

Lip kinematics in long and short stop and fricative consonants

Anders Löfqvist^{b)}

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511 and Department of Logopedics, Phoniatrics and Audiology Lund University, Lund, Sweden

(Received 23 February 2003; revised 22 October 2004; accepted 29 October 2004)

This paper examines lip and jaw kinematics in the production of labial stop and fricative consonants where the duration of the oral closure/constriction is varied for linguistic purposes. The subjects were speakers of Japanese and Swedish, two languages that have a contrast between short and long consonants. Lip and jaw movements were recorded using a magnetometer system. Based on earlier work showing that the lips are moving at a high velocity at the oral closure, it was hypothesized that speakers could control closure/constriction duration by varying the position of a virtual target for the lips. According to this hypothesis, the peak vertical position of the lower lip during the oral closure/constriction should be higher for the long than for the short consonants. This would result in the lips staying in contact for a longer period. The results show that this is the case for the Japanese subjects and one Swedish subject who produced non-overlapping distributions of closure/constriction duration for the two categories. However, the peak velocity of the lower lip raising movement did not differ between the two categories. Thus if the lip movements in speech are controlled by specifying a virtual target, that control must involve variations in both the position and the timing of the target. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1840531]

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

Pages: 858–878

I. INTRODUCTION

This paper examines lip and jaw kinematics in the control of lip closure/constriction duration in labial stop and fricative consonants, where the duration of the oral closure/constriction is varied for linguistic purposes. Based on earlier work on lip movements (Löfqvist and Gracco, 1997), it was hypothesized that speakers could vary the duration of the oral closure by shifting the position of a virtual target for the lips. For example, the target for the lower lip would be at a higher vertical position for a long than for a short consonant. Such a strategy would make the lips stay in contact for a longer time. Preliminary evidence for such a strategy for controlling closure duration was provided for one Swedish speaker by Löfqvist (2000). In addition, Löfqvist and Gracco (1997) found a positive correlation between closure duration and the peak vertical position of the lower lip during the oral closure, but the range of closure durations was not very large in their study.

The results presented by Löfqvist and Gracco (1997) showed that the lips were moving at close to their peak velocities at the instant of labial closure. The high velocity at the impact resulted in tissue compression making the airtight seal for the stop consonant. In addition, mechanical interactions between the lips were observed, with the lower lip pushing the upper lip upward due to its higher velocity. These results were compatible with the idea of a virtual target for the lips that would have them move beyond each other. Such a control strategy would ensure that the lips will make a closure irrespective of variations in their onset positions.

Although there are several acoustic studies of long and short consonants, e.g., Finnish (Lehtonen, 1970), Bengali and Turkish (Lahiri and Hankamer (1988), Swedish (Elert, 1964), and Japanese (Beckman, 1982; Han, 1994), not much is known about possible differences in the articulation of long and short consonants. Electropalatographic studies of Italian stops have shown that the amount of tongue palate contact is larger for geminate than for single stops, and also that there is a general increase in the extent of tongue-palate contact with increasing closure duration (Farnetani, 1990); similar results for American English have been presented by Byrd (1995). An x-ray study of French consonants by Vaxelaire (1995) suggested that the area of tongue palate contact was larger for the long (abutted) stops than for the short ones. Lehiste *et al.* (1973) recorded the activity of the upper lip muscles in a subject producing long and short labial consonants in Estonian, and found that the long cognate was generally produced with two successive peaks in the EMG signal. Dunn (1993) studied lip movements in long and short consonants in Italian and Finnish. Using a derived measure of lip aperture, she found that the lips stayed in contact for a longer period of time for the long sounds; however, the movement kinematics were not consistently different for the long and short consonants. A related study by Smith (1995) examined lip and tongue movements in single and geminate consonants in Japanese and Italian. Of particular interest for the present study, Smith found that the closing movements of the lips were slower for the geminate than for the single consonant.

Although Löfqvist and Gracco (1997) proposed the idea of a virtual target for lip movements, they also noted that it might be applicable to other articulators as well, since whenever two articulators make contact, at least one of them consists of soft tissue. Moreover, recent work on tongue move-

^{a)}Parts of this paper were presented at the First Pan-American/Iberian Meeting on Acoustics, Cancun, Mexico, 2–6 December, 2002.

^{b)}Electronic mail: lofqvist@haskins.yale.edu

ments in speech has explored a similar idea. In this case, the argument has been made that the virtual target for the tongue in the production of a velar stop consonant is above the hard palate in the nasal cavity. Empirical studies of tongue movement kinematics, showing that the tongue can be moving at a high velocity at the instant of oral closure (Fuchs *et al.*, 2001; Löfqvist and Gracco, 2002), and modeling work using a virtual target (Perrier *et al.*, 2003) have shown some support for this idea. It should be pointed out that this study does not address the broader issue of the potential role of virtual targets in speech motor control. The current experiment was only designed to examine a specific hypothesis about virtual targets for the lips.

The hypothesis to be explored in the present study is that a temporal property in speech, i.e., the duration of a labial closure/constriction for stops and fricatives, is governed by varying a spatial control parameter, i.e., lip target position. Although the coordinate system for speech motor control is not known, one proposal views the proper control regime as tract variables, e.g., lip aperture which is implemented jointly by movements of the upper and lower lips and the jaw (e.g., Saltzman and Munhall, 1989). Thus, it is possible that the kinematics of both the upper lip, the lower lip, and the jaw will differ for long and short consonants. Due to the mechanical interactions between the lips discussed above and further elaborated below, measurements will primarily be made of the lower lip movements. However, the peak velocity of the closing movement of the upper lip and the closing movement of a derived measure of lip aperture will be also examined, since they occur before the lips meet. This focus on the lower lip is methodological and should not be taken as evidence that the control is limited to the lower lip.

In its simplest form, the hypothesis about a change in virtual position only considers the movement displacement and assumes that other aspects of the movement, e.g., timing, remain unchanged. For epistemological reasons, such a simple and restricted hypothesis is most easily refuted and thus helpful for a further understanding of lip aperture control in speech. In terms of motor control, the theoretical framework is similar to the model proposed by Gottlieb *et al.* (1989), where the magnitude and the duration of the underlying force pulse can be controlled. In the present case, the assumption is that only the magnitude of the pulse is changed to produce a long consonant.

If the hypothesis that speakers vary the position of a virtual target for the lower lip to produce long and short consonants is correct, the following predictions can be made of the lip movements in long and short consonants:

- (1) The peak position of the lower lip is higher for a long than for a short consonant. If this is true, the following additional predictions can also be made:
- (2) There is a positive correlation between the vertical positions of the upper and lower lips at the point in time when the lower lip reaches its highest vertical position. This follows from the previously observed mechanical interactions between the two lips with the lower lip pushing the upper lip upward (Löfqvist and Gracco, 1997). With the lower lip reaching a higher vertical po-

sition for a long consonant, it would push the upper lip further upward. Hence, the upper lip position at this point in time is higher for a long than for a short consonant.

- (3) A long consonant is produced with a larger lower lip closing displacement than a short consonant. Furthermore, this is due to a difference in the vertical end position of the lower lip and not to a difference in its starting position.
- (4) A long consonant is produced with a higher peak closing velocity of the lower lip. This follows from the strong correlation between movement displacement and peak velocity that has been observed in both speech and non-speech movements (e.g., Cooke, 1980; Ostry *et al.* 1983; Kelso *et al.*, 1985; Vatikiotis-Bateson and Kelso, 1993; Hertrich and Ackermann, 1997; Löfqvist and Gracco, 1997). One exception to this pattern can occur when a movement of one articulator is truncated by a following gesture, so that its movement displacement is decreased (e.g., Munhall *et al.*, 1992; Harrington *et al.*, 1995; Byrd *et al.*, 2000). Interestingly, a different kind of truncation is shown in the lip data presented by Löfqvist and Gracco (1997), when the lowering movement of the upper lip is checked by the rising lower lip. The same phenomenon is observed in the present study.

We should note from the outset that any observed differences will be in the order of a few mm, since the lip movements in labial consonants are usually 1 cm or less, depending on the vowel context (e.g., Löfqvist and Gracco, 1997). To test the hypothesis of variations of the position of a virtual to control closure/constriction duration, speakers of Japanese and Swedish were studied. Both these languages have a contrast between long and short consonants, although the structure of the length contrast differs in the two languages. In Swedish, there is a durational relationship between a vowel and a following consonant. That is, a short vowel is followed by a long consonant, while a long vowel is followed by a short consonant. In addition, there may be differences in vowel quality between the long and short vowels in Swedish (cf., Hadding-Koch and Abramson, 1964; Fant, 1973). Most likely due to these additional components of the length contrast for Swedish consonants, the difference in duration between a long and short consonant is often quite small, and there may be a substantial overlap between the distributions of the closure durations of long and short consonants (Elert, 1964). In Japanese, the difference in closure duration is much larger than in Swedish, with the long ones being about twice as long as the short ones (e.g., Beckman, 1982; Han, 1994).

II. METHOD

A. Subjects

Four female native speakers of Japanese and two native speakers of Swedish, one male and one female, served as subjects. They reported no speech, language, or hearing problems. They were naive as to the purpose of the study. Before participating in the recording, they read and signed a

consent form. (The experimental protocol was approved by the IRB at the Yale University School of Medicine.)

B. Linguistic material

The linguistic material consisted of Japanese and Swedish words, occurring in a short carrier sentence. To keep the phonetic context of the labial consonants as similar as possible, the consonant was placed between the open vowel /a/, whenever possible. The specific reason for keeping the context similar was the predicted larger lower lip vertical displacement for the long than for the short consonants. Unless the position of the lower lip at the start of the closing movement was very similar for the long and short consonants, the observed result might simply be due to differences between the onset positions for the long and short consonants. As a consequence, differences in peak velocity might also be due to differences in displacement related to different onset positions. As will be discussed below, the attempt to control the onset position of the lower lip was not always successful, however.

In Japanese, the following words were used: "napa," "nappa," "sama," "samma," "tofuru," and "daffuru." The labial consonant in the two last words is a bilabial fricative / Φ /. The Swedish subjects produced the following words: "rapa," "rappa," "Saba," "sabba," "rama," "ramma," "slafa," and "haffa." The linguistic material was organized into randomized lists and presented to the subjects in Swedish and Japanese writing, with the words occurring in a short frame sentence with sentence stress on the test word. Fifty repetitions of each word were recorded.

C. Movement recording

The movements of the lips and jaw were recorded using a three-transmitter magnetometer system (Perkell *et al.*, 1992); when proper care is taken during the calibration, the spatial resolution of the system is in the order of 0.5 mm. Receivers were placed on the vermilion border of the upper and lower lip, and on the lower incisors at the gum line. Two additional receivers placed on the nose and the upper incisors were used for the correction of head movements. All receivers were attached using Isodent, a dental adhesive. Care was taken during each receiver placement to ensure that it was positioned at the midline with its long axis perpendicular to the sagittal plane. Two receivers attached to a plate were used to record the occlusal plane by having the subject bite on the plate for a brief interval during the recording. All data were subsequently corrected for head movements and rotated to bring the occlusal plane into coincidence with the *x* axis. This rotation was performed to obtain a uniform coordinate system for all subjects (cf., Westbury, 1994); the origin of the coordinate system was the receiver placed on the upper incisors.

The articulatory movement signals (induced voltages from the receiver coils) were sampled at 500 Hz after low-pass filtering at 200 Hz. The resolution for all signals was 12 bits. After voltage-to-distance conversion, the movement signals were low-pass filtered using a 25-point triangular window with a 3-dB cutoff at 14 Hz; this was done forwards and

backwards to maintain phase. A measure of lip aperture was obtained by calculating the difference between the upper and lower lip vertical receivers. To obtain instantaneous velocity, the first derivative of the position signals was calculated using a three-point central difference algorithm. The velocity signals were smoothed using the same triangular window. Movement onsets and offsets were defined algorithmically at zero-crossings in the velocity signal. The peak movement velocity was also labeled algorithmically. Signal averages were obtained using the onset of oral closure, defined in the acoustic signal, as the line-up point, and with a temporal window extending 100 ms before and 150 ms after the line-up point; these averages were only used for visualization purposes and all measurements were based on the individual tokens. All the signal processing was made using the Haskins Analysis Display and Experiment System (HADES) (Rubin and Löfqvist, 1996).

The acoustic signal was preemphasized, low-pass filtered at 4.5 kHz, and sampled at 10 kHz. The onset and release of the oral closure for the stops and nasals were identified in waveform and spectrogram displays of the acoustic signal. The onset of the closure was identified by the decrease in the signal amplitude. For the fricatives, the cessation and reappearance of voicing were used to identify the constriction phase.

The duration of the oral closure/constriction was measured in the acoustic signal. The vertical receiver positions of the lips and the jaw were measured at the onset and offset of the oral closing movement. The peak vertical velocity of the closing movement was also measured. Finally, the vertical position of the upper lip was measured at the point in time when the lower lip reached its peak vertical position. The displacement of the closing movement was calculated as the difference between the onset and offset vertical positions. The measurements of the upper lip closing movement were problematic due to interactions between the upper and lower lips. This is illustrated in Fig. 1, showing a short (left panel) and a long (right panel) nasal consonant produced by Japanese subject NY. The shaded areas in the left and right panels show the lip movements during the nasal consonant. In the production of the short nasal in "sama" (left panel), there is only one zero-crossing in the upper lip velocity signal during the consonant. However, in the production of the long nasal in "samma" (right panel), there are three zero-crossings in the upper lip velocity signal during the consonant. They occur because the lower lip checks the descent of the upper lip and pushes it upward, as can be seen in the upper lip position signal [see Löfqvist and Gracco (1997) for a detailed analysis of these patterns]. Thus, the upper lip displacement could not be reliably measured, in particular for the long consonants where most of these interactions occurred. The peak velocity of the upper lip closing movement was measured, however, since it occurs before the lips meet; the arrows in Fig. 1 show the peak velocities of the upper and lower lip closing movements. Similarly, the peak closing velocity of the lip aperture signal was measured.

T-tests were used to assess differences between the long and short consonants for each subject. With 50 repetitions of each word, the tests have 98 degrees of freedom. To adjust

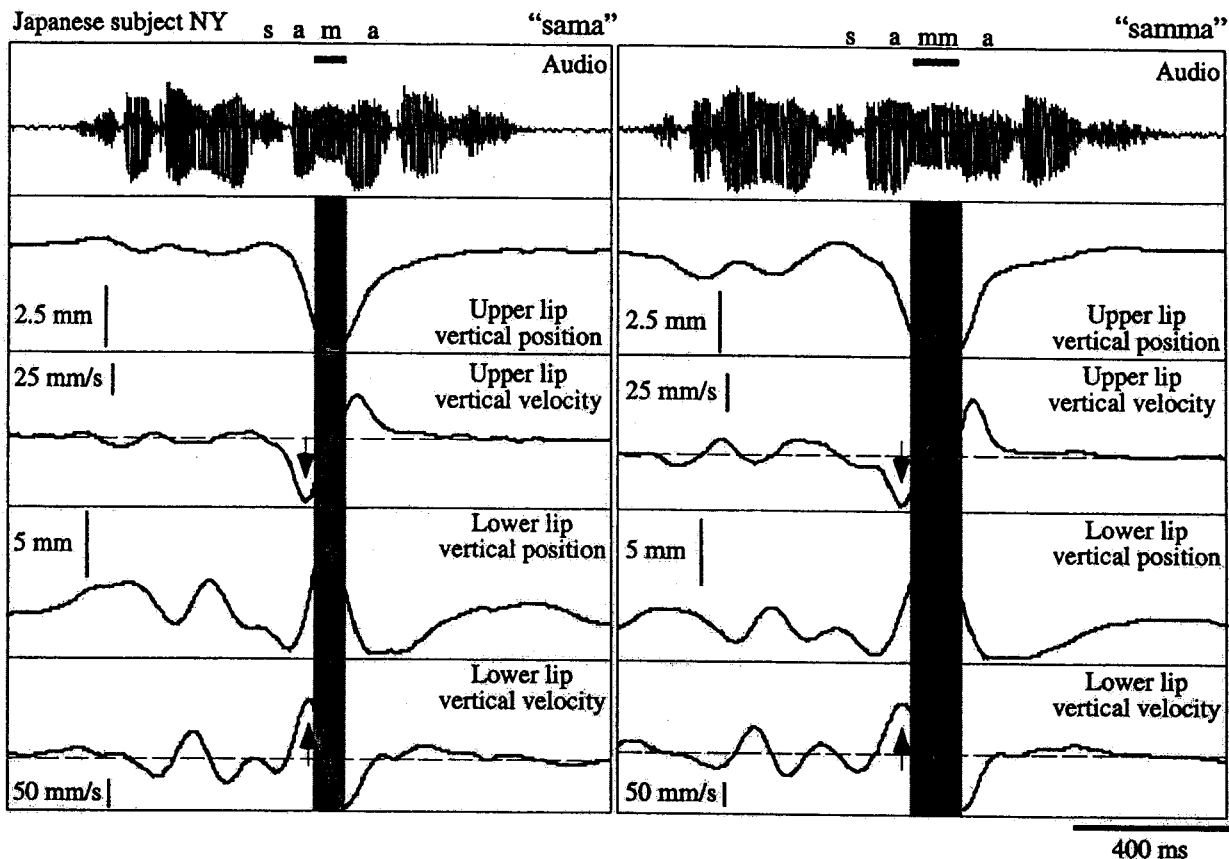


FIG. 1. A short (left panel) and a long (right panel) nasal consonant produced by Japanese subject NY. The shaded area in each panel shows the oral closure for the consonant. The arrows point at the peak velocities of the upper and lower lip closing movements.

for an elevated type I error rate due to multiple comparisons, a conservative α -level of 0.001 was adopted.

III. RESULTS

A. Closure duration

Figure 2 presents the duration of the oral closure/constriction for the Swedish subjects. In Fig. 2, there is considerable overlap between the closure durations for the long and short consonants, in particular for subject AG. The statistical analysis revealed that subject AG did not reliably distinguish the closure durations of the long and short voiceless consonant /p/, $t(98)=2.23$, ns, but did so for the voiced stop /b/, $t(98)=5.01$, $p<0.001$, the nasal /m/, $t(98)=7.87$, $p<0.001$, and the fricative /f/, $t(98)=9.98$, $p<0.001$. In contrast, Swedish subject NR reliably used different closure/constriction durations for all the consonants, $t(98)=15.57$, 7.79, 7.52, and 20.21, for /p,b,m,f/, respectively, with $p<0.001$ in all cases. Figure 3 plots the same results for the Japanese subjects. Here, and in contrast to the Swedish results, there is no overlap between the closure/constriction durations for the long and short consonants for any of the subjects or consonants. The closure/constriction duration of the long consonant is about twice as long as that for the short ones. The statistical analysis showed that all the Japanese subjects produced highly significantly different closure/

constriction durations for the long and short consonants, $t(98)=41.44$, 24.99, 29.38, and 34.18 for the labial stop /p/ for subjects HI, Y, NY, and SS, respectively. The corresponding t values for the labial nasal /m/ were 38.12, 24.89, 29.84, and 32.35 for subjects HI, MY, NY, and SS, respectively, and 34.79, 42.12, 29.7, and 31.2 for the bilabial fricative for subjects HI, MY, NY, and SS, respectively.

B. Movement kinematics

Figure 4 shows signal averages (aligned to the beginning of the acoustic closure) of the lip and jaw movements for the four Japanese subjects, since many of the subsequent analyses will be focused on these subjects. The arrows show the direction of the movement (they have been left out for the jaw since its movement was very small). Note that the window used for signal averaging includes movements during the vowels before and after the short consonants. Since the lip receivers were placed at the vermillion border of the upper and lower lips, the lip receivers are about 0.5–1.5 cm apart vertically when the lips are closed. With one exception, the lower lip of Japanese subject HI in Fig. 4(a), all the lip movements predominantly occur in the vertical dimension. Thus, the focus on the vertical movement dimension is justified. A comparison of the lower lip movement patterns for the long and short consonants in Fig. 4 shows that the lower lip tends to reach a higher vertical position for the long

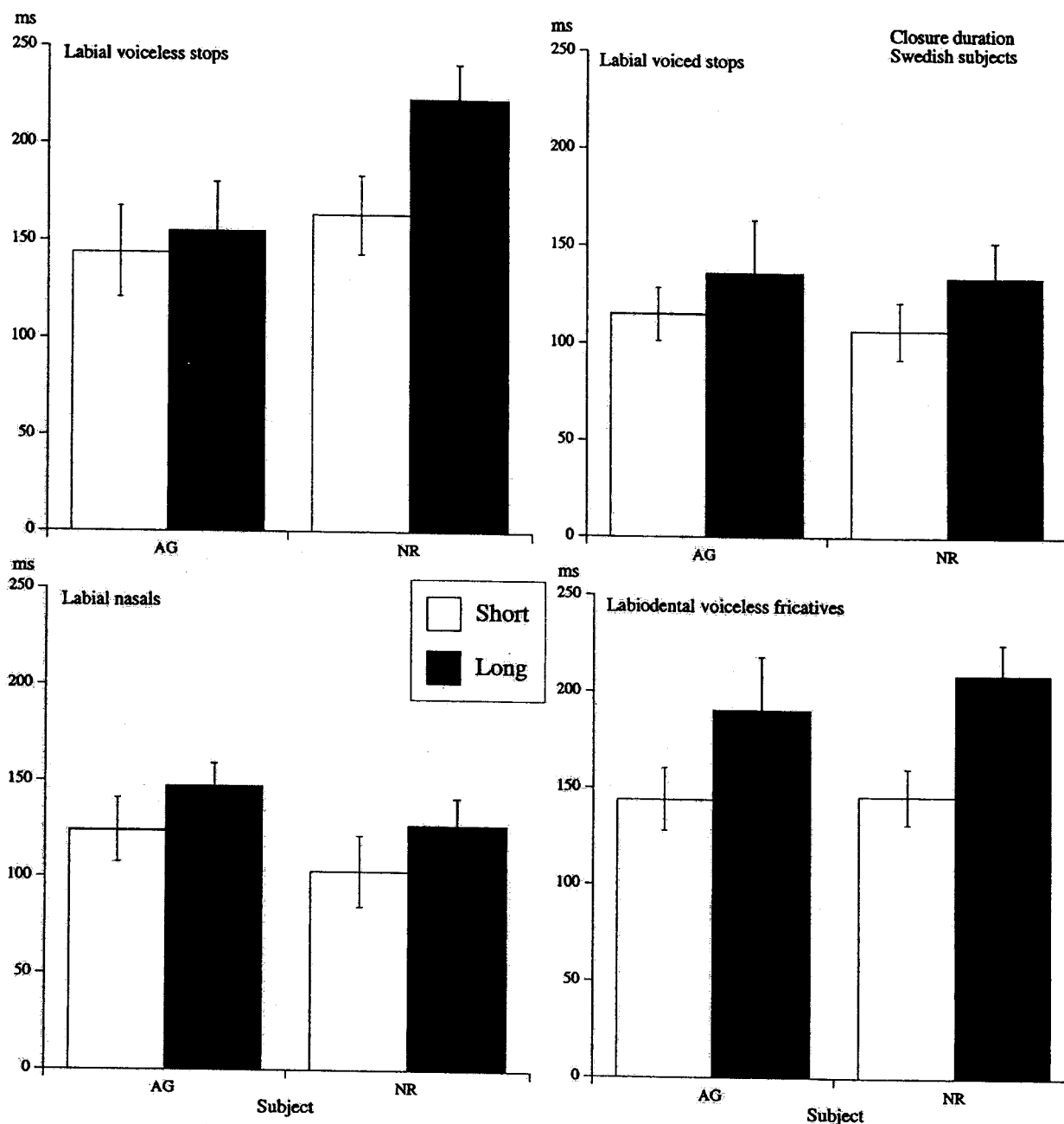


FIG. 2. Closure duration (mean and standard deviation) for the Swedish subjects.

(dashed lines) than for the short ones (solid lines) for all subjects, although the magnitude of the difference varies between subjects. The difference is in the order of a few mm, in particular for Japanese subjects MY, NY, and SS. It is also evident in Fig. 4 that the lip movements differ for the stop and the nasal, on the one hand, and the fricative, on the other; the fricative is bilabial. In particular, the lower lip reaches a lower vertical position for the fricative than for the stop and the nasal, while the upper lip has a higher position during the fricative than during the stop and the nasal.

The main focus of this study is on the kinematics of the lower lip. An analysis of jaw movements revealed no con-

stant differences between the long and short consonants; the jaw movements were very similar. As mentioned above, the analysis of upper lip movements is complicated by interactions between the upper and lower lips. That is, the upper lip lowering movement was often checked by the lower lip raising movement. These interactions also depend on the positions of the lip receivers (cf., Löfqvist and Gracco, 1997, for a more detailed analyses of such interactions). Figure 5 presents the results of the peak upper lip lowering velocity for the Swedish and Japanese subjects' productions of the short and long stop and nasal consonants; for the fricatives, the upper lip movement was very small and often less than 1

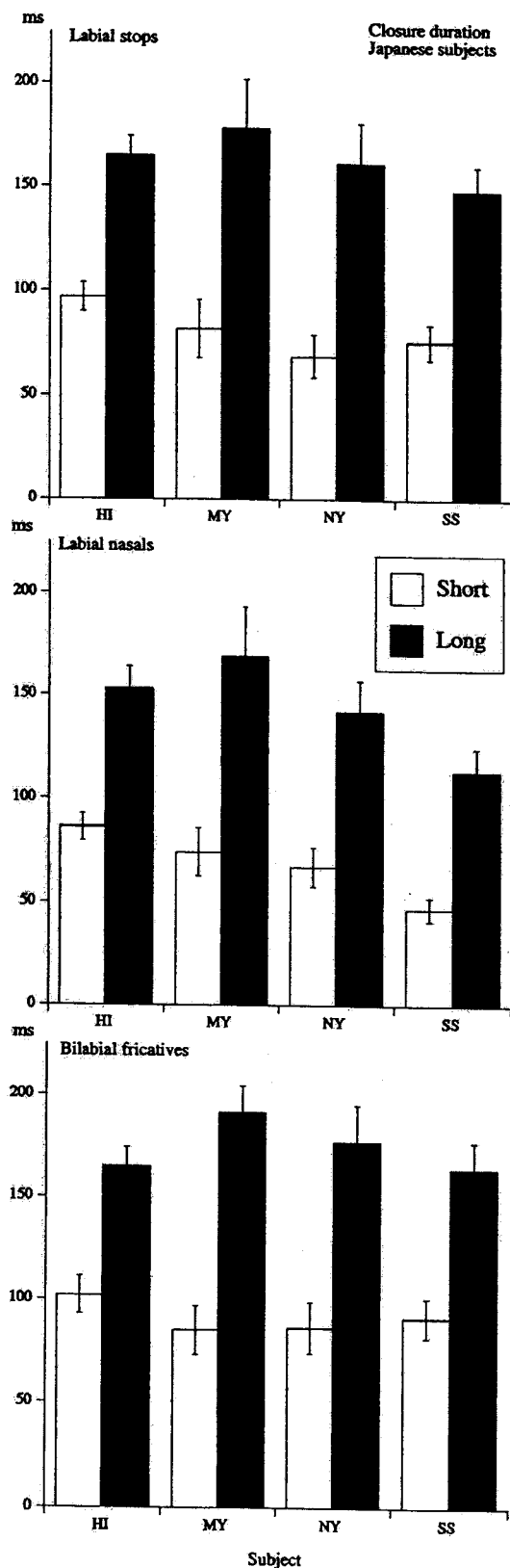


FIG. 3. Closure duration (mean and standard deviation) for the Japanese subjects.

mm. From Fig. 5, it is evident that there is no overall pattern for the speakers. The Swedish speaker AG produced the short and long consonant with almost identical velocities of the upper lip. On the other hand, Swedish speaker NR produced all the long consonants with a significantly higher velocity of the upper lip than the short ones [$t(98)=10.87, 12.62, \text{ and } 12.34$ for /p,b,m/, respectively, $p<0.001$]. In contrast, the Japanese speakers tended to use higher upper lip velocities for the short consonants than for the long ones. For the labial stops, the upper lip velocity difference was significant for subjects HI, NY, and SS, $t(98)=7.12, 4.56, \text{ and } 3.34, p<0.001$. For the nasals, the difference was significant for subjects HI, MY, and SS, $t(98)=15.61, 6.65, \text{ and } 6.01, p<0.001$.

C. Peak vertical lower lip position during the closure

Figure 6 plots the peak lower lip vertical position for the Swedish subjects. According to the hypothesis, the lower lip should reach a higher vertical position for the long than for the short consonants. The results for the two subjects differ, however. A comparison between the peak lower lip position for the long and short consonants showed no significant difference for Swedish subject AG, $t(98)=2.01, 1.61, 1.61, \text{ and } 0.39$ for /p,b,m,f/, but significant differences for Swedish subject NR, with the lower lip reaching a slightly higher position for the long consonants., $t(98)=6.8, 4.9, 7.56, \text{ and } 5.18$ for /p,b,m,f/, with $p<0.001$ in all cases.

Figure 7 plots the same results for the Japanese subjects. Here, the peak lower lip position was significantly higher for the long than for the short consonants for all four Japanese subjects. For HI, $t(98)=4.22, 8.4, \text{ and } 15.58$ for the stop, nasal, and fricative, respectively, with $p<0.001$. The corresponding t statistics for Japanese subject MY were 10.7, 11.99, and 8.51, for Japanese subject NY 17.55, 21.77, and 12.25, and for Japanese subject SS 26.62, 23.38, and 15.12.

These results thus mostly agree with the prediction of a higher peak lower lip vertical position during the closure for the long than for the short consonants. All the Japanese subjects and one of the Swedish subjects produced the long consonants with a higher lower lip vertical position during the closure/constriction for the long than for the short consonants. Due to the overlap in closure/constrictions durations for the Swedish consonants and their more restricted ranges than those of the Japanese subjects, the Swedish data are not particularly useful for testing the original hypothesis. Thus, in the following, the focus will be on the Japanese results.

D. Interactions between the lips

Table I presents the vertical position of the upper lip at the point in time when the lower lip reaches its highest vertical position for the Japanese subjects. As predicted, the upper lip vertical position at the point in time when the lower lip reaches its highest vertical position was significantly higher for the long than the short consonants for subject HI [$t(98)=3.45, 7.04, \text{ and } 7.09$, for the stop, nasal, and fricative, respectively, with $p<0.001$], the stops and nasals of subjects MY [$t(98)=3.54, \text{ and } 5.94$] and SS [$t(98)=11.28 \text{ and } 4.61$], and the stops of subject NY [$t(98)=9.07$]. How-

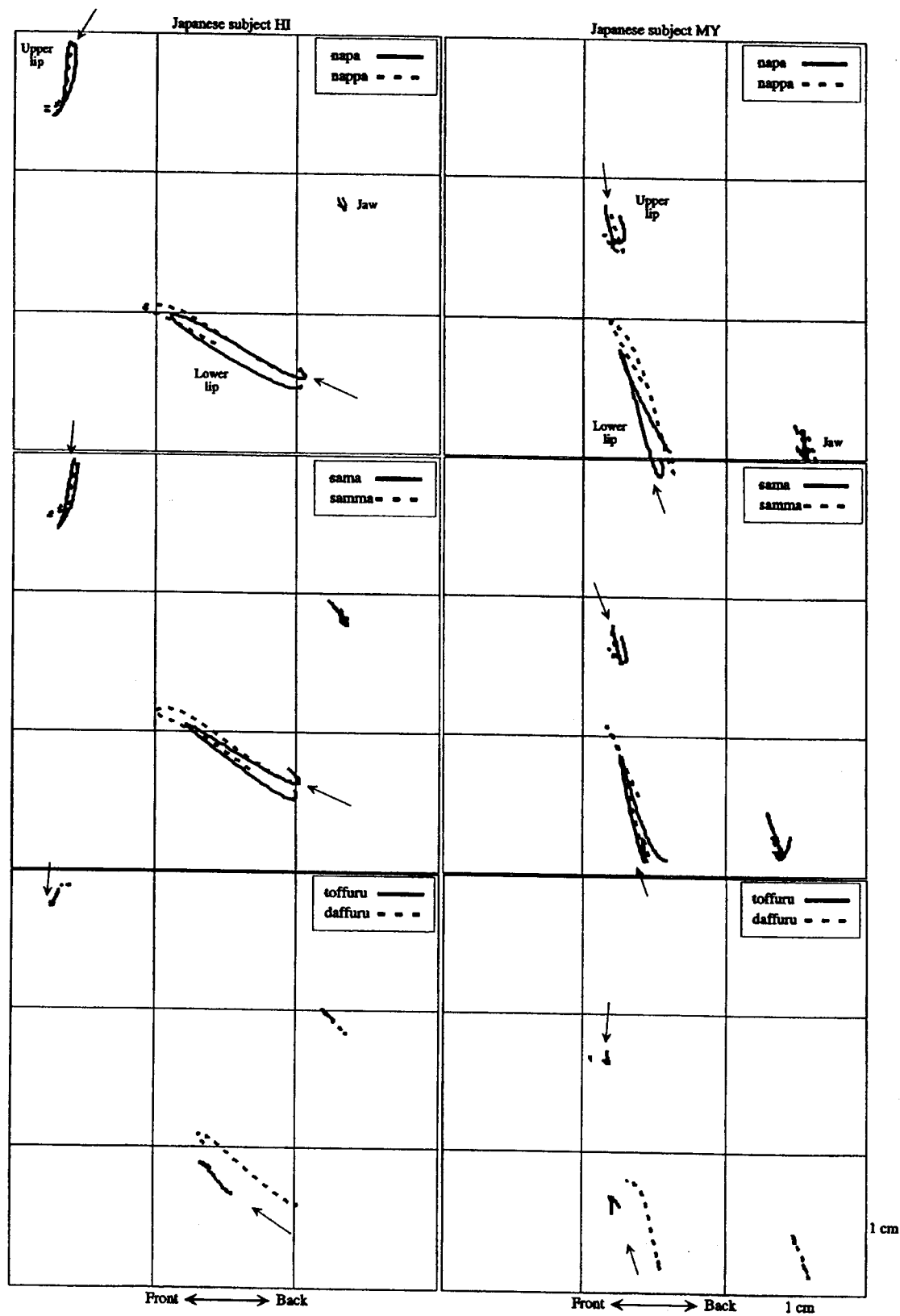


FIG. 4. Average lip and jaw signals for the Japanese subjects HI and MY. The arrows show the direction of movement. The subjects are facing to the left. Average lip and jaw signals for the Japanese subjects NY and SS. The arrows show the direction of movement. The subjects are facing to the left.

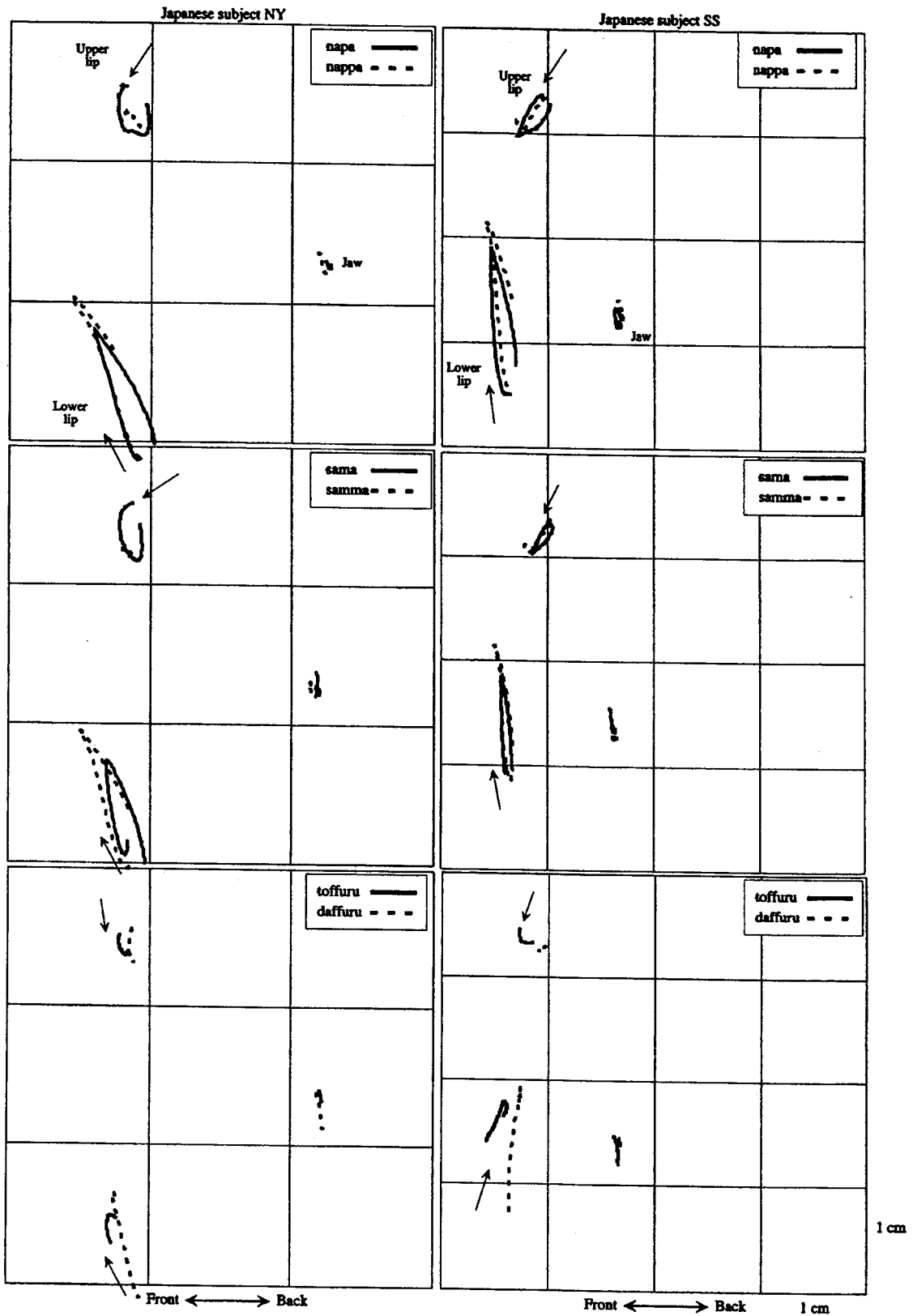


FIG. 4. (Continued.)

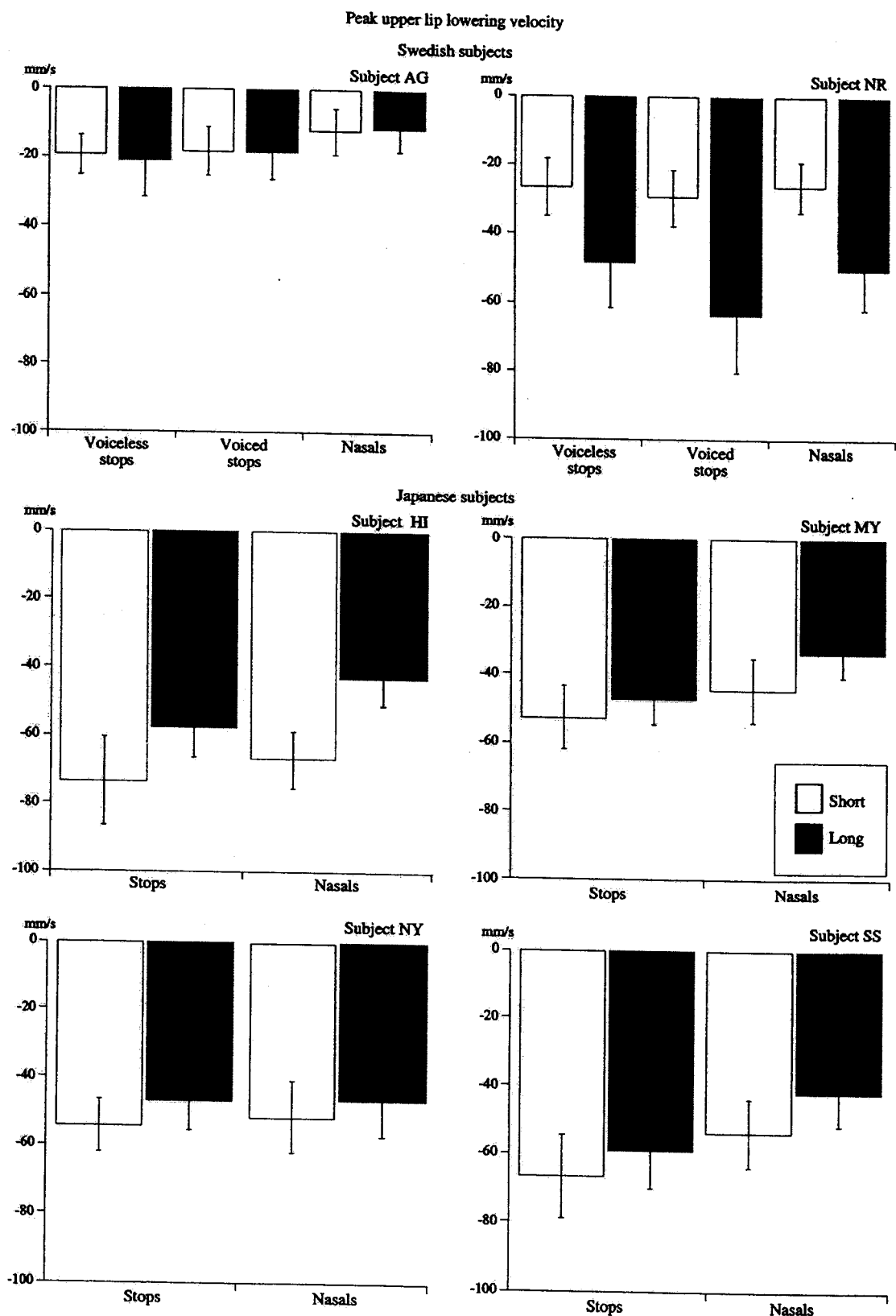


FIG. 5. Peak upper lip lowering velocity for the Swedish and Japanese subjects (mean and standard deviation).

Peak lower lip position during closure
Swedish subjects

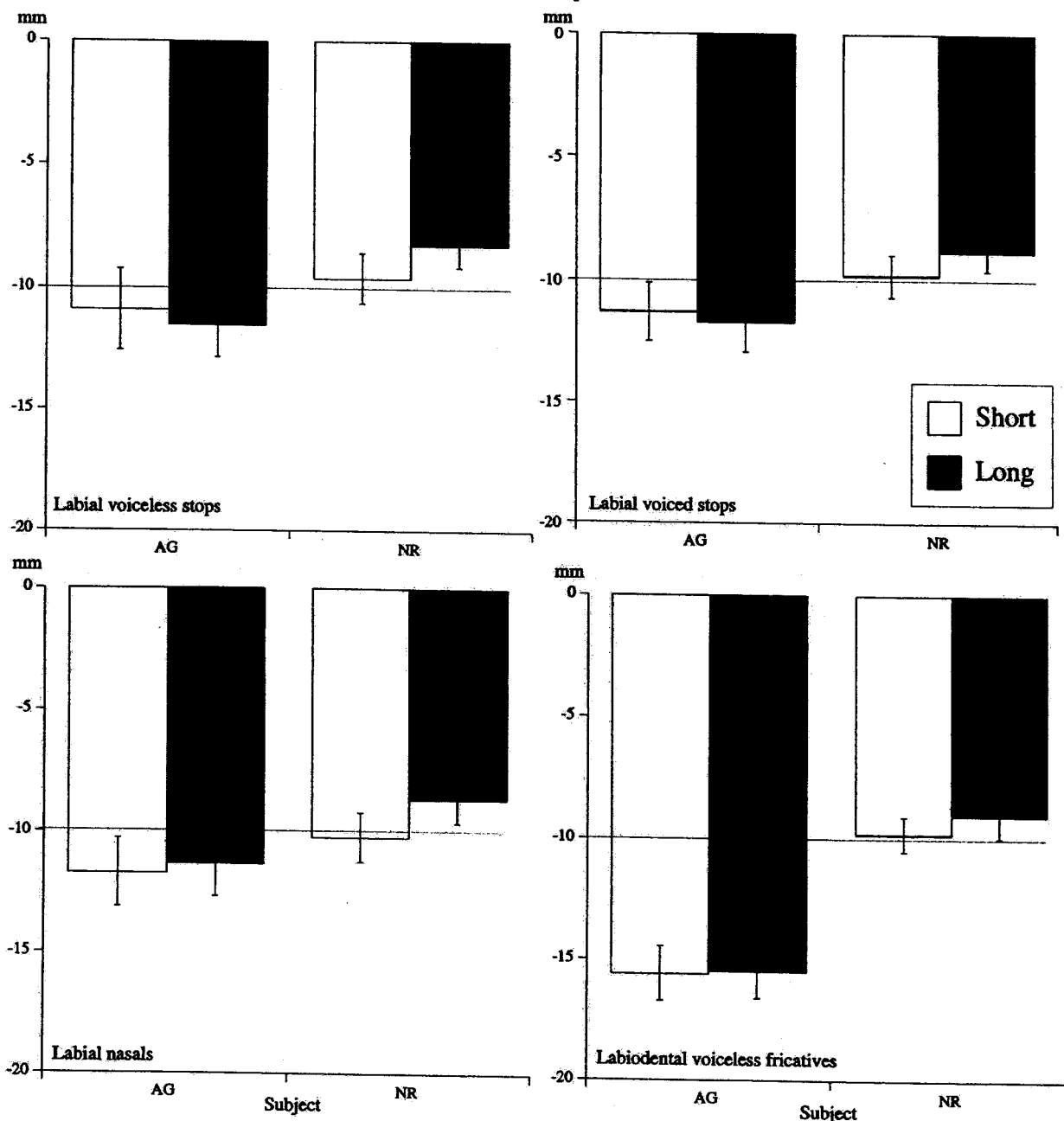


FIG. 6. Peak lower lip vertical position (mean and standard deviation) for the Swedish subjects.

ever, the fricatives of subjects NY and SS showed significant higher upper lip positions for the short than for the long consonant [$t(98)=4.41$, and 3.32 , $p<0.001$]. There was no difference for the fricative of Japanese subject MY [$t(98)=2.34$], and for the nasals of subject NY [$t(98)=2.5$].

E. Lower lip closing displacement

The third hypothesis predicts that the lower lip closing displacement is larger for the long than for the short consonants. In an attempt to make the vertical onset position of the

lower lip closing movement as similar as possible, the vowel context of the labial consonants were made similar. However, both the Swedish subjects produced the long and short consonants with different onset positions of the lower lip. It was always lower for the long consonants, most likely reflecting the more open vowel quality of the preceding short vowel and with higher first and, often, second formant frequencies than its long cognate (Fant, 1973). Thus, the hypotheses about larger lower lip closing displacements and velocities in long consonants could not be assessed for the

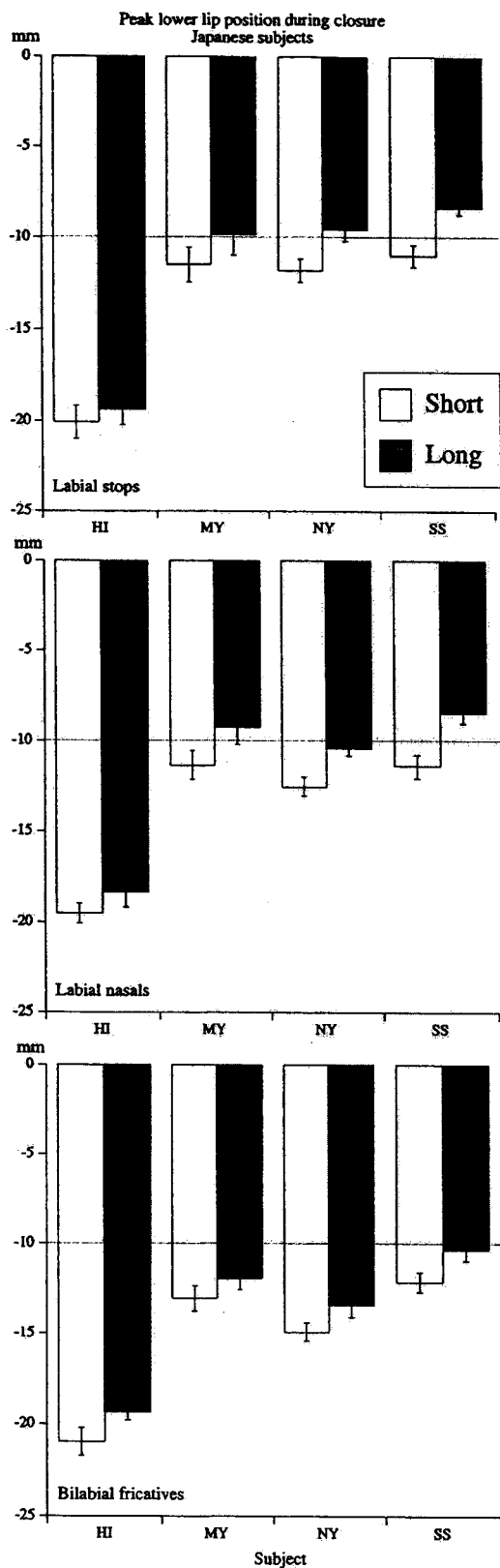


FIG. 7. Peak lower lip vertical position (mean and standard deviation) for the Japanese subjects.

TABLE I. Japanese subjects' upper lip vertical position at the point in time of lower lip peak vertical position (mm). The standard deviation is shown within parentheses.

Subject	p	pp	m	mm	f	ff
HI	-5.9 (0.81)	-5.4 (0.8)	-5.3 (0.6)	-4.3 (0.84)	-2.0 (0.71)	-1.2 (0.42)
MY	-4.7 (1.13)	-4.0 (0.97)	-4.7 (0.72)	-3.7 (0.95)	-3.1 (0.42)	-3.3 (0.5)
NY	2.2 (0.49)	3.4 (0.84)	2.0 (0.73)	2.5 (1.04)	3.8 (0.69)	3.2 (0.7)
SS	-0.3 (0.8)	1.1 (0.4)	0.3 (0.7)	1.1 (0.97)	3.2 (0.61)	2.8 (0.62)

Swedish subjects. Similarly, all four Japanese subjects produced the long bilabial fricative with a significantly lower lip onset position than its short cognate, so these two hypotheses could not be verified for this consonant in the Japanese material.

Table II summarizes the lower lip closing displacement for the stops and nasals for the Japanese subjects. For all subjects, the lower lip displacement is larger for the long than for the short consonants. The t -tests showed all the differences between the lower lip displacement for the long and short stops to be significant [Japanese subject HI: $t(98) = 3.83$; Japanese subject MY: $t(98) = 4.89$; Japanese subject NY: $t(98) = 11.02$; Japanese subject SS: $t(98) = 9.66$, with $p < 0.001$ in all cases]. The same was true for the nasals [Japanese subject HI: $t(98) = 7.33$; Japanese subject MY: $t(98) = 8.35$; Japanese subject NY: $t(98) = 17.41$; Japanese subject SS: $t(98) = 18.12$, with $p < 0.001$ in all cases].

Since this result could be due to a difference in onset position, offset position, or both, Table III shows the vertical onset position of the lower lip closing movement. The t -tests revealed no significant differences in the lower lip onset position between long and short consonants except for the nasals of subjects NY [$t(98) = 6.61, p < 0.001$] and SS [$t(98) = 18.12, p < 0.001$], which had a lower position for the long than for the short consonants.

The hypothesis about a larger lower lip raising displacement for the long than for the short consonants was supported by the results for the Japanese productions of stops and nasals. Moreover, there was no difference in the onset position of the lower lip closing movement between the long and short consonants, with two exceptions. However, inspection of Fig. 7 and Table III shows that the difference in the peak lower lip position for the nasals is 2.1 and 2.9 mm for subjects NY and SS, respectively, while the difference in the

TABLE II. Japanese subjects' lower lip raising displacement (mm). The standard deviation is shown within parentheses.

Subject	p	pp	m	mm
HI	4.5 (0.78)	5.1 (0.73)	4.2 (0.58)	5.3 (0.87)
MY	9.7 (1.53)	11.3 (1.65)	7.6 (1.22)	9.9 (1.5)
NY	9.6 (0.99)	11.6 (0.82)	6.8 (0.85)	10.0 (0.95)
SS	13.9 (0.5)	16.5 (1.46)	9.4 (1.02)	13.1 (1.03)

TABLE III. Japanese subjects' lower lip vertical position at the onset of the closing movement. The standard deviation is shown within parenthesis.

Subject	p	pp	m	mm
HI	-24.6 (0.78)	-24.5 (0.83)	-23.7 (0.5)	-23.6 (0.46)
MY	-21.7 (1.37)	-21.1 (1.09)	-19.0 (1.01)	-19.1 (0.93)
NY	-21.4 (1.01)	-21.1 (0.75)	-19.4 (0.63)	-20.4 (0.84)
SS	-24.9 (1.12)	-24.9 (1.39)	-20.8 (0.88)	-21.6 (1.09)

onset position is 1.0 and .8 mm. Thus, the difference in the peak lower lip position is greater than the difference in the lower lip onset position. Hence, the results the nasals of subjects NY and SS are compatible with the original hypothesis about a larger closing displacement of the lower lip for the long than for the short consonants.

F. Lower lip peak closing velocity

The next hypothesis to be evaluated predicts that the peak closing velocity of the lower lip is higher for the long than for the short consonants. This is based on the commonly found strong relationship between movement displacement and peak velocity (e.g., Cooke, 1980; Ostry *et al.*, 1983; Kelso *et al.*, 1985; Vatikiotis-Bateson and Kelso, 1993; Hertrich and Ackermann, 1997; Löfqvist and Gracco, 1997). Figure 8 plots the peak velocity of the lower lip closing movement for the Japanese subjects' production of stops and nasals. Contrary to the prediction, there were no significant differences between the peak closing velocity of the lower lip between the long and short consonants except for the stops of subject MY [$t(98) = 3.44, p < 0.001$], but here the short stops were produced with a higher lower lip closing velocity. Thus, this particular prediction was not supported by the data. This is puzzling given the commonly observed very strong correlation between movement displacement and peak velocity. To provide a closer view of this particular relationship, Fig. 9 plots the lower lip displacement and peak velocity separately for the long and short consonants, the stops in Fig. 9(a) and the nasals in Fig. 9(b). Interestingly, the expected positive relationship only holds within the long and short consonants, but not between them. All the correlations are significant and all are above 0.9, with the exception of the long stops and nasals of subject NY (lower left panels in Fig. 9). The predicted results would have the data points for the long and short consonants form a continuous function, with the long ones, the open triangles, being higher than the short ones, the filled squares. Thus, these results strongly suggest that these subjects do not control the duration of the oral closure in long and short consonants by only varying the position of a virtual target of the lower lip during the oral closure: The long and short consonants have different control regimes. The original hypothesis is thus wrong. So, how do these subjects control the duration of the oral closure? One potential answer is provided by the slopes of the regressions shown in Fig. 9. This slope is one representation of the stiffness of the movement (cf. Kelso *et al.*, 1985). For all sub-

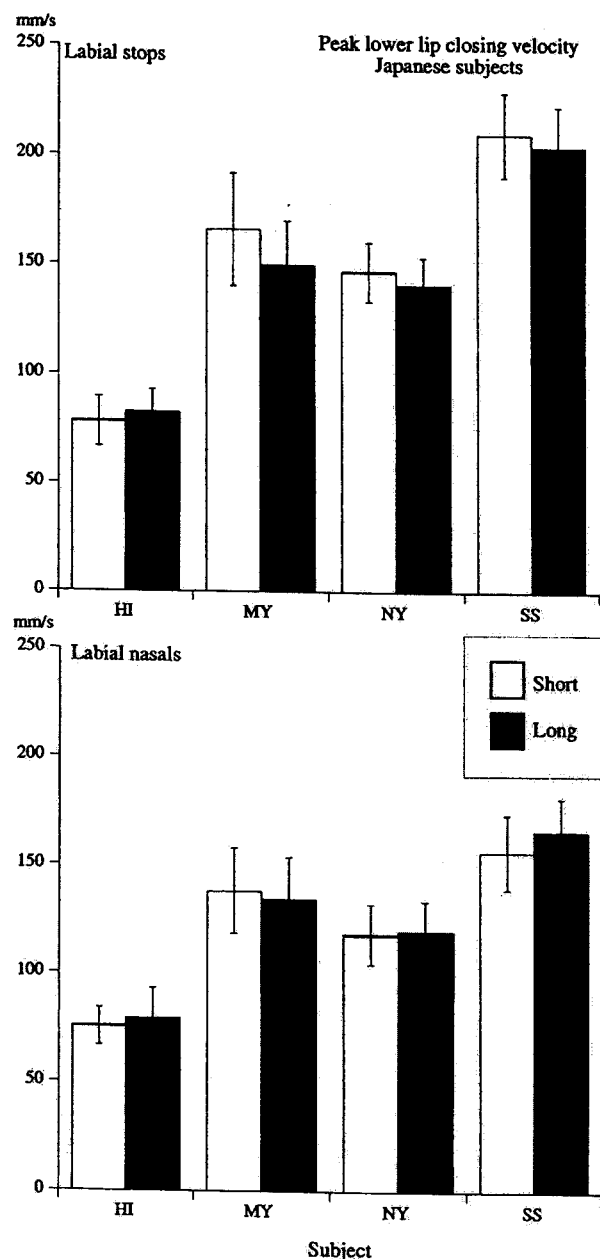


FIG. 8. Peak velocity of the lower lip closing movement (mean and standard deviation) for the Japanese subjects.

jects except HI (where the slopes are almost identical), the slopes are higher for the short than for the long consonants, thus suggesting that the short consonants are produced with stiffer movements. However, examination of the 95% confidence intervals for the slopes of the long and short consonants only showed no overlap for subject MY.

Before examining movement stiffness in more detail, we will look at the peak closing velocity of the derived lip aperture signal. The results are shown in Table IV. The first thing to note is that in most of the cases, the short consonant is produced with a higher lip aperture closing velocity; the only exceptions are the nasals of subject SS, where the long consonant has a higher velocity, and of subject NY, where

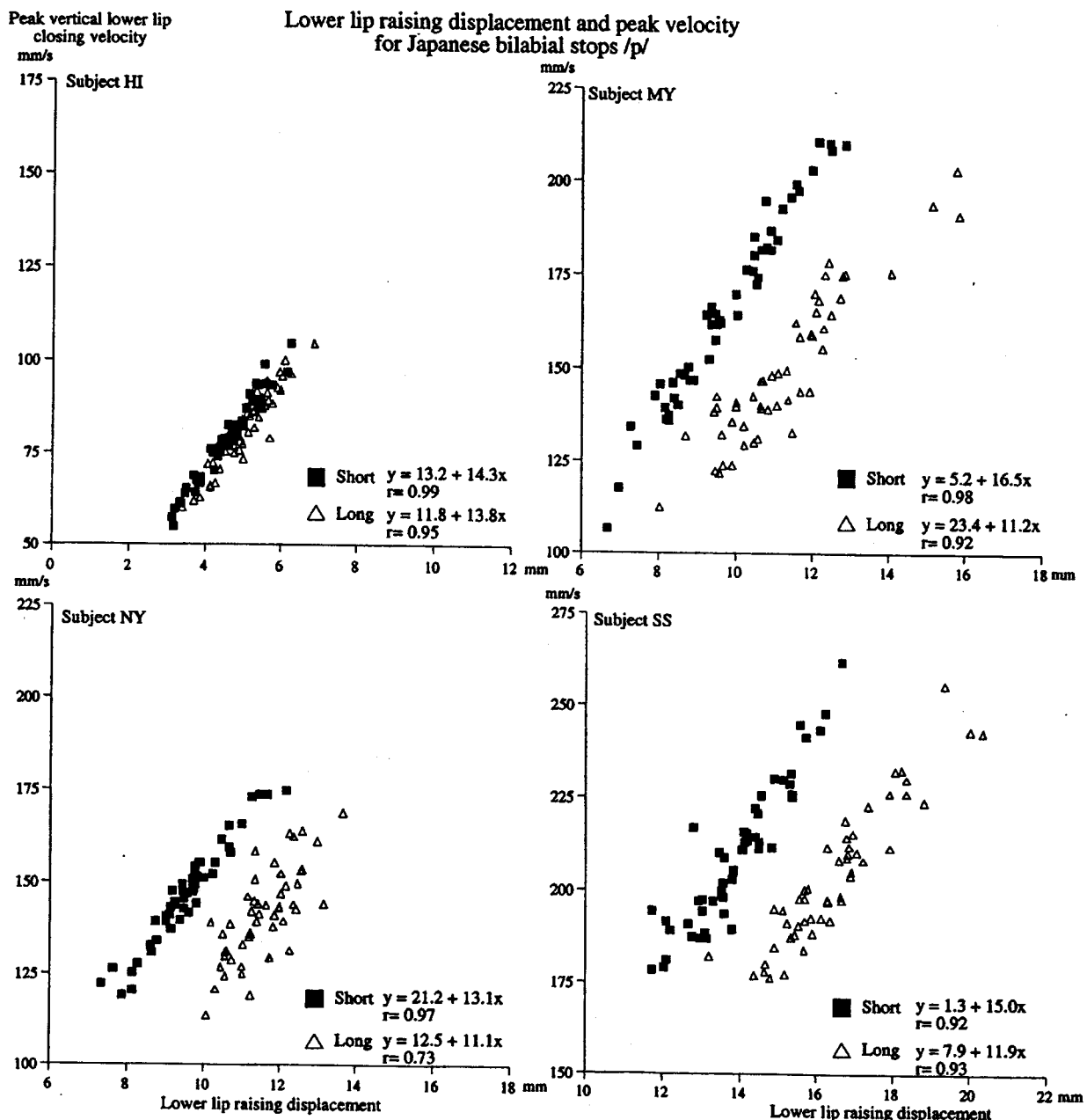


FIG. 9. Lower lip raising displacement and peak velocity for the Japanese subjects' production of labial stops. Lower lip raising displacement and peak velocity for the Japanese subjects' production of labial nasals.

there is no difference. Overall, the difference is not very large, however, and the statistical analysis only showed three of the differences to be significant: the stops for subjects MY [$t(98) = 3.93, p < 0.001$] and NY [$t(98) = 3.38, p < 0.001$], and the nasals of subject HI [$t(98) = 7.01, p < .001$].

G. Velocity and acceleration of the lower lip

Figures 10 and 11 plot averages (aligned to the beginning of the acoustic closure) of the lower lip position, velocity, and acceleration signals for the Japanese subjects' productions of labial stops and nasals, respectively; Fig. 10 shows the stops and Fig. 11 the nasals. The long consonants

are shown by the dashed lines. Several observations can be made about these results. First, the position signals show that the peak position of the lower lip does not occur at the same point in time for the long and short consonants. Only for the stops of subject HI is there a similarity in the timing of the peaks in the position signals for the long and short consonants. For all other cases, the peak lower lip position occurs later in the long than in the short consonant. Second, the lower lip raising velocity signals for the short consonant show the bell-shaped characteristic of simple movements. However, for the long consonants, the velocity of the lower lip shows a change around the second zero crossing, just

Peak vertical lower lip
closing velocity

Lower lip raising displacement and peak velocity
for Japanese bilabial nasals /m/

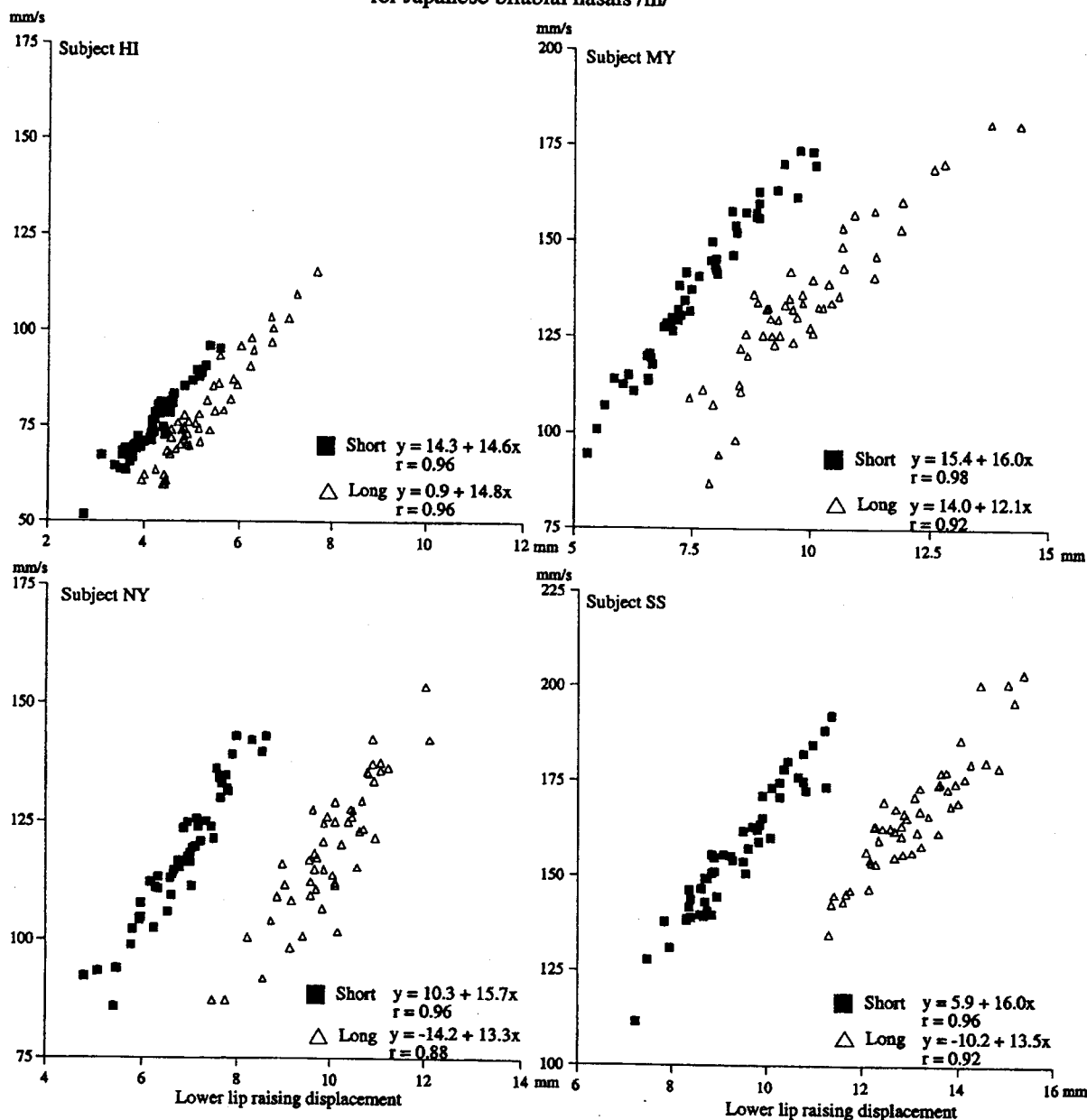


FIG. 9. (Continued.)

before it for subjects MY, NY, and SS, and just after it for subject HI, so that the velocity curve is not symmetric. Third, the acceleration signals indicate that the deceleration of the lower lip is momentarily reduced for the long consonants. This adjustment is apparently made to maintain the lower lip in a high position (subject HI) or keep it moving upwards (subjects MY, NY, and SS) and thus in contact with the upper lip for a longer period of time. For the long stops of subject HI, the deceleration almost reaches zero [lower left panel in Fig. 10(a)].

TABLE IV. Peak closing velocity of the lip aperture signal. The standard deviation is shown within parenthesis.

Subject	p	pp	m	mm
HI	-139.2 (13.79)	-128.8 (10.04)	-131.6 (11.16)	-113.5 (14.53)
MY	-189.8 (23.11)	-173.3 (18.67)	-157.8 (17.69)	-149.0 (17.44)
NY	-183.4 (14.75)	-173.2 (15.24)	-157.0 (19.51)	-157.1 (20.38)
SS	-255.6 (24.44)	-243.9 (25.03)	-191.0 (18.9)	-195.5 (18.2)

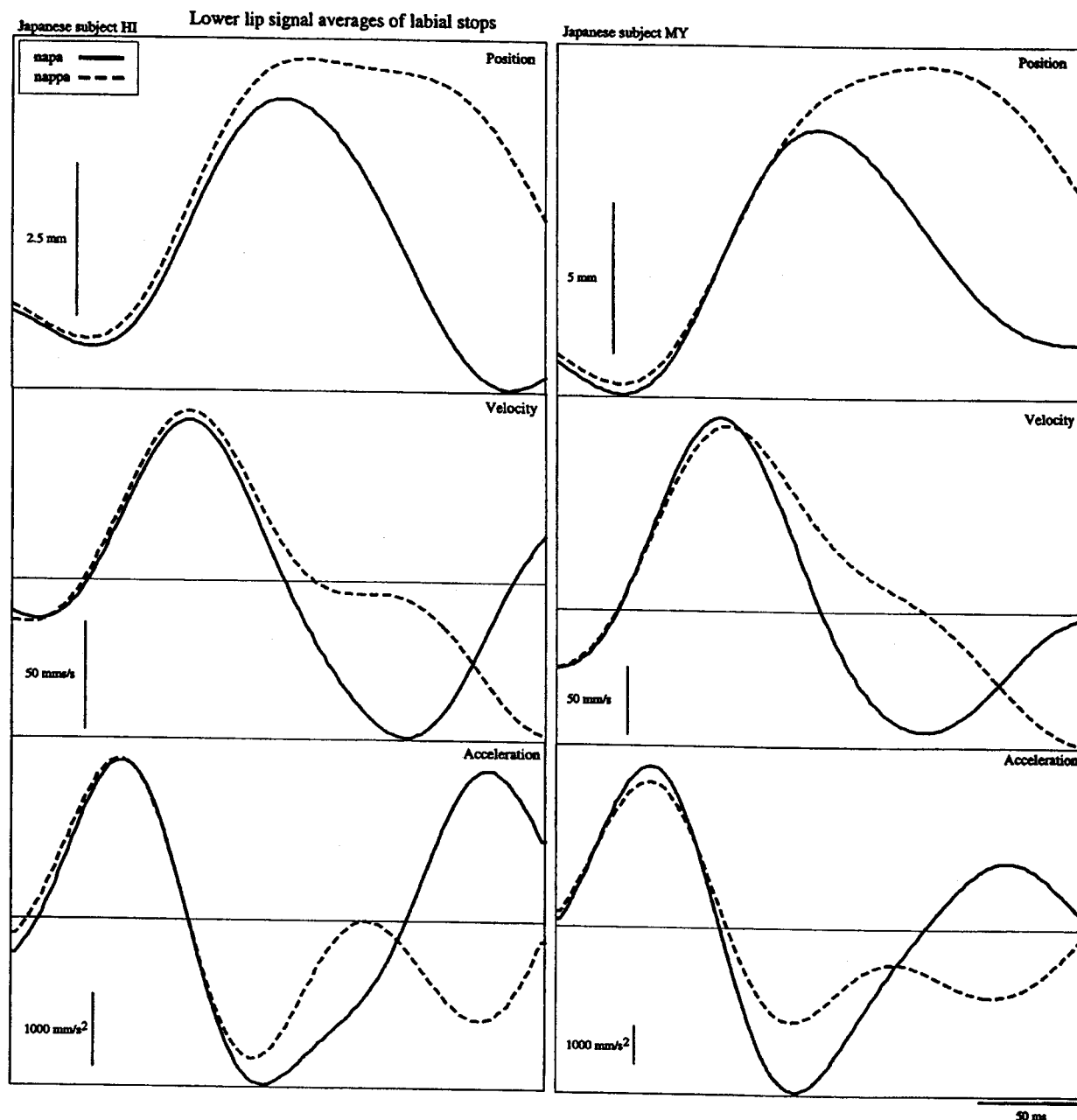


FIG. 10. Signal averages of lower lip position, velocity, and acceleration for the labial stops produced by the Japanese subjects HI and MY. The vertical line in the velocity and acceleration panels represents zero velocity and acceleration. Signal averages of lower lip position, velocity, and acceleration for the labial stops produced by the Japanese subjects NY and SS. The vertical line in the velocity and acceleration panels represents zero velocity and acceleration.

H. Movement stiffness

The regressions shown in Fig. 9 between the displacement and peak velocity of the lower lip closing movement suggest that there might be a difference in the stiffness of the movements for the long and short consonants. To pursue this issue in more detail, another measure of stiffness was calculated as the temporal interval between movement onset and peak velocity (cf. Adams *et al.*, 1993; Hertrich and Ackermann, 1997). This was applied to both the lower lip and lip aperture closing movements. The results are shown in Fig. 12. In all cases, this interval is shorter for the short than for

the long consonants, suggesting a stiffer movement for the short than for the long ones. The statistical analysis showed this difference to be significant for the lower lip and the stops [Fig. 12(a)], $t(98) = 3.15, 10.65, 9.36$, and 9.42 for Japanese subjects HI, MY, NY, and SS, respectively, with $p < 0.001$ in all cases. The same was true for the lower lip and the nasals, $t(98) = 16.59, 6.67, 17.63$, and 9.33 for Japanese subjects HI, MY, NY, and SS, respectively, with $p < 0.001$ in all cases. For the lip aperture, shown in Fig. 12(b), this interval was

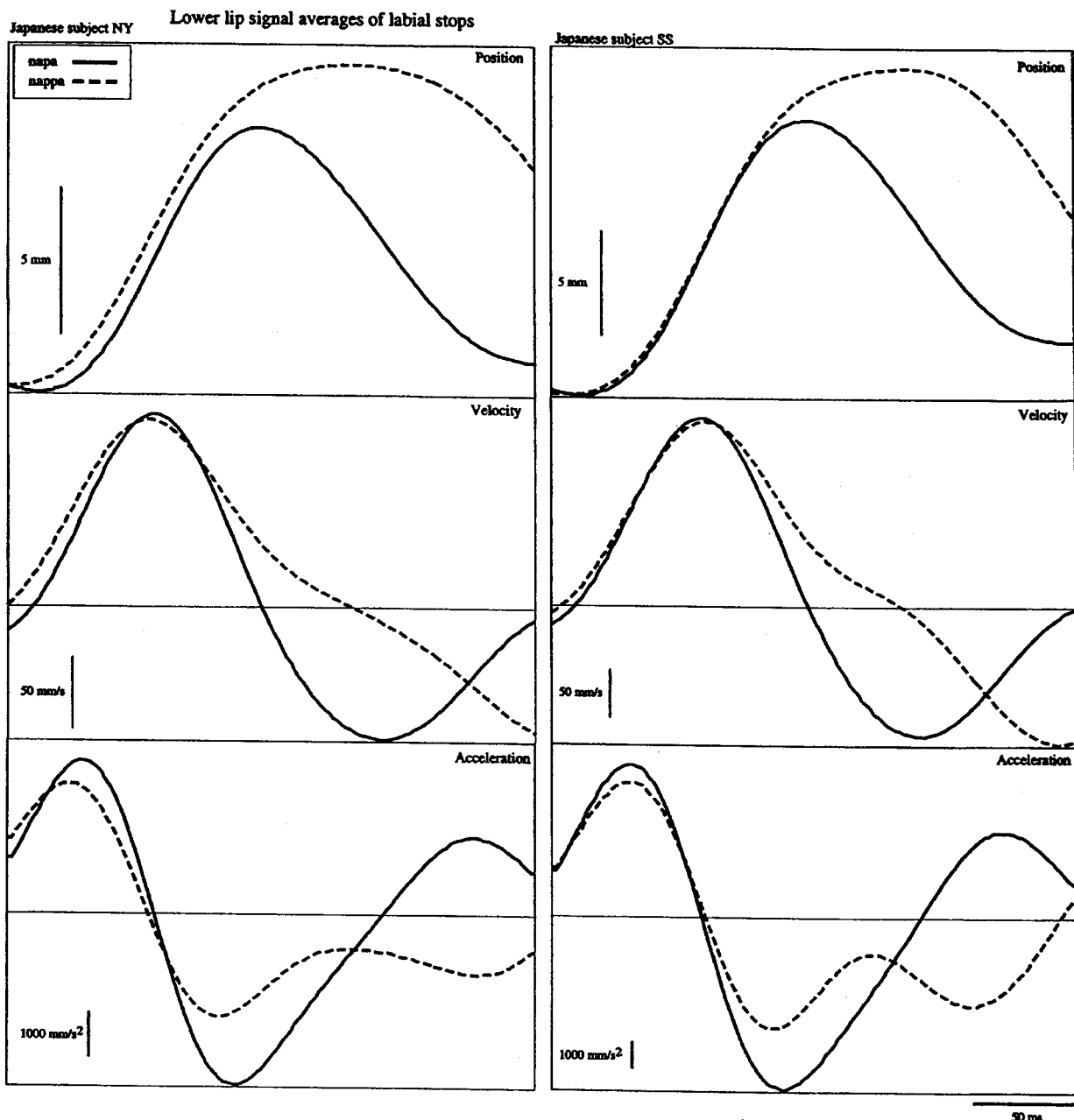


FIG. 10. (Continued.)

reliably shorter for the short than for the long stops in the three subjects MY, NY, and SS [$t(98)=6.4$, 6.87, and 9.47, with $p<0.001$ in all cases]. The large standard deviation for subject HI is due to the fact that it was sometimes hard to find a proper zero crossing in her lip aperture velocity signal. For the nasals, the interval was again shorter in the short than in the long consonants in three subjects HI, NY, and SS [$t(98)=8.16$, 10.96, and 8.25, with $p<0.001$ in all cases]. Subject MY did not show a statistical difference for the nasals.

IV. DISCUSSION

Consistent with the different structure of the length contrast in Swedish and Japanese, the distributions of closure/

constriction duration for long and short consonants showed considerable overlap in the Swedish data but not in the Japanese data. As a consequence, the distributions of closure/constriction durations spanned a smaller temporal range in Swedish than in Japanese, making the Japanese data more suitable for testing the original hypothesis about virtual targets. As noted, in Swedish, there are variations in the duration and quality of the preceding vowel (cf. Hadding-Koch and Abramson, 1964; Fant, 1973). Overall the results for closure/constriction duration in Swedish and Japanese found in the present study are in close agreement with those obtained in other studies.

The results for all four Japanese speakers, and most of

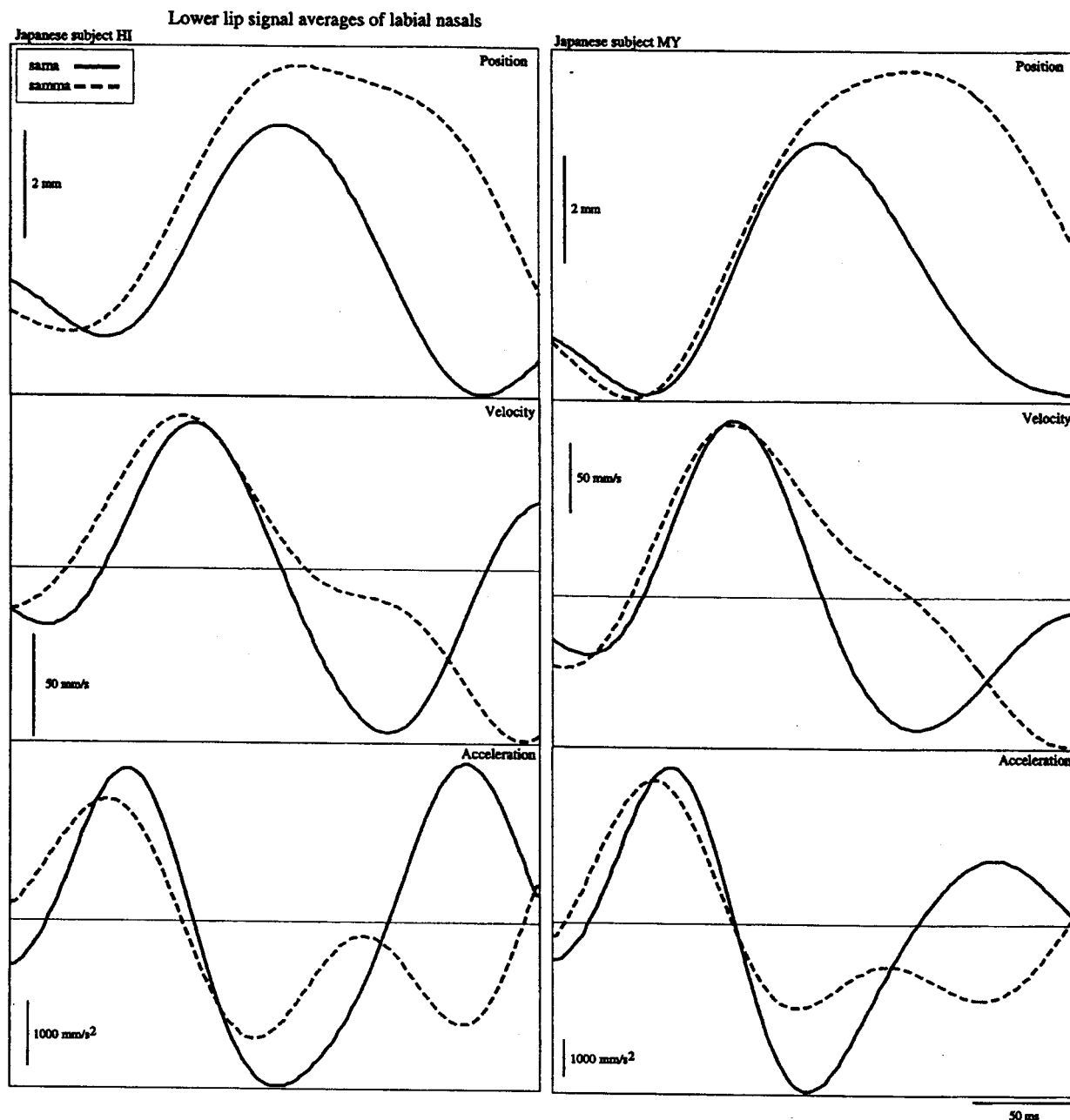


FIG. 11. Signal averages of lower lip position, velocity, and acceleration for the labial nasals produced by the Japanese subjects HI and MY. The vertical line in the velocity and acceleration panels represents zero velocity and acceleration. Signal averages of lower lip position, velocity, and acceleration for the labial nasals produced by the Japanese subjects NY and SS. The vertical line in the velocity and acceleration panels represents zero velocity and acceleration.

them for one of the Swedish speakers (NR), confirm three of the four predictions stated in the introduction: The lower lip reaches a higher vertical position during the closure for long than for short labial consonants (see the results for Swedish subject NR in the right half of the panels in Fig. 6, and all the results for the Japanese subjects in Fig. 7). As shown in Table I, the vertical position of the upper lip is higher at the point in time when the lower lip is at its highest position during the closure for the long stops and nasals, but generally not for the fricatives. This is most likely due to the complete oral

closure for the stops and the nasal, whereas the fricative requires a narrow constriction. The closing movement of the lower lip has a larger displacement for long than for short consonants (Table II) which is generally due to a difference in the end position (Fig. 7) and not in the onset position (Table III). Crucially, however, the final prediction was not supported by the empirical results. That is, the predicted higher peak closing velocity of the lower for the long than for the short consonants was not found (Fig. 8). Thus, the observed differences in lower lip movement between long

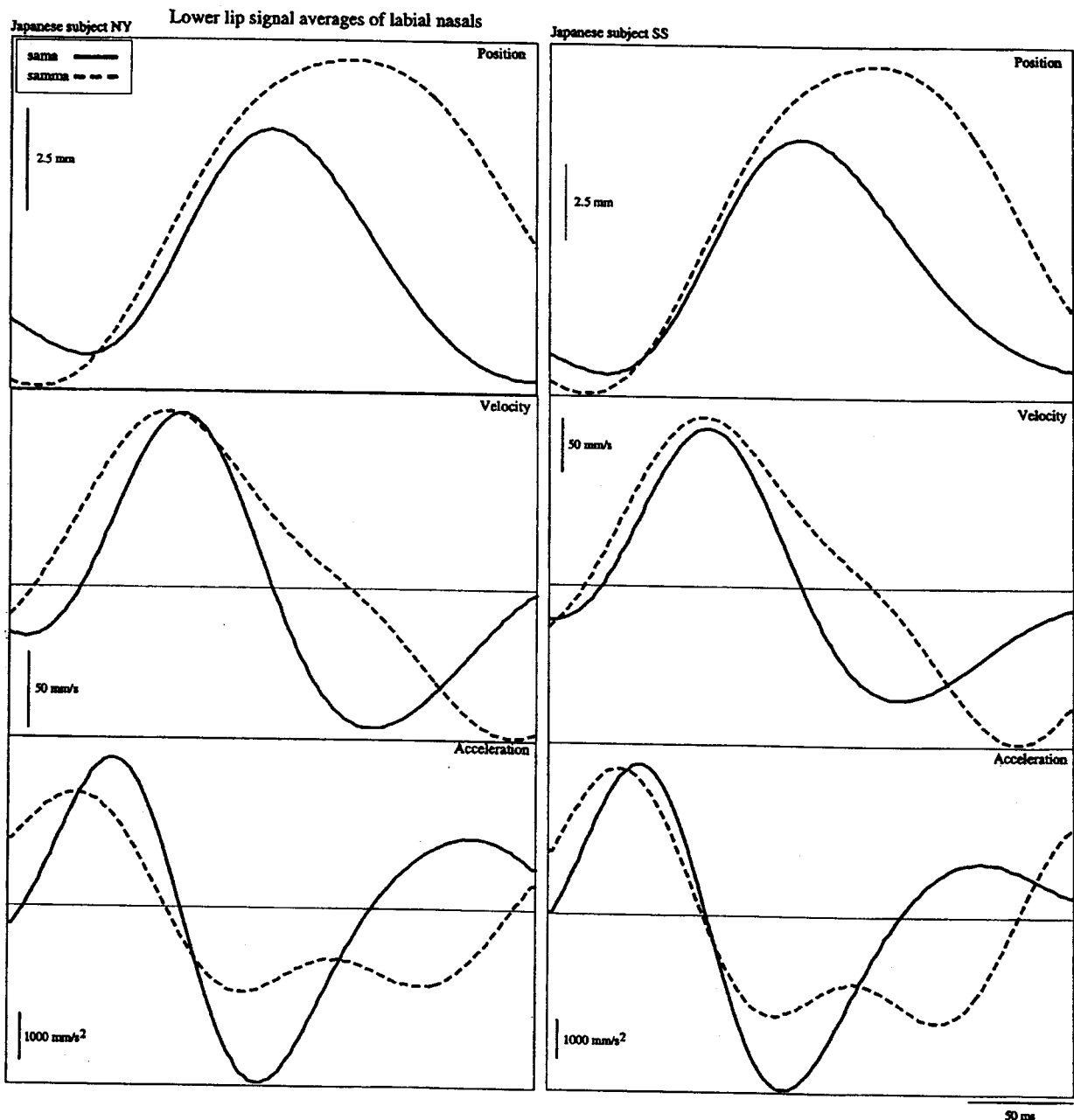


FIG. 11. (Continued.)

and short consonants are not due to a change only in virtual target position. Interestingly, the expected close relationship between movement displacement and peak velocity was only found to hold within the long and short consonants separately (Fig. 9) but not across them, as originally predicted. Assuming that speakers use virtual targets for consonant movements, these results suggest that the subjects did not produce the different closure/constriction durations by only changing the position of a virtual target, but rather by changing both its position and timing. In particular, they modified the deceleration of the lower lip movement to keep it in contact with the upper lip for a longer period of time (Figs. 10 and 11). The original hypothesis assumed a change only

in the displacement of the underlying excitation pulse (cf. Gottlieb *et al.*, 1989), but the results also suggest that the duration of the pulse is changed. Although no electromyographic recordings were attempted in this study, such recording might show a longer duration of activity in the lower lip muscles for the long than for the short consonants (cf. Lehiste *et al.* 1973). At the same time, the lower lip movement pattern during the closure is affected by its contact with the upper lip. Recordings of the contact pressure between the lips could help in further clarifying their interactions.

The results shown in Fig. 2 for the Japanese subjects indicate that the lip movements predominantly occur in the vertical dimension, except for subject HI. Interestingly, the

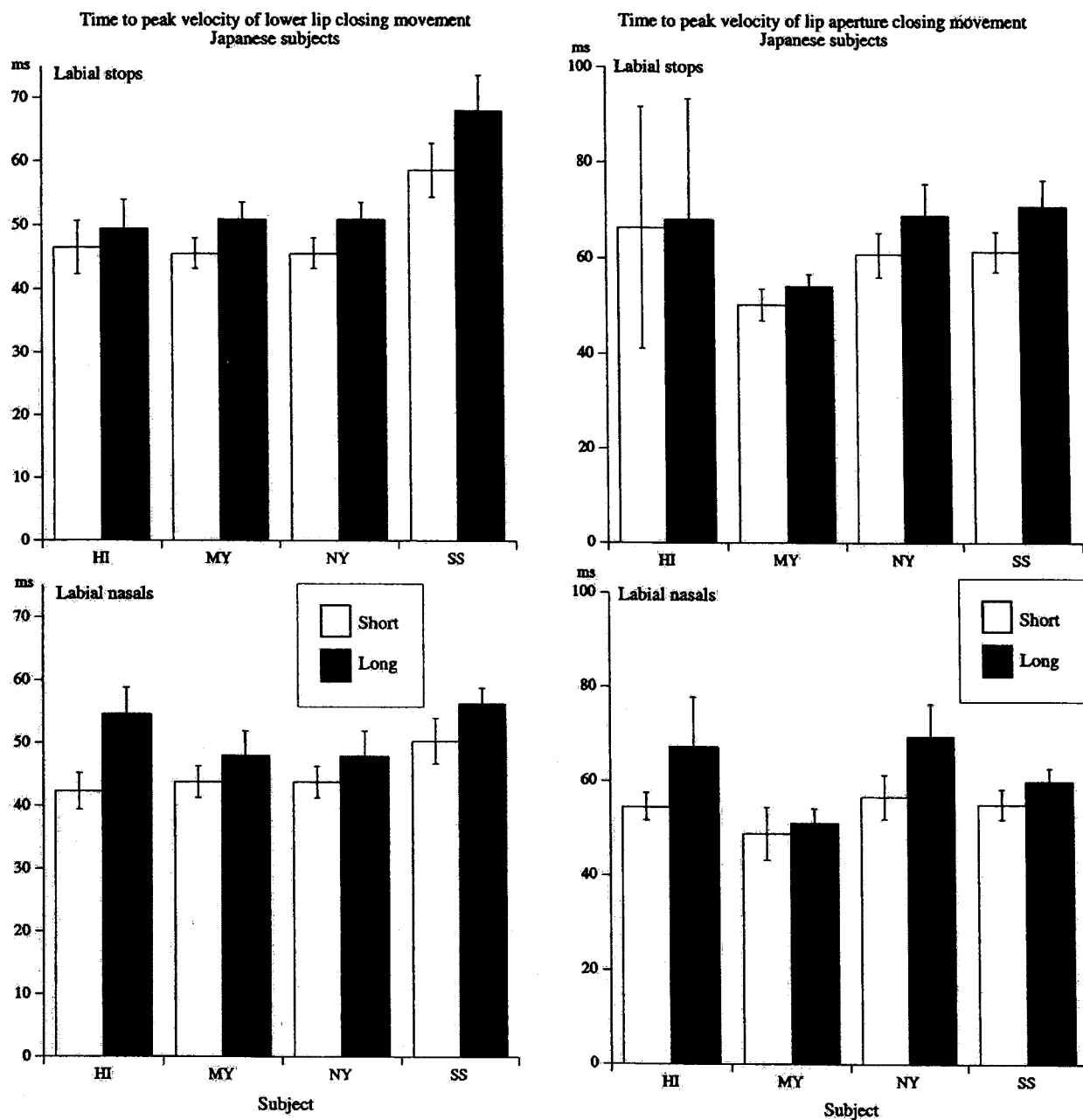


FIG. 12. The interval from movement onset to peak velocity (mean and standard deviation) of the lower lip closing movement for the Japanese subjects. The interval from movement onset to peak velocity (mean and standard deviation) of the lip aperture closing movement for the Japanese subjects.

movement pattern of her lower lip [shown in the left panels of Figs. 10(a) and 11(a)] shows that its peak position occurred approximately at the same time for the long and short consonants; for the other Japanese subjects the lower lip reached its highest position later for the long than for the short consonants. The results for subject HI also differed from those of the other three Japanese subjects in the upper lip movement pattern during the oral closure when the two lips interact. Specifically, all the upper lip movement patterns of subject HI were similar to the one shown in the left panel of Fig. 1, i.e., the lip kinematics during the production of a short labial nasal; this was also true for the long consonants.

In contrast, the other three subjects generally showed a pattern like the one shown in the right panel of Fig. 1, where there is more than one zero-crossing in the upper lip velocity signal during the closure for a long consonant. Possibly, anatomical differences in bite type may explain these differences between the subjects. The movement trajectories shown in Fig. 4 indicate that the vertical separation between the upper and lower lip receivers during the oral closure is slightly larger for subject HI than for the other subjects. Although differences in receiver position will influence the nature and amount of observed lip interaction during the oral closure, the results shown in Table I do not indicate that subject HI

generally shows a different pattern of upper and lower lip interaction that might be due to different receiver placement.

Although no systematic differences in jaw kinematics were found between the long and short consonants, the upper lip lowering movement peak velocity was higher for the short than for the long consonants in the Japanese subjects (Fig. 5), but not always statistically significant. Similarly, the lip aperture closing velocity tended to be higher for the short than the long consonants, although the differences were only significant in three cases (Table IV). These trends are similar to the results presented by Smith (1995). At the same time, the Swedish subject who produced the long and short consonants with different durations (NR) produced the long ones with a higher upper lip closing velocity than the short ones, cf. Fig. 5. Interestingly, Smith (1995) found a similar difference between Japanese and Italian.

The lip movements of the short consonants tended to be made with higher stiffness, measured both in the lower lip and in the lip aperture signal (Fig. 12). This might be due to the shorter movement times of the lips for the short consonants, since movement stiffness tends to decrease with movement duration, at least when such durational changes have been due to changes in speaking rate (e.g., Adams *et al.*, 1993).

In summary, the present results suggest that the original hypothesis about only a change in the position of a virtual target for the lips in the control of closure duration needs to be modified. Clearly, and as shown in Figs. 10 and 11, both the position and the timing of the target is changed. At the same time, the idea of a spatial control of a durational property in speech is interesting and receives some qualified support in the present study. That is, a speaker can use variations in lip displacement to make changes in closure duration for labial consonants.

To further understand the production of long and short consonants, some additional studies can be considered. One is to study the development of this contrast in children's speech, since lip movements can be recorded using noninvasive procedures (e.g., Green *et al.*, 2000). Another interesting issue in the production of long and short labial consonants is the coordination of lip and tongue movements. Löfqvist and Gracco (1999) showed that more than 50% of the tongue movement from the first to the second vowel in a VCV sequence with a labial stop occurred during the stop closure. The material presented by Smith (1995) suggests that the tongue movement might be altered so that the tongue moves slower during the closure for a long consonant in Japanese. Finally, since the tongue is moving during the closure for a lingual stop consonant (e.g., Löfqvist and Gracco, 2002), one might predict that the tongue movement during the closure is slower for a long than for a short consonant in order to maintain the contact between the tongue and the palate. Löfqvist (2003, 2004) presents evidence that this is indeed the case.

ACKNOWLEDGMENTS

The author is grateful to Mariko Yanagawa for help with the Japanese material, and to Peter Assman, Dani Byrd, and an anonymous reviewer for comments on an earlier version of the manuscript. This work was supported by Grant No.

DC-00865 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

- Adams, S., Weismer, G., and Kent, R. (1993). "Speaking rate and speech movement velocity profiles," *J. Speech Hear. Res.* 36, 41–54.
- Beckman, M. (1982). "Segment duration and the mora in Japanese," *Phonetica* 39, 113–135.
- Byrd, D. (1995). "Articulatory characteristics of single and blended lingual gestures," in *Proceedings of the XIIIth International Congress of Phonetic Sciences*, edited by K. Elenius and P. Branderud, Stockholm, Vol. 2, pp. 438–441.
- Byrd, D., Kaun, A., Narayanan, S., and Saltzman, E. (2000). "Phrasal signatures in articulation," in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, edited by M. B. Broe and J. B. Pierrehumbert (Cambridge U.P., Cambridge), pp. 70–87.
- Cooke, J. (1980). "The organization of simple, skilled movements," in *Tutorials in Motor Behavior*, edited by G. Stelmach and J. Requin (North-Holland, Amsterdam), pp. 199–212.
- Dunn, M. H. (1993). "The Phonetics and Phonology of Geminate Consonants: A Production Study," unpublished doctoral dissertation, Yale University.
- Elert, C.-C. (1964). *Phonologic Studies of Quantity in Swedish* (Almqvist & Wiksell, Uppsala).
- Fant, G. (1973). "Acoustic description and classification of phonetic units," in *Speech Sounds and Features*, edited by G. Fant (Ed.) (MIT, Cambridge, MA), pp. 32–83.
- Farnetani, E. (1990). "V-C-V lingual coarticulation and its spatiotemporal domain," in *Speech Production and Speech Modelling*, edited by W. Hardcastle and A. Marchal (Kluwer, Dordrecht), pp. 93–130.
- Fuchs, S., Perrier, P., and Mooshammer, C. (2001). "The role of the palate in tongue kinematics: An experimental assessment in VC sequences," *Proc. Eurospeech 2001*, Aalborg, pp. 1487–1490.
- Gottlieb, G., Corcos, D., and Agarwal, G. (1989). "Strategies for the control of voluntary movements with one mechanical degree of freedom," *Behav. Brain. Sci.* 12, 189–250.
- Green, J. R., Moore, C. A., Higashikawa, M., and Steeve, R. (2000). "The physiological development of speech motor control: Lip and jaw coordination," *J. Speech Lang. Hear. Res.* 43, 239–255.
- Hadding-Koch, K., and Abramson, A. (1964). "Duration versus spectrum in Swedish vowels: Some perceptual experiments," *Studia Linguistica* 18, 94–107.
- Han, M. (1994). "Acoustic manifestations of mora timing in Japanese," *J. Acoust. Soc. Am.* 96, 73–82.
- Harrington, J., Fletcher, J., and Roberts, C. (1995). "Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data," *J. Phonetics* 23, 305–322.
- Hertrich, I., and Ackermann, H. (1997). "Articulatory control of phonological vowel length contrasts: Kinematic analysis of labial gestures," *J. Acoust. Soc. Am.* 102, 523–536.
- Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E., and Kay, B. (1985). "A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling," *J. Acoust. Soc. Am.* 77, 266–280.
- Lahiri, A., and Hankamer, J. (1988). "The timing of geminate consonants," *J. Phonetics* 16, 327–338.
- Lehtonen, J. (1970). *Aspects of Quantity in Standard Finnish*, K. J. Gummerus Jyväskylä.
- Lehiste, I., Morton, K., and Tatham, M. (1973). "An instrumental study of consonant gemination," *J. Phonetics* 1, 131–148.
- Löfqvist, A. (2000). "Control of closure duration in stop consonants," in *Proceedings of the 5th Seminar on Speech Production: Models and Data* (Institut für Phonetik und Sprachliche Kommunikation, Munich), pp. 29–32.
- Löfqvist, A. (2003). "Control of closure/constriction duration in lingual consonants," *J. Acoust. Soc. Am.* 114, 2397(A).
- Löfqvist, A. (2004). "Making a vocal tract closure longer and shorter," in *From Sound to Sense: Fifty+ Years of Discoveries in Speech Communication*, edited by J. Slifka, S. Manuel, and M. Matthies (Res. Lab. Electronics, MIT, Cambridge, MA), pp. C169–C174.
- Löfqvist, A., and Gracco, V. (1997). "Lip and jaw kinematics in bilabial stop consonant production," *J. Speech Lang. Hear. Res.* 40, 877–893.
- Löfqvist, A., and Gracco, V. (1999). "Interarticulator programming in VCV sequences: Lip and tongue movements," *J. Acoust. Soc. Am.* 105, 1864–1876.

- Löfqvist, A., and Gracco, V. (2002). "Control of oral closure in lingual stop consonant production," *J. Acoust. Soc. Am.* **111**, 2811–2827.
- Munhall, K., Hawkins, S., Fowler, C. A., and Saltzman, E. (1992). "Compensatory shortening' in monosyllables of spoken English," *J. Phonetics* **20**, 225–239.
- Ostry, D., Keller, E., and Parush, A. (1983). "Similarities in the control of speech articulators and the limbs: Kinematics of tongue dorsum movements during speech," *J. Exp. Psychol. Hum. Percept. Perform.* **9**, 622–636.
- Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., and Jackson, M. (1992). "Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements," *J. Acoust. Soc. Am.* **92**, 3078–3096.
- Perrier, P., Payan, P., Zandipour, M., and Perkell, J. (2003). "Influence of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study," *J. Acoust. Soc. Am.* **114**, 1582–1599.
- Rubin, P., and Löfqvist, A. (1996). "HADES: Haskins Analysis Display and Experiment System," *Haskins Laboratories Status Report on Speech Research* (available at www.haskins.yale.edu/HASKINS/SR/sr.html).
- Saltzman, E., and Munhall, K. (1989). "A dynamical approach to gestural patterning in speech production," *Ecological Psychol.* **1**, 333–382.
- Smith, C. L. (1995). "Prosodic patterns in the coordination of vowel and consonant gestures," in *Laboratory Phonology IV: Phonology and Phonetic Evidence*, edited by B. Connell and C. Arvaniti (Cambridge U.P., Cambridge), pp. 205–222.
- Vatikiotis-Bateson, E., and Kelso, J. A. S. (1993). "Rhythm type and articulatory dynamics in English, French, and Japanese," *J. Phonetics* **21**, 231–265.
- Vaxelaire, B. (1995). "Single vs. double (abutted) consonants across speech rate: X-ray and acoustic data for French," *Proc. XIII Int. Conf. Phonetic Sci.*, Stockholm, Vol 1, pp. 384–387.
- Westbury, J. (1994). "On coordinate systems and the representation of articulatory movements," *J. Acoust. Soc. Am.* **95**, 2271–2273.