Journal of Phonetics (2002) **30,** 281–302
doi:10.1006/jpho.2002.0172
Available online at http://www.idealibrary.com on IDEAL®

Compensatory responses of articulators to unexpected perturbation of the palate shape

Masaaki Honda* and Akinori Fujino

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

Tokihiko Kaburagi

Kyushu Institute of Design, Japan and CREST, Japan Science and Technology Corporation, Japan

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

This paper describes compensatory articulatory behavior in response to an unexpected perturbation of the oral cavity. A mechanical device was used to change dynamically the thickness of an artificial palate; the thickness of the un-inflated artificial palate was increased or the thickness of the already inflated artificial palate was decreased. The palatal perturbations were introduced under conditions of normal and masked auditory feedback. Acoustic, perceptual, and speech movement changes were examined in response to the experimental conditions. An EMA system was used to look for evidence of compensatory articulation during the utterance of repeated CV syllables containing fricative /ʃ/ and affricate /tʃ/. When perturbation by palatal inflation was randomly given just before the initial syllable in the repeated fricative syllable, there was frequent overshoot of the tongue in forming the constriction, causing speech errors in the first syllable with the fricative. Compensation by the tongue for the unexpected palatal inflation became evident around the second syllable and the speech error disappeared in the successive syllables. Even when auditory feedback was masked, rapid compensatory articulation of the tongue was observed around the second syllable, but speech errors randomly occurred in the successive syllables. This fact suggests that tactile feedback gathered by sensing contact between tongue and inflated artificial palate is primarily used to develop the rapid compensation, and that auditory feedback is used in finely adjusting articulation with a longer time lag.

*Address correspondence to M. Honda, Wakamiya 3-1, Morinosato, Atugi, Kanagawa 243-0198, Japan. E-mail: hon@idea.brl.ntt.co.jp

0095–4470/02/$ - see front matter

1. Introduction

Studies on oral-articulatory perturbations have provided valuable insights into motor equivalence, or how different muscle activation patterns can achieve the same articulatory or acoustic goals. In particular, previous studies of steady-state structural perturbations caused by insertion into the mouth of a bite-block (Lindblom, Lubker & Gay, 1979; Fowler & Turvey, 1980; Flege, Fletcher & Homiedan, 1988; McFarland & Baum, 1995, Baum, McFarland & Diab, 1996) or a thick artificial palate (Hamlet & Stone, 1976, 1978; McFarland, Baum & Chabot, 1996) have revealed that articulatory compensation made by adjusting articulator positions allowed the phoneme-specific acoustic goal to be achieved.

In these perturbation studies, important issues have been whether or not the articulatory compensation is immediate and complete and how sensory and auditory feedback is employed in the compensation process. In the earliest bite-block study, talkers succeeded in producing near-normal speech even in their first speaking trial just after the insertions when vowel production was disturbed by fixing the jaw opening with a bite block (Lindblom *et al.*, 1979). Observations of the immediate compensation suggested that speech motor control processes use predictive control in developing compensatory articulation for fixation of the lower jaw. Subsequent studies (Flege *et al.*, 1988; McFarland & Baum, 1995), however, reported that the compensation is neither immediate nor complete as there were significant differences in the acoustic parameters and perceptual data for vowels and consonants produced under bite-block and normal conditions. They also revealed that compensatory strategies may develop over time by employing sensory feedback in the compensation process.

On the other hand, insertion of an artificial palate into the mouth has been found to particularly disturb the articulation of the alveolar fricative. Compensatory articulation, though incomplete, was observed just after the insertion and it continued to gradually improve over the course of 2 weeks (Hamlet & Stone, 1978). A subsequent study (McFarland *et al.*, 1996) supported these findings for incomplete immediate compensation and the improvement over time, and suggested a role of auditory and sensory feedback in the adaptation processes.

These studies revealed the potential contribution of sensory information from auditory and tactile feedback in the immediate compensation as well as in the long-term adaptation process. These previous experiments on structural perturbation used steady-state modification of the vocal tract during the utterance by means of a fixed bite-block or artificial palate; the perturbation was not unexpected, because the talkers were aware of the insertion of these appliances into their mouths. Therefore, it is less clear whether sensory feedback is immediately utilized for the articulatory adjustment when the vocal tract structure is unexpectedly perturbed.

On the other hand, studies of the compensatory response of the upper lip to an unexpected mechanical perturbation of the mandible or lower lip, have suggested that a cortical feedback mechanism is involved in immediate adjustment of the articulation (Folkins & Abbs, 1975; Abbs & Gracco, 1984; Gracco & Abbs, 1985). Other studies on mechanical perturbation, however, revealed a contribution of the mechanical linkage between lips, jaw, and tongue to the immediate compensation (Kelso, Tuller, Vatikiotis-Bateson & Fowler, 1984; Ito, Gomi & Honda, 2000). In contrast to the mechanical perturbation of the articulator, there is little contribution of the mechanical linkage to the articulatory adjustments following structural perturbation. Therefore, examining the compensatory response to unexpected structural perturbation may provide valuable evidence of short-time sensory feedback control in ongoing speech production.

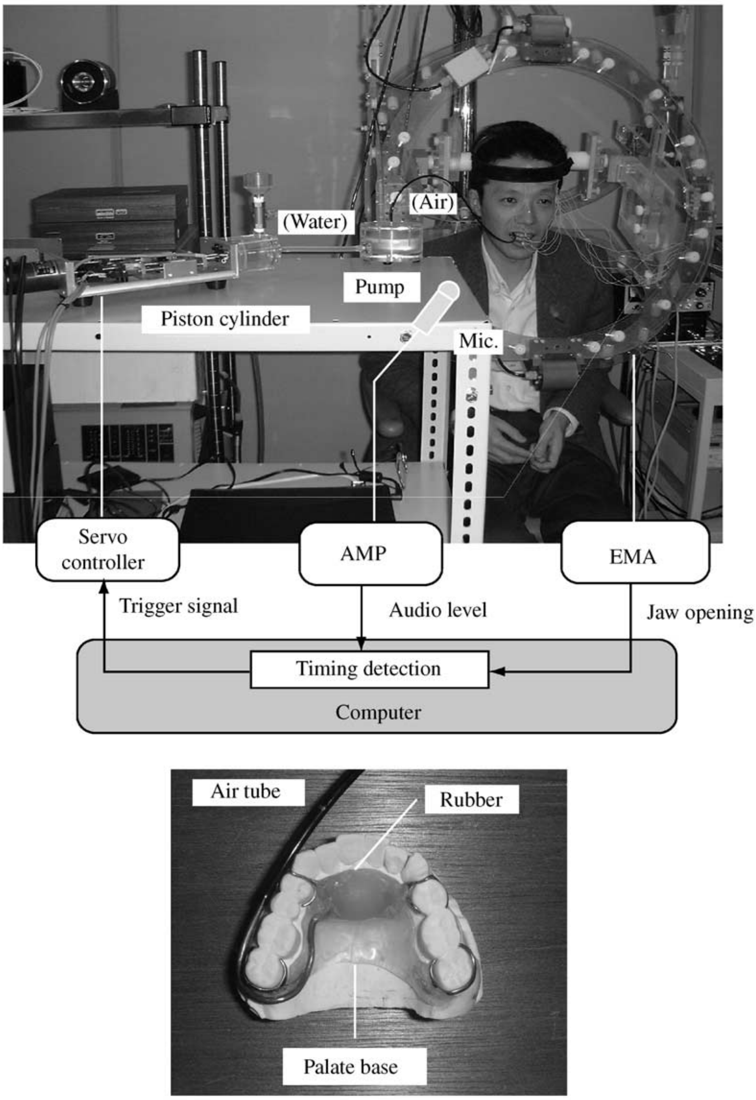
The present study examines compensatory articulation during an unexpected perturbation of the oral cavity. To provide the unexpected perturbations of the oral cavity, we constructed an artificial palate whose thickness can be dynamically changed during speech; the thickness of the un-inflated artificial palate was increased or the thickness of the already inflated artificial palate was decreased. In this study, we focus on the immediate compensation and subsequent adaptation process within repeated CV syllables. Palatalized consonants /ʃa/ and /tʃa/ were selected, because they would be more selectively impaired by the change of palate shape than other sound classes. In order to distinguish the contribution of the tactile information from the auditory information in adjusting the articulation, the compensatory responses are examined under both normal and masked auditory-feedback conditions. The compensatory responses in consonant production are studied in terms of the perceptual, acoustic, and articulatory variables.

2. Method

2.1. *Experimental setup*

For the unexpected perturbation, we designed an artificial palate whose thickness can be changed during speech (Honda & Kaburagi, 2000). The artificial palate, shown in Fig. 1, has a 1-mm-thick acrylic base and a semi-circular 1-mm-thick rubber piece. The size of the rubber part is dependent on the palate size of the subject, but is typically about 2.6 cm in width and 1.6 cm in length. The rubber is attached to the base in the alveolar region using cyanoacrylate adhesive glue to form a balloon. The thickness of the palate can be controlled by inflating the balloon by external air pressure through a lead tube. The top of the balloon is located at the mid-line of the palate 8-mm behind the incisors, and the balloon can be inflated from the un-inflated position to a height of 4 mm. Hence, the total thickness of the artificial palate is the sum of the 1-mm thick base and the 4-mm high balloon, i.e., 5 mm. Previous studies have used a 4-mm thick artificial palate (Hamlet & Stone, 1978) and 6 and 3 mm thick artificial palates (McFarland *et al.*, 1996). The latter study suggested that an alveolar thickness of 6 mm was the largest that could be used without interfering with normal occlusion. Thus, we chose a 5 mm thick artificial palate so as to provide a reliable palatal perturbation without interfering with normal occlusion.

The air pressure was controlled by a piston cylinder that was actuated by a servo-actuator through a ball screw. To improve the response time of the pressure control, a pump located near the subject's mouth was used to convert the exhaust water pressure from the piston cylinder into air pressure. The inflation time of the palate from the rest position to maximum thickness or *vice versa* is approximately 60 ms. Because of the nonnegligible inflation time as compared with velocity of the tongue motion, the perturbation was initiated during production of an open vowel so that the subject would be unaware of the perturbed palate shape until the tongue contacted the palate. The timing of the perturbation was controlled in real time by



Air-controlled artificial palate

Figure 1. Experimental setup and the air-controlled artificial palate.

using both the jaw-opening level recorded by electro-magnetic articulography (EMA) (Schönle, Gräbe, Wenig, Höhne, Schrader & Conrad, 1987; Perkell, Cohen, Svirsky, Matthies, Garabieta & Jackson, 1992) and the acoustic power level recorded by microphone. Also, during the experiment, the pressure control unit was placed in an acoustic housing to prevent the subject hearing it. This eliminated the chance that noise from the unit would signal the perturbation.

The articulatory movements for the utterance with the artificial palate were recorded on an electro-magnetic articulographic (EMA) system (Carstens AG-100) (Kaburagi & Honda, 1994). The receiver coils of the EMA were attached to the lower jaw, the lips, and at four positions along the tongue surface; namely, 8 mm posterior to the tongue apex and at intervals of approximately 14 mm (see Fig. 3). The seven positions on the articulators and two reference positions, one on the nose and the other on the upper teeth, were sampled at a rate of 250 Hz. The auditory signal was also recorded and sampled at 16 kHz.

2.2. *Experiments*

The repeated syllables /ʃa ʃa ⋯ ʃa/ and /tʃat tʃa ⋯ tʃa/ were used as speech samples. Each syllable was uttered 8 times with the leading syllable /iya/ within a test utterance. The fricative and affricate pairs, which have a different manner of articulation but the same place of articulation, were selected to investigate the compensatory articulation caused by the perturbation of the palate shape. The subjects were two male adults.

Four palatal conditions were examined in the experiment: steady-state deflated, steady-state inflated, inflation, and deflation. In the steady-state condition, the thickness of the artificial palate was kept constant during the utterance. In the inflation condition, the thickness of the deflated artificial palate was increased dynamically. On the contrary, in the deflation condition, the thickness of the inflated artificial palate was decreased dynamically. The inflated and deflated perturbations were given at the moment of the production of the open vowel /a/ in the leading syllable /iya/ so that the subject would be unaware of the perturbed palate shape until the tongue contacted the palate in the transition to the following fricative or fricative-stop. Then, the palatal perturbation continued to the end of the utterance. In order to minimize expectation of the perturbation by the subject, the perturbation was randomly given with an occurrence probability of 0.2.

Palatal perturbations were examined under normal and masked auditory-feedback conditions. For auditory-feedback masking, subjects were exposed to pink noise of about 88 dB SPL intensity through binaural earphones during the utterance. The intensity level was adjusted by the subjects so that the auditory feedback through both air and bone conduction was masked.

The experiment was carried out in four sessions. In the first session, the inflation palatal perturbation was examined in the normal auditory feedback condition. In this session, the artificial palate was normally deflated and randomly inflated. Accordingly, articulatory and acoustic data for the steady-state deflated and inflated palatal conditions were collected in this session. In the second session, the inflation palatal perturbation was examined in the masked auditory feedback condition. The masking noise was given to the subject during the entire session. In the third and fourth sessions, the deflation palatal perturbation was examined in normal and masked auditory feedback conditions in that order. In these sessions, the artificial palate was normally inflated and randomly deflated, providing articulatory and acoustic data for the steady-state inflated and deflation palatal conditions. For each session, there were 30 trials for each test utterance: 6 perturbed trials and 24 steady-state trials.

The subject was permitted to practice the utterance with artificial palate deflated before the first session and with it inflated before the third session. Practice was 5 min for each case. The perturbing effect of the artificial palate was perceptually noticed in first attempts at speaking with the steady-state deflated artificial palate, but the perturbing effect apparently declined after several minutes practice. The artificial palate, however, may still perturb normal speech production in the subsequent experimental session (Hamlet & Stone, 1978). EMA coils may also have a perturbing effect on fricative articulation (Weismer & Bunton, 1999). In spite of the potential perturbing effect of the artificial palate and EMA coils, compensatory articulation was examined by comparing articulatory and acoustic data obtained in the steady-state deflated/inflated condition with those obtained in the inflation/deflation perturbed condition.

A perceptual test was conducted on the speech errors that occurred in the perturbed trials. The six stimuli were randomly selected from each of the four conditions. The total number of stimuli was 192, i.e., six repetitions, four perturbation conditions, two auditory-masking conditions, two consonants, and two subjects. Four listeners participated in the perceptual test. The task was to identify each consonant in the repeated syllable using any consonant category of Japanese. When speech was not identified as a clear phonemic category, the listener was allowed not to judge it as a phonemic category. Because of the difficulty in identifying every consonant in the eight-syllable sequences after one listening, each listener was permitted to listen to each stimulus repeatedly until a confident judgment could be made. No *a priori* information about stimuli was provided to the listeners, although they were able to easily guess that only two target sequences were involved.

3. Results

3.1. *Perceptual analysis*

Mean error identification scores for every consonant in the repeated syllable for four palatal perturbation conditions and normal auditory-feedback conditions are shown in Table I. Each column shows the time course for the repeated syllable from the 1st to the 8th syllable.

In the inflation condition, speech errors were particularly apparent in the initial fricative /ʃ/ for both subjects, and the fricative was mostly misidentified as the affricate /tʃ/. The mean error identification score was 94% for subject A and 72% for subject B. Speech errors disappeared in the 2nd syllable for subject A and in the 3rd syllable for subject B, and remained absent up to the end of the utterance. In contrast, there were very few speech errors in the deflation condition. The errors were observed only in the first fricative for subject B.

For the affricate /tʃ/, speech errors were few for every perturbation condition. All speech errors for the affricate were observed for subject B, and there were no errors for subject A. For every perturbation condition, affricate /tʃ/ was mostly misidentified as the stop /t/, and most errors occurred on successive syllables in one particular trial by subject B. In the steady-state deflated condition, there was no speech error for either subject or either consonant. This suggests that any disturbing effects of the deflated artificial palate and EMA coils are marginal in producing correctly perceived speech compared to the disturbing effect in other perturbation conditions.

Mean error identification scores for the masked auditory-feedback condition are shown in Table II. When the auditory-feedback was masked, speech errors for the

Table I. Mean percentage error score for consonant identification in the repeated syllable for the normal auditory-feedback condition

		Syllable no.	1	2	3	4	5	6	7	8
Consonant	Perturbation	Subject								
/ʃ/	Steady-state deflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	0	0	0	0	0
	Inflation	Sub. A	94	28	0	0	0	0	0	0
		Sub. B	72	0	0	0	0	0	0	0
	Deflation	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	17	0	0	0	0	0	0	0
	Steady-state inflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	0	0	0	0	0
/tʃ/	Steady-state deflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	0	0	0	0	0
	Inflation	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	17	17	12	6	12	12	12	12
	Deflation	Sub. A	0	0	0	0	0	6	0	0
		Sub. B	3	12	0	0	0	0	6	6
	Steady-state inflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	6	6	17	17	6

Table II. Mean percentage error score for consonant identification in the repeated syllable for the masked auditory-feedback condition

		Syllable no.	1	2	3	4	5	6	7	8
Consonant	Perturbation	Subject								
/ʃ/	Steady-state deflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	0	0	0	0	0
	Inflation	Sub. A	100	78	44	44	33	56	45	67
		Sub. B	44	0	33	44	22	28	22	33
	Deflation	Sub. A	0	0	0	3	3	8	11	17
		Sub. B	0	0	0	0	0	3	0	0
	Steady-state inflated	Sub. A	78	56	28	78	78	67	67	67
		Sub. B	17	11	33	22	11	22	22	22
/tʃ/	Steady-state deflated	Sub. A	0	0	0	0	0	0	0	0
		Sub. B	0	0	0	6	6	6	6	17
	Inflation	Sub. A	17	0	17	28	17	28	37	37
		Sub. B	0	0	0	0	0	0	0	0
	Deflation	Sub. A	17	17	17	11	11	22	22	28
		Sub. B	0	0	0	0	0	0	0	0
	Steady-state inflated	Sub. A	0	17	28	39	28	50	56	67
		Sub. B	0	45	39	61	67	72	83	83

fricative /ʃ/ in the inflation condition mostly occurred in the initial syllable and some errors remained in the last syllables, in which /ʃ/ was mostly misidentified as /tʃ/. This tendency was commonly observed for both subjects, except that there was no error in the second syllable for subject B. The speech errors did not occur successively, but randomly over the repeated syllable.

For the deflation condition, there were fewer speech errors for the fricative than for the inflation condition. There were some errors, i.e., /ʃ/ was misidentified as /tʃ/, in the last several syllables for subject A. An increase of speech errors in the last syllables was commonly observed for subject A in all perturbation conditions. This may be related to the difficulty in producing long syllable repetition without auditory feedback. The affricate /tʃ/ was mostly misidentified as /t/ for the deflation and inflation conditions and as /ʃ/ for the deflation condition for subject A, while there were no speech errors for subject B in either condition.

In the steady-state inflated condition, speech errors for both consonants occurred frequently in every syllable position except for the initial syllable for the affricate. The fricative /ʃ/ was mostly misidentified as /tʃ/ and affricate /tʃ/ was mostly misidentified as /t/. The speech errors occurred successively in the repeated syllable for many trials. Errors were negligible for the steady-state deflated condition.

The perceptual test showed that the effects of auditory feedback masking on the production of correctly perceived speech were more apparent in the inflation and the steady-state inflated conditions; it was relatively small in the deflation and steady-state deflated conditions. This means that auditory feedback is more effective when speaking in an unusual (inflated palatal) condition than in the usual (deflated palatal) condition. Also, the auditory-feedback masking effect seemed to be somewhat subject-dependent.

3.2. *Acoustic analysis*

Acoustic analysis was performed on speech segments of fricative and affricate consonants in the repeated syllable. A 30-ms speech segment was taken so that the segment began at 16 ms after the articulatory timing at which the vertical position of the 2nd tongue coil T_2 showed the local maximum in the time trajectory. After low-pass filtering the speech signal with a cutoff frequency of 8 kHz, six LPC cepstral coefficients, which represent a gross feature of the spectrum, were obtained by a 12-pole LPC analysis. In order to test for the significant acoustical differences between the steady-state deflated condition and other conditions, multi-variable analysis of variance was performed on the six cepstral parameters of the acoustic segments for the initial and middle (fourth) syllable in the repeated syllable. Separate analyses were conducted for the two auditory-feedback conditions, because the increased voice effort when auditory-feedback was masked by pink noise caused a significant difference in speech loudness and voice source.

Fig. 2 shows a typical example of acoustic waveforms and spectra for the initial syllable segment of /ʃ/ in the steady-state deflated and the inflation conditions. The speech waveform in the inflation condition, which was perceived as /tʃ/, is characterized by a silence followed by a short burst and a frication component. Similar characteristics were commonly observed when speech errors occurred in the inflation condition. When the frication component of the perturbed speech was weak or the duration was short, the stop consonant /t/ was perceived. As shown in Fig. 2, the lack of the frication component resulted in reduced high-frequency components in the spectrum compared with the spectrum of the correctly perceived syllable segment.

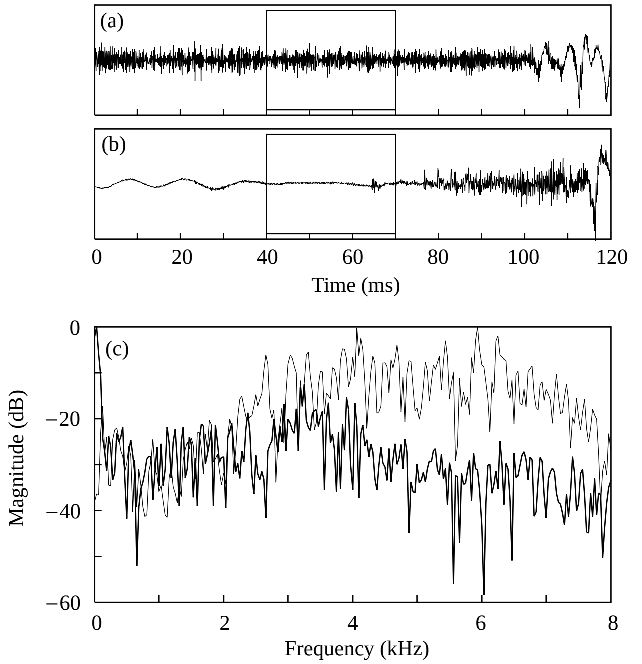


Figure 2. Speech waveforms and spectra in the initial syllable /ʃ/ for the steady-state deflated and the inflation conditions. (a) speech waveform in the steady-state deflated condition, (b) speech waveform in the inflation condition, and (c) spectra for the consonant segment in the steady-state deflated condition (thin line) and the inflation condition (thick line). The boxes on the waveform show the speech segment at which the acoustic analysis is performed.

The chi-squared values obtained from the multi-variable analysis, i.e., the logarithmic values of the determinant ratios of the total covariance matrix to the between-covariance matrix, are listed in Table III. Separate variance analysis was performed for each subject and the chi-squared values are averaged over two subjects. As seen in Table III, the relative magnitude of log likelihood ratio shows good correspondence with the speech error rate in the perceptual test. For the normal auditory-feedback condition, the acoustical difference of the fricative between the steady-state deflated and the inflation conditions is significant in the initial syllable but not significant in the middle syllable. The acoustical difference of the affricate is not significant in both the initial and the middle syllables. When auditory feedback is masked, the acoustical differences between the steady-state deflated and the inflation conditions are similar to those in the normal auditory condition. In the steady-state inflated condition, there are significant differences from the steady-state deflated condition for both consonants.

3.3. *Articulatory analysis*

In this section, we first compare the steady-state inflated, inflation and deflation conditions with the steady-state deflated condition under normal auditory feedback. Then, we examine the corresponding conditions under masked auditory feedback.

TABLE III. Log likelihood ratios for the multi-variable analysis on LPC cepstral parameters for normal and masked auditory-feedback conditions

			Initial	Mid
Consonant	Normal	Perturbation		
/ʃ/		Inflation	14.76*	4.81
		Deflation	5.54	3.97
		Steady-state inflated	7.53	8.08
	Masked	Perturbation		
		Inflation	13.79*	9.73
		Deflation	6.10	4.87
		Steady-state inflated	11.52	12.63*
/tʃ/	Normal	Perturbation		
		Inflation	7.73	5.35
		Deflation	9.65	3.67
		Steady-state inflated	6.60	8.40
	Masked	Perturbation		
		Inflation	10.44	12.03
		Deflation	10.11	12.73*
		Steady-state inflated	12.82*	14.13*

*denotes significant difference from the steady-state deflated condition ($p < 0.05$).

3.3.1. Steady-state inflated perturbation

In order to examine the speech movements for evidence of compensatory responses to the palatal perturbation, the articulator positions for the steady-state inflated condition were compared with those for the steady-state deflated condition in the normal auditory-feedback condition. Fig. 3 shows scatter plots of the articulatory positions in the midsagittal plane for consonants /ʃ/ and /tʃ/ for the two subjects. These samples were collected from all speaking trials with the steady-state deflated and steady-state inflated palates. Each position was sampled at the moment at which the vertical position of the tongue point T_2 showed the local maximum on the time trajectory. In the figures, the ellipses represent a range of two standard deviations along the principal components of variation. The palate shapes, which are illustrated in Fig. 3, were measured by tracing the EMA coil along the deflated and inflated palates.

When the palate shape was inflated in the alveolar region, the front two positions on the tongue T_1 and T_2 were lowered and the back two positions T_3 and T_4 moved backward for /ʃ/. The vertical shifts in T_1 and T_2 positions were significant for subjects A ($p < 0.01$ in F-test) and B ($p < 0.05$). The horizontal shifts of T_3 and T_4 were significant only for subject A ($p < 0.01$). It seemed that the lingual constriction in ʃ was formed near the tongue position T_2 for subject A, while it was formed at a more anterior position of the tongue surface between T_1 and T_2 for subject B. Similar compensation by the tongue was observed for /tʃ/, although the backward displacement of the tongue was small. The vertical shifts of T_1 and T_2 were significant only for subject A ($p < 0.01$). The jaw position slightly lowered for the inflated palate for both consonants, and the jaw lowering was significant for the subject A (in $p < 0.01$).

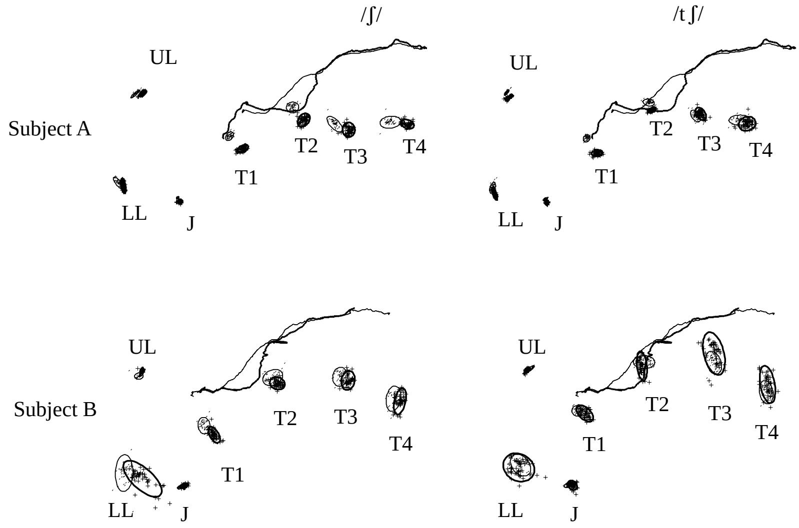


Figure 3. Plots of jaw (J), lips (UL, LL) and tongue (T1,T2,T3, T4) positions for repeated syllable /ʃa/ uttered in the steady-state deflated and the steady-state inflated conditions. The deflated (thin line) and inflated palate (thick line) shapes were obtained by tracing an EMA coil along the palate.

The time course of compensatory behavior for the steady-state inflated trial was compared with that for the steady-state deflated trial. The time trajectories of the horizontal and vertical positions of the jaw and the tongue for the repeated syllable /ʃa/ for subject A are shown in Fig. 4. The articulatory trajectories were averaged over six speaking trials with the line-up point set as the onset of the tongue constriction in the initial consonant. When the palate was steady-state inflated, the horizontal positions of the tongue shifted backward at the beginning of the utterance.

The vertical positions T_1 and T_2 were also lowered at the beginning of the utterance, except during the open vowel /a/ in the leading syllable /iya/. The tongue shows little evidence of overshooting an appropriate constriction width when moving towards the inflated palate in the transition from /a/ to /ʃ/. The compensatory trajectories were observed even in the first speaking trial with the steady-state inflated condition. This suggests that the compensatory articulation is well developed during the 5-min practice interval and that predictive motor control is employed to adapt tongue movement to the modified palate shape.

3.3.2. *Inflation perturbation*

The compensatory behaviors for the unexpected inflation condition were examined in terms of their dynamical responses in comparison with those for the steady-state deflated condition.

Fig. 5 shows the time trajectories of the jaw and the tongue for the repeated syllable /ʃa/ for subject A in the inflation and the steady-state deflated conditions. These trajectories were averaged over every six trials for each condition with the line-up point set as the onset of the tongue constriction in the initial consonant. The

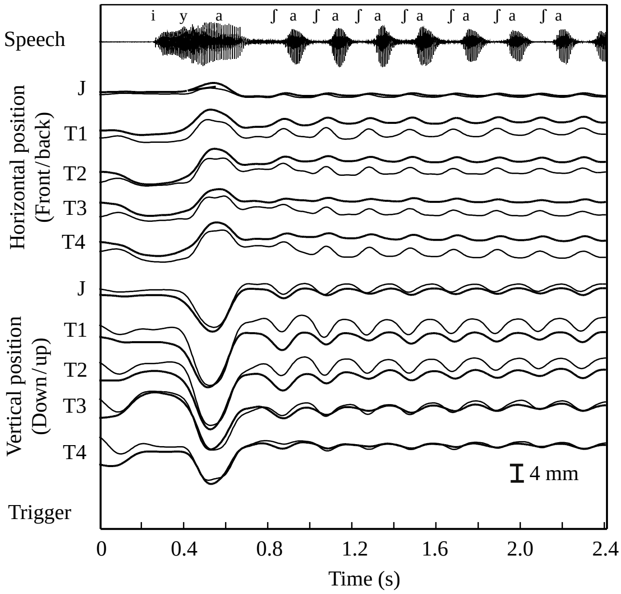


Figure 4. Articulatory trajectories of the jaw (J) and tongue (T_1, T_2, T_3 and T_4) in steady-state deflated condition (thin line) and the steady-state inflated condition (thick line) for repeated syllable /ʃa/ in normal auditory-feedback condition (Subject A). Larger value of the horizontal position indicates posterior position. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

speech signal for a typical trial in the steady-state deflated condition is displayed (it is thus not precisely synchronized with the movement traces). The onset of the trigger signal in the figure corresponds to the instant at which the perturbation was given. The vertical line indicates the onset of the tongue–palate contact in the transition. The contact onset, was determined by having the subjects gradually raise the tongue position toward articulation of /ʃ/ and sustain the articulation at the onset of the tongue–palate contact. Then, the contact onset was estimated from the vertical position of T_2 of the sustained articulation data.

The standard deviation of the time trajectories over six trials for both perturbation conditions are shown in Fig. 6; The amplitude scale is magnified to twice that of the articulatory trajectory. The standard deviation of tongue position averaged over the interval of the repeated syllable ranged from 0.58 to 0.90 mm for the steady-state deflated condition and 0.42–0.92 mm for the inflation condition.

Statistical analysis was performed to identify the moment at which the difference between the averaged articulatory trajectories in the steady-state deflated and the inflation conditions became significant ($p < 0.05$ in the F-test) after the inflated perturbation was given. The moment is indicated by an arrow on each time trajectory in Fig. 5. The time lag of compensatory response from the tongue–palate onset is listed in Table IV.

As seen in Fig. 5, the articulatory trajectories for the inflated trial were almost identical to those for the steady-state deflated one in the interval from the beginning of the utterance to the onset of articulating /ʃ/. Then, during production of the

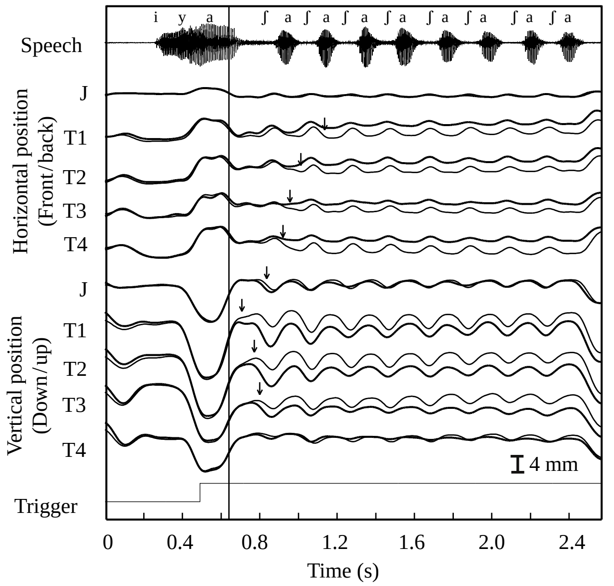


Figure 5. Articulatory trajectories of the jaw (J) and tongue (T_1, T_2, T_3 and T_4) in steady-state deflated condition (thin line) and the inflation condition (thick line) for repeated syllable /ʃa/ for subject A in normal auditory-feedback condition. Trigger signal shows perturbation onset and the vertical line shows onset of the tongue–palate contact. The arrow on the trajectory shows the time point at which the difference between the two trajectories becomes significant ($p<0.05$ in the *F*-test). Larger value of the horizontal position indicates posterior position. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

initial fricative, the vertical positions of the tongue were immediately lowered by about 4 mm for T_1 and T_2 and 2 mm for T_3 and remained at the lower positions during the following syllables. The horizontal positions of the tongue T_1–T_4 remained nearly the same as those for the steady-state deflated trial during the initial syllable and moved backward by about 3 mm during the second syllable. The time lag, measured with respect to tongue–palate onset, at which a significant departure of the tongue from the steady-state deflated trajectory occurred, was different among the vertical and horizontal directions as well as the tongue position. The time lag for the vertical position of T_1, T_2, and T_3 ranged from 72 to 164 ms, and the time lag for their horizontal position ranged from 284 to 500 ms (see Table IV). The differences emerged earlier in the vertical direction than in the horizontal direction, and earlier in the front tongue position for the vertical direction and in the back position for the horizontal direction. There was a significant difference in the jaw's vertical position in the transition from the initial consonant to the following vowel; it was not significant for the subsequent repeated syllables. The time lag for subject B ranged from 100 to 140 ms for the vertical and horizontal positions. It was longer in the vertical position of T_1 and was shorter in the horizontal positions than those for subject A, which resulted in small differences in the time lag between the vertical and the horizontal positions as well as between the front and the back positions.

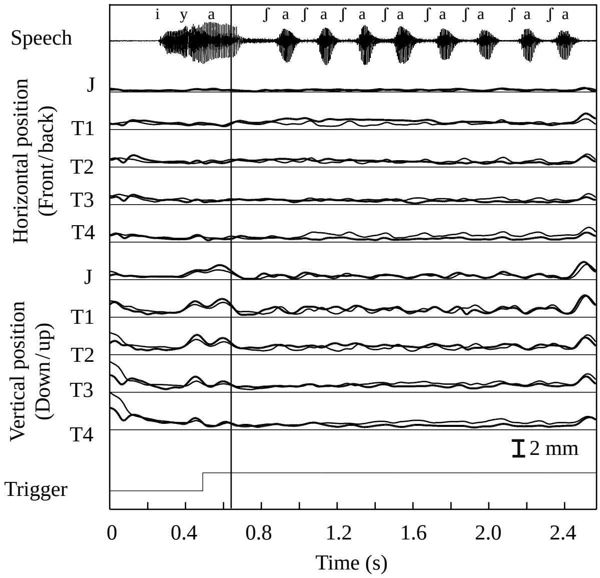


Figure 6. Standard deviation of articulatory trajectories over speaking trials in the steady-state deflated condition (thin line) and the inflation condition for repeated syllable /ʃa/ for subject A in normal auditory-feedback condition. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

TABLE IV. Time lag in ms of the compensatory response of the tongue and jaw from the tongue–palate onset for the repeated syllable /ʃa/ in inflation and deflation conditions for normal and masked auditory feedback

		Subject A				Subject B			
		Normal		Masked		Normal		Masked	
		Inflation	Deflation	Inflation	Deflation	Inflation	Deflation	Inflation	Deflation
Hor.	J	—	—	—	—	—	—	—	—
	T1	500	368	228	344	132	144	184	172
	T2	376	352	240	376	—	160	204	180
	T3	320	396	228	316	124	176	220	236
	T4	284	316	244	340	100	180	224	256
Ver.	J	200	—	320	—	—	—	—	—
	T1	72	76	76	76	124	130	176	160
	T2	136	108	144	108	128	140	180	188
	T3	164	148	192	140	124	176	100	180
	T4	—	—	—	—	—	—	—	—

— denotes no significant difference in trajectories between the steady-state deflated/inflated and the inflation/deflation conditions.

Speech errors in this condition mostly involved the initial fricative being perceived as an affricate (see Table I). This can be interpreted as overshoot of the tongue relative to the modified palate shape. The vertical position of T_2 at the initial fricative onset was slightly higher than that in the subsequent fricative articulation. The overshoot in the vertical position of T_2 caused vocal-tract closure, which resulted in the lack of a fricative component at the fricative onset, as seen in the acoustic waveform in Fig. 2. Adequate compensation by the tongue had been established by the second fricative and continued to the final syllable. The observed compensatory adjustments are consistent with the perceptual and acoustic results.

For the repeated syllable of affricate /tʃa/, the compensatory response was similar to that for the fricative /ʃa/, although the inflated displacement of the tongue was smaller. The affricate is relatively insusceptible to inflated perturbation, as the overshoot response of the tongue against the perturbed palate does not disturb the vocal tract closure.

3.3.3. Deflation perturbation

Movement patterns for the unexpected deflation condition were examined in comparison with those for the steady-state inflated condition. Fig. 7 shows the averaged articulatory trajectories for the repeated syllable /ʃa/ for subject A in the deflation and the steady-state inflated conditions. The arrows indicate the moment at which the difference between the trajectories for the two conditions becomes significant. As seen in Fig. 7, the articulatory trajectories for the deflated trial were almost identical to those for the steady-state inflated trial in the interval from the

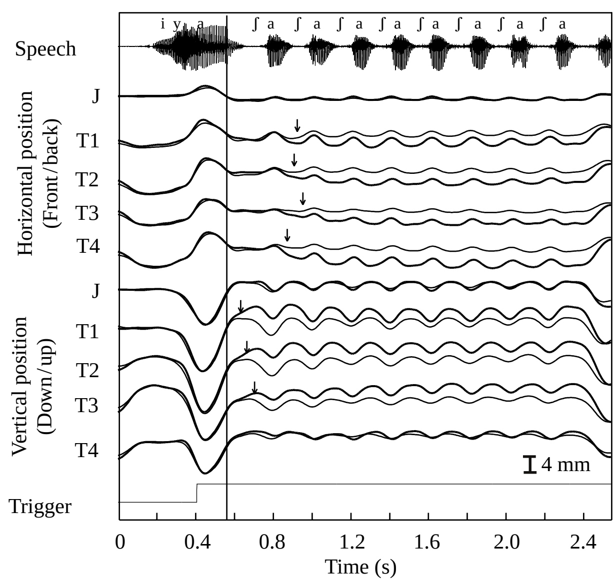


Figure 7. Articulatory trajectories of the jaw (J) and tongue (T_1, T_2, T_3 and T_4) in the steady-state inflated condition (thin line) and the deflation condition (thick line) for repeated syllable /ʃa/ for subject A in normal auditory-feedback condition. Trigger signal shows perturbation onset and the vertical line shows onset of the tongue–palate contact. The arrow on the trajectory shows the time point at which the difference between the two trajectories becomes significant ($p<0.05$ in the *F*-test). Larger value of the horizontal position indicates posterior position. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

beginning of the utterance to the onset of articulating /ʃ/. The perturbed trajectories of the vertical positions T_1–T_3 immediately restored to an adequate position relative to the deflated palate in the transition towards the initial fricative, and did not show any undershoot to the palate. The vertical positions of T_1, T_2, and T_3 were higher by about 4, 4 and 3 mm than those in the steady-state inflated condition. The vertical displacement shifts of tongue positions T_1 and T_2 were almost the same as the displacement change in the palatal height. The horizontal positions of the tongue T_1–T_4 moved forward by about 2–4 mm during the second syllable.

The time lag from tongue–palate onset of significant tongue-trajectory differences ranged from 76 to 148 ms for the vertical position of T_1, T_2 and T_3, and was about 350 ms for their horizontal positions for subject A. There was no significant difference in the jaw position (see Table IV). The changes in tongue position began earlier in the vertical direction than in the horizontal direction, and in the front tongue position for the vertical direction. The time lag for subject B ranged from 130 to 180 ms, and there were small differences between the vertical and horizontal positions.

When the trajectories in the deflation condition are compared with those in the steady-state deflated condition shown in Fig. 5, the trajectories in the two figures show no significant differences in any vertical positions of the tongue after the transition toward the initial fricative. The horizontal trajectories of the inflation condition reverted more gradually (over about a three-syllable interval) to the path followed in the steady-state deflated condition. A rapid restoration to the vertical position of the steady-state deflated trial presumably explains the low speech error rate in the perceptual test and the small acoustical differences. A few speech errors occurred in the initial syllable for subject B (see Table I). The response in these cases had involved a nonphonemic category, suggesting that adequate vocal tract constriction had not been achieved.

3.3.4. *Masked auditory-feedback effect*

The auditory-feedback effect on the immediate compensation and successive adaptation process in the inflated and deflated perturbed trials was examined. Fig. 8 shows the time trajectories of the jaw and the tongue for the repeated syllable /ʃa/ for subject A in the inflation and the steady-state deflated conditions when the auditory-feedback was masked. Fig. 9 shows the standard deviations of the time trajectories over six trials for both perturbation conditions. When auditory feedback was masked, the changes in the horizontal and the vertical displacements of the repeated syllables were larger than those in the normal auditory feedback condition (see Figs. 5 and 8). This suggests that articulation effort as well as phonation effort increases when auditory feedback is masked. The standard deviation of tongue position averaged over the interval of the repeated syllable ranged from 0.66 to 1.09 mm for the steady-state deflated condition and 0.56–1.10 mm for the inflation condition. The trial variances for the masked auditory feedback condition were slightly larger than those for the normal auditory feedback condition shown in Fig. 6.

The trajectories in the inflation condition are almost identical to those in the steady-state deflated condition until onset of tongue–palate contact. Changes in the vertical positions of the tongue T_1 to T_3 occurred immediately after the contact

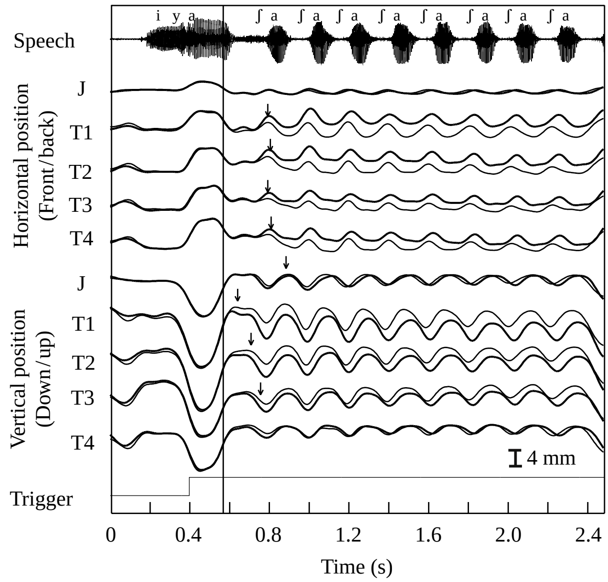


Figure 8. Articulatory trajectories of the jaw (J) and tongue (T_1, T_2, T_3 and T_4) in the steady-state deflated condition (thin line) and the inflation condition (thick line) for repeated syllable /ʃa/ for subject A in masked auditory-feedback condition. Trigger signal shows perturbation onset and the vertical line shows onset of the tongue–palate contact. The arrow on the trajectory shows the time point at which the difference between the two trajectories becomes significant ($p<0.05$ in the F-test). Larger value of the horizontal position indicates posterior position. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

onset even when the auditory feedback was masked. The time lag from the tongue–palate onset ranged from 76 to 192 ms for the vertical positions of T_1, T_2 and T_3, and was around 230 ms for their horizontal positions for subject A (see Table IV). The time lag for subject B ranged from 176 to 220 ms and there were small differences between the vertical and horizontal positions. The time lag for the vertical positions was similar to that for the normal auditory-feedback condition. Although the temporal compensatory profile was similar to that in the normal auditory-feedback condition, the lowering displacement of the tongue was 2–3 mm, which was smaller than that in the normal condition. The trajectories of the tongue often showed some fluctuation over trials, with random overshoot in the vertical position for the consonants in the repeated syllable, which caused vocal tract closure and a lack of frication component.

Fig. 10 shows the time trajectories for subject A in the deflation and the steady-state inflated conditions. The perturbed trajectories of vertical positions T_1–T_3 immediately restored to an adequate position relative to the deflated palate in the transition towards the initial fricative, and those of the horizontal positions also restored to adequate positions by the second syllable. These tendencies were very similar to those in the normal auditory-feedback condition (see Figs. 7 and 10). When auditory feedback was masked, the vertical positions of the tongue for the

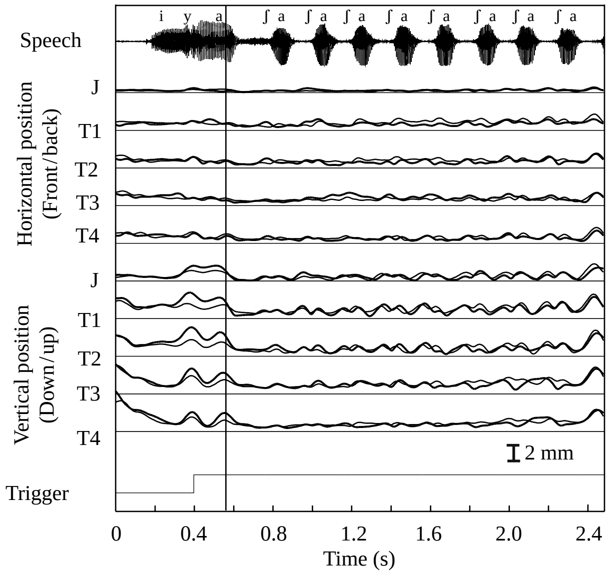


Figure 9. Standard deviation of articulatory trajectories over speaking trials in the steady-state deflated condition (thin line) and the inflation condition for repeated syllable /ʃa/ for subject A in masked auditory-feedback condition. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

fricative consonant in the steady-state inflated condition were higher than those in the normal condition, which frequently resulted in articulation error in producing the fricative consonant (see Table II).

4. Discussion

Previous studies using artificial palates revealed that immediate compensation for the palatalized fricative consonant occurred in the speaking trial just after the insertion of the appliances into the mouth, although small but significant acoustical and phoneme identification differences from the normal speaking condition still occurred (Flege *et al.*, 1988; McFarland & Baum, 1995). The steady-state inflated condition in our experiment corresponds to the post-adaptation condition in the previous studies, since the subject was permitted to practice the utterance with the inflated artificial palate for 5 min. In the post-adaptation condition, articulatory measurement data showed that the compensatory articulation was implemented by lowering the tongue blade position and shifting backward the entire tongue position. The palatalized fricative and affricate consonants showed no significant phoneme identification and acoustic differences between the steady-state deflated and inflation conditions, which is similar to previous results (McFarland & Baum, 1995).

Our inflation and deflation conditions provide evidence of how compensatory articulation develops in ongoing speech motor control when the palatal shape change cannot be predicted *a priori*. When the inflated palatal perturbation was

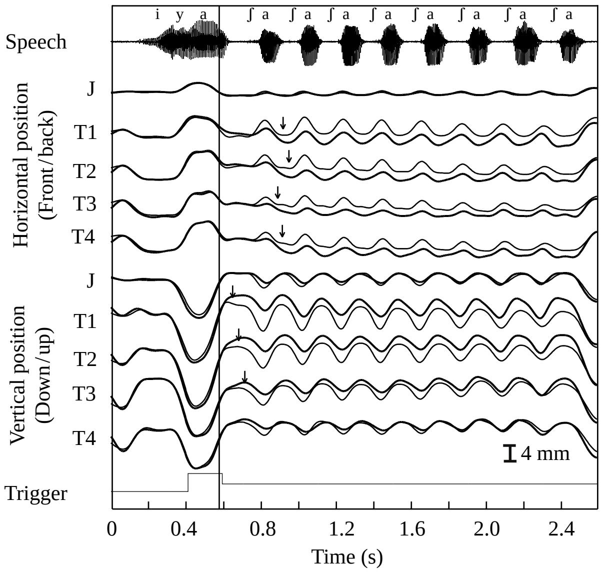


Figure 10. Articulatory trajectories of the jaw (J) and tongue (T_1, T_2, T_3 and T_4) in the steady-state inflated condition (thin line) and the deflation condition (thick line) for repeated syllable /ʃa/ for subject A in masked auditory-feedback condition. Trigger signal shows perturbation onset and the vertical line shows onset of the tongue–palate contact. The arrow on the trajectory shows the time point at which the difference between the two trajectories becomes significant ($p<0.05$ in the F-test). Larger value of the horizontal position indicates posterior position. The speech waveform is from a single typical utterance, and thus not precisely aligned with the averaged movement traces.

unexpectedly given just before a repeated fricative syllable utterance, the first syllable was mostly incorrectly produced due to overshoot of the tongue relative to the inflated palate shape, and immediate compensatory adjustments of the tongue, resulting in correctly perceived fricatives, became evident by the second syllable and continued to the final syllable. The response time lag from the tongue–palate contact onset was in a range of 72–164 ms for the vertical positions of the tongue and in the range of 284–500 ms for the horizontal positions for one subject, and was in a range of 100–140 ms for the vertical and the horizontal positions for other subjects. Similar rapid compensatory response was observed even when the auditory feedback was masked.

The latter fact in particular suggests that the compensatory response is driven by tactile rather than auditory information. The rather short time-lags (approx. 100 ms) for vertical differences might at first sight seem to reinforce this interpretation, since the time-lag for auditorily-driven responses can be expected to be well over 100 ms (Kawahara, 1994). However, it is quite conceivable that a large part of the short lag differences is due to the simple mechanical effect of the inflated palate preventing the tongue from following its unperturbed path.[1]

[1]The fact that these vertical lag values tend to increase from T1 to T3 is also not easy to explain on the basis of active adjustments, but could simply reflect the order in which the mechanical obstacle starts affecting different parts of the tongue.

The mechanical reaction effect, however, would not explain the lowering shift kept in the subsequent syllables and the relatively slow response in the horizontal movements shifting the tongue to a more posterior position. In order to clarify active vs. mechanical contributions to the rapid lowering response of the tongue, it would be necessary to examine the tongue muscle activities during the compensatory articulation, and perhaps to examine compensatory behavior under conditions of oral sensory deprivation.

Nonetheless, it is tempting to assume that it is precisely the sensory experience associated with being forced off the unperturbed trajectory in the first fricative that drives the compensatory response that has fully developed by the time the second fricative is articulated.

Thus, even if it is difficult to be sure at present whether the shortest lags we measured already represent part of the active response, it seems highly plausible that tactile information plays a central role in planning the response. For the time scales of interest here it is clearly an advantage that proprioceptive feedback from mechanoreceptors is much faster than auditory feedback. In a mechanical perturbation study, when a load was applied to the jaw during the closing gesture for production of /z/ in /bæz/, a quick response in the EMG activity of the tongue muscle (genioglossus posterior) was observed with a latency of 20–30 ms after the load onset (Kelso *et al.*, 1984). In another study, the autogenic reflex response in EMG to mechanical stimulation of the tongue blade was 30–40 ms (Weber & Smith, 1987). The time lag of the compensatory response we observed for the tongue blade is longer than those latencies, but may be compatible with them if the additional delay required from EMG activity to displacement is considered.

When the auditory-feedback was masked, incorrect production often occurred for the inflation and the steady-state inflated perturbations, whereas there were few errors for the deflation and the steady-state deflated conditions. One possible explanation is that acquisition of motor control adjustments necessary to cope with the modified artificial palate is still incomplete after the 5 min practice. A previous study revealed that compensation for the fricative to the modified artificial palate improved with practice over weeks (Hamlet & Stone, 1978). With insufficient acquisition of the motor control appropriate to an unusual speaking environment, auditory feedback would be effectively used in finely adjusting ongoing articulation by comparing the produced sound with the acoustic goal.

When the artificial palate was unexpectedly deflated from the inflated position, changes in movement pattern occurred at a time lag that was similar to that in the inflated perturbation condition. Movements immediately reverted to those in the steady-state deflated conditions, and showed no undershoot to the deflated palate. This fact can be explained by the exploitation of biomechanical saturation effects in motor control (Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico,Wozniak & Guiod, 1997). When one articulator contacts another, the position of the articulator saturates for a continuous change in the muscle activation level. Saturation effects are often found in tongue–palate contact in fricative as well as in stop consonants. Palatographic data on alveolar and palatal fricatives show that stable constriction could be achieved by pressing the sides of the tongue having a central groove against the hard palate (Hamlet & Stone, 1978). When the palate shape is unexpectedly returned to the deflated position, the subject expects the inflated palate until the perturbation onset. If motor control aims for the position control of the tongue relative to the inflated palate, some changes in the trajectories caused by the compensatory adjustment would be found when the palate is suddenly deflated. The displacement and velocity of the perturbed trajectory of the vertical position of the tongue blade for /ʃ/ showed, however, no significant difference from that in the steady-state deflated trial (un-inflated palate). One possible explanation is that motor control utilizes saturation effects by the tongue–palate contact rather than precise positioning relative to the modified palate shape. If the fricative is articulated by the tongue with a deeper central groove in the inflated palate condition, saturation effects would lead to the correct articulation without employing any compensation. When the palate is inflated from the deflated position, the inflated palate disturbs the central groove of the tongue even if the saturation effect is employed. This leads to the incorrect production of /ʃ/ as shown in the perceptual test.

5. Summary

Compensatory articulatory responses to unexpectedly inflated and deflated palatal perturbations were examined in normal and masked auditory-feedback conditions. The experimental results showed that the compensatory articulation occurred in ongoing speech production with a response time lag of 76–500 ms for tongue positions. The rapid compensatory response found in the inflated palatal perturbation suggests that tactile feedback is primarily used. The results also show that auditory feedback is used in ongoing speech production when adaptation to the unusual structural perturbation is insufficient. Evidence for closed-loop sensory feedback in ongoing speech production could be made even more apparent by also examining the muscle activation levels related to unexpected structural perturbations of the kind presented here.

References

Abbs, J. H. & Gracco, V. L. (1983) Sensorimotor actions in the control of multi-movement speech gestures, *TINS*, **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,** 705–723.

Baum, S. R., McFarland, D. H. & Diab, M. (1996) Compensation to articulatory perturbation: perceptual data, *Journal of the Acoustical Society of America*, **99,** 3791–3794.

Flege, J. E., Fletcher, S. G. & Homiedan, A. (1988) Compensating for a bite block in /s/ and /t/ production: palatographic, acoustic, and perceptual data, *Journal of the Acoustical Society of America*, **83,** 212–228.

Folkins, J. & 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. & 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,** 1225–1233.

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

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,** 418–432.

Hamlet, S. L. & Stone, M. (1976) Compensatory vowel characteristics resulting from the presence of different types of experimental dental prostheses, *Journal of Phonetics*, **4,** 199–218.

Hamlet, S. L. & Stone, M. (1978) Compensatory alveolar consonant production induced by wearing a dental prosthesis, *Journal of Phonetics*, **6,** 227–248.

Hamlet, S. L., Cullison, B. L. & Stone, M. (1979) Physiological control of sibilant duration: insights afforded by speech compensation to dental prostheses, *Journal of the Acoustical Society of America*, **65,** 1276–1285.

Honda, M. & Kaburagi, T. (2000) Speech compensation to dynamical structural perturbation of the palate shape. In *Proceedings of 5th Seminar of Speech Production* & *CREST Workshop on Models of Speech Production*, Kloster Seeon, Germany, pp. 21–24.

Ito, T., Gomi, H. & Honda, M. (2000) Task dependent jaw-lip coordination examined by jaw perturbation during bilabial-consonant utterances. In *Proceedings of 5th Seminar on Speech Production & CREST Workshop on Models of Speech Production*, Kloster Seeon, Germany, pp. 41–44.

Kaburagi, T. & Honda, M. (1994) Determination of sagittal tongue shape from the positions of points on the tongue surface, *Journal of the Acoustical Society of America*, **96,** 1356–1366;

Kawahara, H. (1994) Interaction between speech production and perception under auditory feedback perturbations on fundamental frequencies, *Journal of the Acoustical Society of Japan*, **15,** 201–202.

Kelso, J. A. S., Tuller, B., Vatikiotis-Bateson, E. & Fowler, C. A. (1984) Functionally specific articulatory adaptation to jaw perturbations during speech: evidence for coordinative structures, *Journal of Experimental Psychology*, **10,** 812–832.

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.

McFarland, D. H. & Baum, S. R. (1995) Incomplete compensation to articulatory perturbation, *Journal of the Acoustical Society of America*, **97,** 1865–1873.

McFarland, D. H., Baum, S. R. & Chabot, C. (1996) Speech compensation to structural modifications of the oral cavity, *Journal of the Acoustical Society of America*, **100,** 1093–1104.

Perkell, J. S., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I. & Jackson, M. T. T. (1992) Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements, *Journal of the Acoustical Society of America*, **92,** 3078–3096.

Perkell, J. S., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J. & Guiod, P. (1997) Speech motor control: acoustic goals, saturation effects, auditory feedback and internal models, *Speech Communication*, **22,** 227–250.

Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J. & Conrad, B. (1987) Electromagnetic articulography: use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract, *Brain and Language*, **31,** 26–35.

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,** 70–79.

Weismer, G. & Bunton, K. (1999) Influence of pellet mancers on speech production behavior: Acoustical and perceptual measures. *Journal of the Acoustical Society of America*, **105,** 2882–2894.