness values, and c_1 and c_2 are the viscosity values of the corresponding muscles (including perioral soft tissues), respectively. The $\hat{x}_u$ and $\hat{x}_j$ denote the equilibrium positions of the upper lip and jaw, which vary according to the muscle activities. Considering these relationships, the variational equation of Equation (1) can be represented as

$$m_u \delta\ddot{x}_u = k_1\delta x_u - c_1\delta\dot{x}_u - k_2(\delta x_j - \delta x_u) + c_2(\delta\dot{x}_j - \delta\dot{x}_u) \tag{3}$$

If it were possible to obtain sufficiently excited variational components of all terms in Equation (3) (i.e., the upper lip acceleration, velocity and position, and of the jaw velocity and position), we could estimate the ratios of all unknown parameters in this equation as demonstrated in Gomi & Kawato (1996). During a very brief period (<50 ms) in speech tasks, however, it may be difficult to apply rich perturbation without causing a large disturbance to articulation. For the sake of simplicity, we assume that the stiffness components (position dependent terms in Equation (3)) are relatively dominant in this relationship at a certain time, and then we obtain

$$k_2(\delta x_j - \delta x_u) \cong k_1\delta x_u \tag{4}$$

Note that we do not assume that this relationship is always valid, rather we assume it to be valid when dynamic forces are relatively small or cancel each other out. By using dynamic simulation of the interaction of the upper lip and jaw, we confirmed that this assumption is mostly valid 40 ms after the perturbation (unpublished observation). If we could measure the force twitching the muscle between the upper lip and jaw, the stiffnesses, k_1 and k_2, could be directly estimated by using Equation (4). We cannot, however, decompose the perturbation force applied to the jaw into the force pulling the muscles between the jaw and upper lip and that pulling the muscles between the jaw and head.

We will now consider the variation of k_1, which represents the stiffness of the upper-lip elevation muscles (ULE). Fig. 8 shows the audio signal, and EMG activities of the ULE, OOS, and DAO (mean of 10 trials) during the utterance of "kono /aΦaΦa/ mitai" without any perturbation.

The EMGs of OOS and DAO were roughly synchronized to each other for forming bilabial constrictions, suggesting that the stiffness of these muscles (k_2) increased. On the other hand, there was no obvious change in the EMG of the ULE. Even in the perturbed trials, the ULE showed a small and very brief activation change (preliminary observation), which may not cause a great change in the ULE stiffness. This difference in EMG response to the perturbation may be supported by Smith, McFarland, Weber & Moore (1987) who demonstrated that the EMG response is greatly dependent on the mechanical stimulus location. Based on

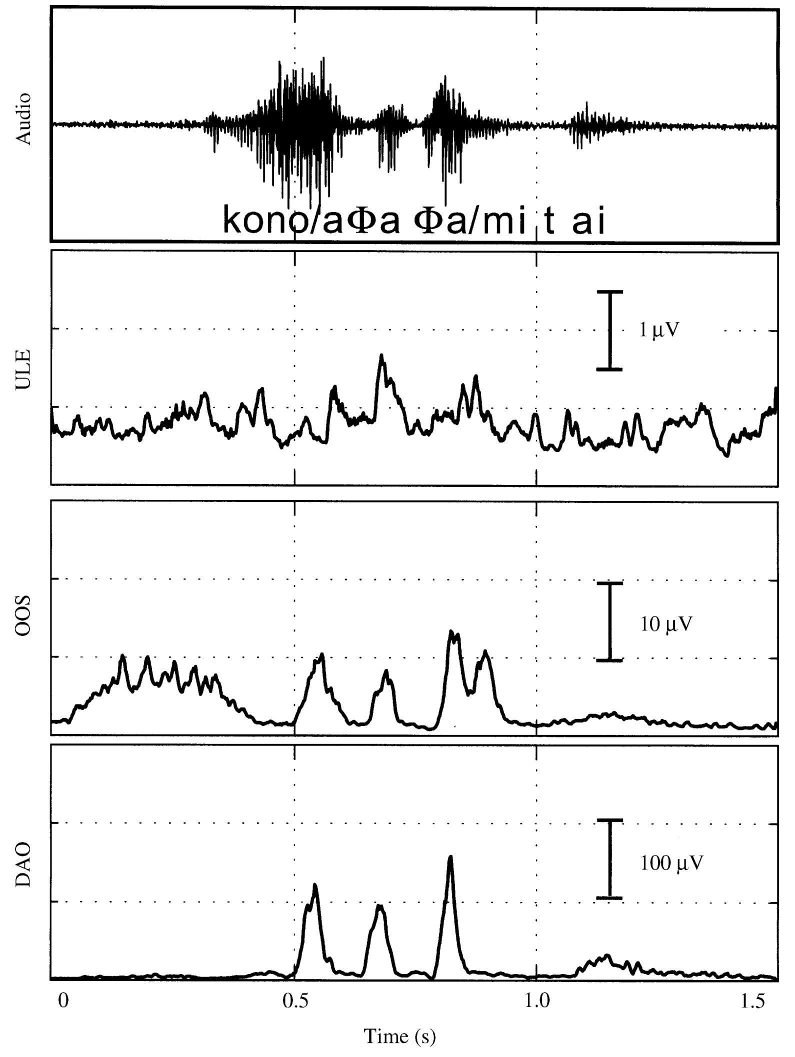

Figure 8. Audio signal and EMG (rectified and smoothed with 10 ms temporal averaging) of ULE, OOS and DAO during the sentence production of subject A.

these observations, we assume that the stiffness k_1 is constant during the utterances. Under this assumption, we can characterize the ratio between the muscle stiffness values k_2 at two different phonemes (u_1, u_2) as follows:

$$\frac{k_2(u_2)}{k_2(u_1)} = \frac{\delta x_u(u_2)(\delta x_j(u_1) - \delta x_u(u_1))}{\delta x_u(u_1)(\delta x_j(u_2) - \delta x_u(u_2))} \tag{5}$$

Here, δx_u and δx_j express the displacements of the upper lip and jaw caused by the perturbation, respectively. This relationship enables us to estimate the stiffness variation (i.e., relative stiffness) of the linkage connecting the upper lip and jaw for all different phonemes.

3.4. *Stiffness variation*

By putting the displacements caused by the perturbations during the productions /Φ/ and /a/ into Equation (5), we can calculate the ratio between the stiffness values of the linkage connecting the upper lip and jaw during these productions. For the sustained (static) productions, given the displacements shown in Table I, the values of the stiffness ratio $k(\Phi_s)/k(a_s)$ were 4.65 ± 0.20 (Subject A) and 4.78 ± 0.27 (Subject B). For the sentence (dynamic) productions, the stiffness ratios $k(\Phi_d)/k(a_d)$ for four subjects are listed in Table II. Here, subscripts "*s*" and "*d*" denote "static" and "dynamic" conditions. Note that the means and standard deviation (S.D.) of the stiffness ratios were obtained by using a resampling (Bootstrap) method with 1000 random replications (Efron & Tibshirani, 1993). All these stiffness ratios are significantly larger than one ($p < 0.01$ for the null-hypothesis confirmed by the bootstrap statistical test), indicating that the stiffness of the linkage connecting the upper lip and jaw for /Φ/ was greater than that for /a/ during sustained and sentence productions.

Similarly, we can compare the stiffness in the sustained (static) and sentence (dynamic) productions. The ratios $k(\Phi_d)/k(\Phi_s)$ and $k(a_d)/k(a_s)$ for Subject A were

TABLE II. Stiffness ratio $k(\Phi)/k(a)$ (mean $\pm$ S.D.) during the sentence production for four subjects

	Subj. A	Subj. C	Subj. D	Subj. E
$k(\Phi)/k(a)$	1.95 ± 0.35	3.47 ± 0.27	3.86 ± 1.35	2.18 ± 0.12

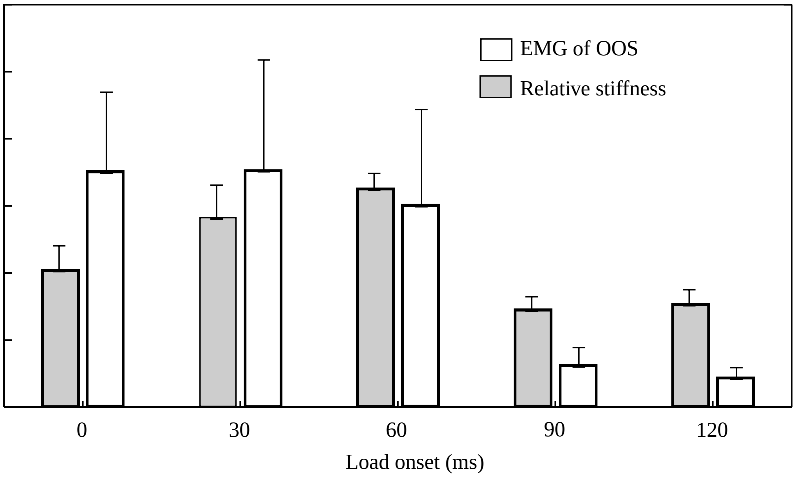

Figure 9. Temporal change in stiffness of the linkage between the upper lip and jaw characterized from the displacements of the upper lip and jaw caused by five kinds of perturbations, and temporal change in the EMG of the OOS (30 ms before each load onset) (Subject A). Stiffness values for the second to fifth perturbations (load onsets: 30, 60, 90, 120 ms) were obtained relative to the mean stiffness for the first perturbation (load onset: 0 ms). The error bar of the stiffness denotes the standard deviation of the bootstrap estimates, and the error bar of the EMG denotes the standard deviation of trial variation.

0.91 ± 0.17 and 2.16 ± 0.31, respectively. This increase in the dynamic stiffness of /a/ may be ascribed to the coarticulatory effect in the utterance where the vowel occurred between the repetitions of the same consonant.

Fig. 9 shows the variations of the relative stiffness ($k(t_n)/k(t_1)$, $n = 1, 2, 3, 4, 5$) and the EMG activity (OOS) (30 ms before the load onset as a rough estimation of mechanochemical delay) for each perturbation applied at the five different times during sentence production (Subject A). Here, $t_n (n = 1, 2, 3, 4, 5)$ indicates the perturbation timing ($0, 30, 60, 90, 120$ ms). For this subject, the second (t_2, 30 ms) and fourth (t_4, 90 ms) perturbations were applied around the first /Φ/ and the second /a/ productions, respectively. The stiffness increases around the /Φ/ production and decreases around the /a/ production as shown in this figure. Additionally, the stiffness variation nicely correlates with the EMG activity of OOS ($r = 0.767$ mean for Subjects A, C, D, and E). This result suggests that the compensatory movement for maintaining the labial constriction is realized by passive dynamics regulated by muscle activation according to the speech task.

4. Discussion

4.1. *Contributions of passive dynamics and sensory feedback*

In the preceding sections, we have shown the task-dependent variation of the upper lip compensatory movements and have calculated the stiffness of the muscle linkage between the upper lip and jaw. In addition, the EMG of OOS was highly correlated with the stiffness variation. Based on these observations and analyses, we ascribe the quick downward movement of the upper lip for the jaw perturbation to the increased stiffness of the muscle linkage between the upper lip and jaw.

One could claim that the downward shifts of the upper lip after the perturbation can be ascribed to the downward shift of the lower lip preventing the upper lip downward movement during /Φ/ production. As shown by the solid line in Fig. 5, the downward shift of the upper lip after the perturbation (first panel) occurred in the increasing phase of the labial distance (second panel). This means that the upper lip moved upward faster than the lower lip, suggesting that the upper lip downward shift was not caused by removing a movement block of the lower lip. In addition, as observed in the response perturbed around /a/ (dash-dot line) in Fig. 5, the upper lip was accelerated downward. Considering that the upper and lower lips were parted in this phase, the downward acceleration of the upper lip may be ascribed to a twitch through the mechanical linkage between the upper lip and jaw.

A major merit of mechanical linkage is a fast reaction speed, which is crucial for real-time control, whereas the latencies of neural transmission and mechanochemical dynamics cannot be avoided in the responses caused by neural linkages. Additionally, passive dynamics of the linkage connecting articulators would automatically compensate for fluctuations in motor command and perturbations caused by body and head movements, thus robustness of articulation may increase. For example, in an articulation of /Φ/, by increasing the stiffness of the linkage connecting the upper and lower lips/jaw, a bilabial constriction would not be violated by perturbations and context variations.

Although we focus on the advantage of controlling passive dynamics here, we do not deny the contributions of heterogenic neural linkage for cooperative actions.

Several studies (Folkins & Abbs, 1975; Folkins & Zimmermann, 1982; Abbs & Gracco, 1983, 1984; Kelso *et al.*, 1984; Gracco & Abbs, 1985; Shaiman, 1989) have suggested that the compensatory movement of the upper lip associated with jaw or lower-lip perturbation is induced by neural linkage using sensory information. The observed latencies in muscle response of the upper lip or tongue during bilabial stops or lingo-dental fricatives or stops were 20 ± 18(S.D.) ms (OOS), 15–35 ms (OOI) in (Kelso *et al.*, 1984), and 22–75 ms (OOS and OOI) in (Abbs & Gracco, 1984), which were longer than latencies of the perioral reflex (12–18 ms in Abbs & Gracco (1984); 14–17 ms in Weber & Smith (1987)). The latencies of the EMGs in the present study were comparable with those in the Kelso and Abbs studies, and the displacement of the upper lip was much faster than the EMG response as shown in Fig. 3. Therefore, it may be impossible to generate quick displacements by using neural linkage because of latencies in neural transmission and mechanochemical dynamics. Instead, we suggest that the passive dynamics contribute to generate a quick phase of compensatory behavior, and neural reflex feedback and voluntary modifications regulate the slow phase of compensatory behavior. This is partly supported by a simulation study (Ito, Gomi & Honda, 2000*b*) in which passive-dynamics was found to reproduce the initial phase of the compensatory movement for the perturbation, but not to perfectly mimic its later phase. The present study suggests that, not only by regulating neural linkage but also by controlling passive dynamics (namely stiffness), fluent continuous articulatory movements can be generated under a variety of conditions.

4.2. *Functional organization of the perioral muscle stiffness*

An adjustment in coordinated articulatory movements due to passive dynamics was reported by Kelso *et al.* (1984). They found that the additional downward movements of the upper lip and upward movements of the lower lip without any additional EMG increases were caused by a jaw opening perturbation. Their interpretation was that these responses were passive overshoots caused by "momentum" rather than by nonautogenic neural linkages.

Unlike the effects due to inertial dynamics, passive dynamics of "stiffness" can be changed by altering muscle activation. As demonstrated above, muscles connecting the upper lip and jaw (OOS and DAO) show large activity for the bilabial fricative /Φ/. As a result of these muscle activities, the upper lip is depressed downward to form a bilabial constriction for /Φ/. Additionally, this increase in the EMG could be, as mentioned in the Results section, accompanied by an increase in stiffness of the linkage between the upper lip and jaw due to muscle inherent characteristics, and then the upper lip easily moves together with the jaw when the jaw is depressed by a perturbation.

On the other hand, as compared in the Results section, for the /a/ productions, the downward displacement of the upper lip was small because of the low stiffness of the linkage between the upper lip and jaw resulting from the low muscle activations. Unlike /Φ/, the /a/ production does not require keeping a particular gap between the upper and lower lips, implying that the stiffness is regulated according to task requirements. Since the stiffness of the linkages among organs defines their interaction, the regulation of stiffness may be beneficial in controlling articulatory cooperative behaviors. Therefore, the stiffness regulation mechanism can be regarded as a strategy for accomplishing speech tasks under various conditions.

Note that even though our results suggest the importance of muscle spring-like properties, the classical mass–spring model (Fowler & Turvey, 1980; Feldman & Levin, 1995) in which the motion target is encoded by an invariant parameter, i.e., the equilibrium position, of each articulator is not advocated. Levelt (1993) pointed out that the mass–spring model in its simple form fails in handling compensatory articulatory adjustments, one of which we demonstrated here. The stiffness regulation we argue here is not for generating movement to achieve the equilibrium position. Instead, by describing the interaction among articulatory organs rather than the movement of each organ itself, articulatory behaviors can be depicted by smaller degrees of freedom of motor commands, and then the articulatory gestures would be robust to contextual change of phonations and unanticipated perturbation. In this sense, our stiffness regulation hypothesis is in accord with the task-dynamics model proposed by Saltzman (1986) in which each task is embedded in the total dynamics of articulators and controllers. Additionally, in the learning process, error detection between prediction and consequence of sensory information (Lindblom, Lubker & Gay, 1979) may be important in changing internal models for control. Further investigation is needed in order to examine these hypotheses and model the general control mechanisms for articulatory movements.

We thank K. Ishii and T. Hirahara (NTT CS Labs.) for their continuing encouragement. We also thank K. Honda (ATR) for his technical assistance in EMG recording, M. Wakumoto (Showa University) for his advice in developing experimental devices, and T. Konno and M. Sawada (NTT-AT) for their support in software development.

References

Abbs, J. H. & Gracco, V. L. (1983) Sensorimotor actions in the control of multi-movement speech gestures, *Trends in Neurosciences*, **6**(9), 391–395.

Abbs, J. H. & Gracco, V. L. (1984) Control of complex motor gestures: orofacial muscle responses to load perturbations of lip during speech, *Journal of Neurophysiology*, **51**(4), 705–723.

Efron, B. & Tibshirani, R. J. (1993) *An Introduction to the Bootstrap*, Vol. 57. London/Boca Raton FL: Chapman & Hall/CRC.

Feldman, A. G. & Levin, M. F. (1995) The origin and use of positional frames of reference in motor control, *Behavioral and Brain Sciences,* **18,** 723–806.

Folkins, J. W. & Abbs, J. H. (1975) Lip and jaw motor control during speech: responses to resistive loading of the jaw, *Journal of Speech and Hearing Research*, **18**(1), 207–220.

Folkins, J. W. (1978) Lower lip displacement during in-vivo stimulation of human labial muscles, *Archives of Oral Biology*, **23**(3), 195–202.

Folkins, J. W. & Zimmermann, G. N. (1982) Lip and jaw interaction during speech: responses to perturbation of lower-lip movement prior to bilabial closure, *Journal of the Acoustical Society of America*, **71**(5), 1225–1233.

Fowler, C. A. & Turvey, M. T. (1980) Immediate compensation in bite-block speech, *Phonetica*, **37,** 306–326.

Gomi, H. & Kawato, M. (1996) Equilibrium-point control hypothesis examined by measured arm-stiffness during multi-joint movement, *Science*, **272,** 117–120.

Gomi, H. (1998) Anisotropic stiffness reduction during constrained multijoint arm movement. In *Proceeding of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology society*, Hong Kong, pp. 2336–2337.

Gomi, H. & Osu, R. (1998) Task-dependent viscoelasticity of human multijoint arm and its spatial characteristics for interaction with environments, *The Journal of Neuroscience*, **18**(21), 8965–8978.

Gottlieb, G. L. & Agarwal, G. C. (1988) Compliance of single joints: elastic and plastic characteristics, *Journal of Neurophysiology*, **59**(3), 937–951.

Gracco, V. L. & Abbs, J. H. (1985) Dynamic control of the perioral system during speech: kinematic analyses of autogenic and nonautogenic sensorimotor processes, *Journal of Neurophysiology*, **54**(2), 418–432.

Hogan, N. (1984) Adaptive control of mechanical impedance by coactivation of antagonist muscles, *IEEE Transactions on Automatic Control*, **AC-29**(8), 681–690.
Ito, T., Gomi, H. & Honda, M. (2000*a*) Task dependent jaw–lip coordination examined by jaw perturbation during bilabial-consonant utterances. In *Proceedings of the 5th seminar on speech production*, Kloster Seeon, Germany, pp. 41–44.
Ito, T., Gomi, H. & Honda, M. (2000*b*) Model of the mechanical linkage of the upper lip–jaw for the articulatory coordination. In *Proceedings of the 6th international conference on spoken language processing*, Beijing, China, pp. 889–892.
Kearney, R. E. & Hunter, I. W. (1990) System identification of human joint dynamics, *Critical Review in Biomedical Engineering*, **18,** 55–87.
Kelso, J. A. S., Tuller, B., Bateson, E. V. & Fowler, C. A. (1984) Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordicative structures, *Journal of Experimental Psychology: Human Perception and Performance*, **10**(6), 812–832.
Levelt, W. J. M. (1993) *Speaking: from intention to articulation*. Boston: MIT Press.
Lindblom, B., Lubker, J. & Gay, T. (1979) Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation, *Journal of Phonetics*, **7,** 147–161.
Mussa-Ivaldi, F. A., Hogan, N. & Bizzi, E. (1985) Neural, mechanical, and geometric factors subserving arm posture in humans, *The Journal of Neuroscience*, **5**(10), 2732–2743.
O'Dwyer, N. J., Quinn, P. T., Guiter, B. E., Andrews, G. & Neilson, P. D. (1981) Procedures for verification of electrode placement in EMG studies of orofacial and mandibular muscles, *Journal of Speech and Hearing Research*, **24,** 273–288.
Osu, R. & Gomi, H. (1999) Multijoint muscle regulation mechanisms examined by measured human arm stiffness and EMG signals, *Journal of Neurophysiology*, **81,** 1458–1468.
Saltzman, E. (1986) *Task dynamic coordination of the speech articulators: a preliminary model*, Experimental Brain Research Series, Vol. 15, pp. 129–144. Berlin: Springer-Verlag.
Shaiman, S. (1989) Kinematic and electromyographic responses to perturbation of the jaw, *Journal of Acoustical Society of America*, **86**(1), 78–88.
Smith, A., McFarland, D. H., Weber, C. M. & Moore, C. A. (1987) Spatial organization of human perioral reflexes, *Experimental Neurology*, **98**(2), 233–248.
Weber, C. M. & Smith, A. (1987) Reflex responses in human jaw, lip, and tongue muscles elicited by mechanical stimulation, *Journal of Speech and Hearing Research*, **30**(1), 70–79.