



Published in final edited form as:

*J Acoust Soc Am.* 1999 March ; 105(3): 1864–1876.

## Interarticulator programming in VCV sequences: Lip and tongue movements

Anders Löfqvist<sup>a)</sup> and Vincent L. Gracco

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

### Abstract

This study examined the temporal phasing of tongue and lip movements in vowel–consonant–vowel sequences where the consonant is a bilabial stop consonant /p, b/ and the vowels one of /i, a, u/; only asymmetrical vowel contexts were included in the analysis. Four subjects participated. Articulatory movements were recorded using a magnetometer system. The onset of the tongue movement from the first to the second vowel almost always occurred before the oral closure. Most of the tongue movement trajectory from the first to the second vowel took place during the oral closure for the stop. For all subjects, the onset of the tongue movement occurred earlier with respect to the onset of the lip closing movement as the tongue movement trajectory increased. The influence of consonant voicing and vowel context on interarticulator timing and tongue movement kinematics varied across subjects. Overall, the results are compatible with the hypothesis that there is a temporal window before the oral closure for the stop during which the tongue movement can start. A very early onset of the tongue movement relative to the stop closure together with an extensive movement before the closure would most likely produce an extra vowel sound before the closure.

### INTRODUCTION

Interarticulator programming, i.e., the temporal and spatial coordination of different articulators during speech, has been the focus of much work in speech motor control. One example of this line of research is the coordination of oral and laryngeal articulatory events in the production of voiceless consonants (e.g., Gracco and Löfqvist, 1994; Löfqvist, 1980; Löfqvist and Yoshioka, 1981; Löfqvist and Yoshioka, 1984). In the production of these sounds, a glottal abduction/adduction gesture occurs while a closure or constriction is made in the oral cavity. Variations in the temporal programming of the oral and laryngeal gestures are used to produce contrasts of voicing and aspiration (e.g., Löfqvist, 1995). The phasing of the oral and laryngeal articulations is also important for the management of air pressure and air flow (e.g., Koenig *et al.*, 1995; Löfqvist *et al.*, 1995; Löfqvist and McGowan, 1992; McGowan *et al.*, 1995). Here, an increase in oral air pressure is made to drive the noise source. Another example of interarticulator programming that has received particular attention is the control of velar movements in the production of nasal consonants (e.g., Bell-Berti and Krakow, 1991; Clumeck, 1976; Kollia *et al.*, 1995; Krakow, 1989). Also in this case, the timing of the velar movements relative to other oral articulators is important for producing the correct sound sequence. The timing patterns between the velar and other articulatory movements appear to vary across languages, in particular with respect to the presence or absence of nasal vowels in a language.

This paper presents a study of another case of interarticulator programming, i.e., the timing of lip and tongue movements in sequences with a vowel, a bilabial stop consonant, and a vowel.

In producing such a sequence, a speaker has to do two or three tasks: move the tongue between the positions for the first and second vowels; close and open the lips for the stop consonant; if the stop is voiceless, open and close the glottis. Here, the focus is on the temporal coordination of the two first tasks. In addition, the kinematics of the tongue movement between the two vowels is evaluated.

Although there are acoustic studies of such VCV sequences (e.g., Öhman, 1966; Magen, 1997) that have examined the anticipatory coarticulation from the second vowel to the first, only a few published studies appear to have addressed the timing of lip and tongue movements using records of articulator motion. Houde (1968), in an x-ray study of one subject, noted that there was a tendency “for the lip to begin its vertical transition to closure before the tongue points begin their transitions from their initial target positions to the next target position” (Houde, 1968, p. 61). He also added, however, that the evidence was meager and that it was hard to establish the time at which a transition began. Gay (1977), also using x rays and two subjects, made a much stronger and precise statement: “The movement of the tongue body from the first to the second vowel does not begin until after closure for the intervocalic consonant is completed” (Gay, 1977, p. 187). However, an earlier study of Gay (1974) appears to suggest that the tongue movement does in fact start before the oral closure (see Fig. 6 in Gay, 1974). Lubker *et al.* (1977) examined the hypothesis that there was a fixed relationship between the onsets of the tongue movement and the rounding of the lips for a vowel. Using x ray, they studied one subject and concluded that their notion of synchronous programming between the tongue and the lips was not supported by the data.

Alfonso and Baer (1982) combined x ray, electromyography, and acoustic/perceptual analyses to study tongue movements in the sequence /əpVp/ produced by a single subject, where the vowel (V) was one of several American English vowels. Since they were making frame-by-frame measurements of the x-ray films, they measured the horizontal and vertical movements separately. Their results suggested that the horizontal and vertical movements of the tongue dorsum start at different times during the sequence (see Figs. 10 and 11 in Alfonso and Baer, 1982). The vertical tongue movement trajectories for the different vowels began to diverge after the oral closure for the first stop consonant, while the horizontal trajectories diverged during the first schwa vowel /ə/. The perceptual results also showed that listeners could identify the upcoming vowel from the schwa vowel. Of interest to note in their results is that the magnitude of the horizontal tongue dorsum movement was much larger than the vertical movement for this subject. Thus, the timing difference may be related to the movement magnitude.

The purpose of this study was thus to examine the timing of lip and tongue movements using a larger number of subjects, a more varied speech material, and a state-of-the-art movement transduction system. In addition, tongue movement kinematics is evaluated as a function of consonant voicing and vowel context. Since we have recently conducted a detailed study of lip and jaw kinematics in bilabial stop consonant production (Löfqvist and Gracco, 1997), the kinematics of lip and jaw movements is not specifically addressed here.

The timing of the lip and tongue movements in the present study can appropriately be discussed within the theoretical framework of anticipatory coarticulation. The production of a bilabial stop consonant does not in itself require any tongue movements, although they may occur as a result of jaw movements made for the closure and release of the stop consonant. Thus, the tongue movement can, in principle, start anywhere before, or during, the stop closure. Studies of anticipatory coarticulation have been particularly concerned with how long in advance of an upcoming segment its acoustic and/or articulatory properties can be anticipated. For example, several studies of the anticipatory coarticulation of lip rounding have been made. Some of these studies (e.g., Benguerel and Cowan, 1974; Lubker, 1981; Sussman and

Westbury, 1981) suggested that the onset of lip rounding started earlier as the number of consonants before the rounded vowel increased, supporting what has been called a look-ahead model of coarticulation. Other studies (e.g., Bell-Berti and Harris, 1979, 1981, 1982) showed that the onset of lip rounding was constrained to start within a relatively fixed temporal window, supporting what has come to be known as a frame mode of coarticulation. Later, some of the early studies were questioned since they had failed to use the proper control conditions of placing the consonants between unrounded vowels. When such control conditions are included, the extent of anticipatory lip rounding appears to be limited. In addition, recent studies (e.g., Perkell and Matthies, 1992) have suggested that both an unconstrained look-ahead model and a constrained frame model of anticipatory labial coarticulation are untenable and should be replaced by hybrid models. In a hybrid model, the onset of the lip rounding gesture for a vowel follows the look-ahead model, while the maximum acceleration of the rounding movement obeys a frame model. Since there is considerable variability between subjects in their patterns of coarticulation, Abry and Lallouache (1995) propose that anticipatory labial coarticulation should be modeled separately for each subject. Earlier work on anticipatory coarticulation of lip rounding tended to focus only on factors of motor control. More recent work also suggests that perceptual factors may well play a role in limiting the degree of coarticulation to maintain the integrity of the speech signal.

It is not entirely clear what such models of anticipatory coarticulation predict for the onset of the tongue movement from the first to the second vowel in a VCV sequence with a labial stop consonant. One methodological problem is that it is not easy to define a boundary between the two vowels based on acoustic or articulatory records. Studies of anticipatory coarticulation usually use an acoustically defined landmark, such as the onset of a vowel, as a base for the measurements. In an articulatory sense, such a point is arbitrary, since the tongue movement for the vowel starts before such an acoustically defined point. In addition, the tongue movement between the two vowels in a VpV sequence is continuous. One might argue that the onset of the tongue movement should be constrained in the sense that the speaker avoids to start it so much in advance of the stop consonant that an “extra” vowel is produced and perceived. A perceptual study by Carré *et al.* (1996), using an acoustic model of the vocal tract, showed this to happen when the movement to the second vowel started very much in advance of the bilabial stop closure. Based on this hypothesis, one might further hypothesize that in those instances where the tongue movement onset occurs well in advance of the lip closure for the stop consonant, the magnitude of the tongue movement should be small, so as not to introduce an extra segment. To test this hypothesis, the present study examined the relationship between the interval from the onset of the tongue movement from the first to the second vowel to oral closure for the stop and the magnitude of the tongue movement from its onset to the stop closure. If this hypothesis were correct, we would expect a negative correlation between these two variables, i.e., a tongue movement starting well before the stop closure should show a small displacement before the oral closure. Another hypothesis would relate the onset of the tongue movement from the first to the second vowel to the magnitude of the tongue movement trajectory itself. Based on this hypothesis, one might expect a positive correlation between the interval from tongue movement onset to lip closing onset and the size of the tongue movement trajectory between the two vowels.

To examine these specific hypotheses and other factors influencing the tongue movement trajectory, measurements were made of the onset of the tongue movement from the first to the second vowel relative to the onset of the lip closing movement for the stop consonant and also relative to the acoustically defined oral closure for the stop. The magnitude of the tongue movement trajectory was also measured to see if it was related to the timing of the tongue and lip movement. In addition, the duration of the tongue movement trajectory was measured to examine the relationship between movement amplitude and duration. Finally, a calculation was made of the percentage of the tongue movement from the first to the second vowel that

occurred during different intervals of the VCV sequence, such as during the first vowel, during the stop closure, and during the second vowel.

## I. METHOD

### A. Subjects

Two female (LK, DR) and two male subjects (VG, AL) participated. All subjects had normal speech and hearing and no history of speech or hearing disorders. Three of the subjects (LK, DR, VG) are native speakers of American English. Subjects LK and DR grew up in the midwest, while subject VG grew up in Florida; they all currently live in the northeast. Speaker AL is a native speaker of Swedish who is also fluent in English. Subjects VG and AL are the two authors.

### B. Linguistic material

The linguistic material consisted of  $V_1CV_2$  sequences, where the first and second vowels ( $V_1$  and  $V_2$ ) were always one of /i, a, u/, and the consonant (C) one of /p, b/. The sequences were placed in the carrier phrase “Say ... again” with sentential stress occurring on the second vowel ( $V_2$ ) of the sequence. Ten repetitions of each sequence were recorded. Only the sequences with asymmetric vowel contexts were included in the analysis, since it is virtually impossible to define movement onsets and offsets in sequences with symmetrical vowel contexts due to the small and inconsistent amount of tongue movement.

### C. Procedure

The movements of the lips, the jaw, and the tongue were recorded using a three-transmitter magnetometer system (Perkell *et al.* 1992). Receivers were placed on the upper and lower lips, on the lower incisors, and on four positions on the tongue. The tongue receivers will be referred to as tongue tip, tongue blade, tongue body, and tongue rear. The lip receivers were placed below and above the vermilion border of the upper and lower lip, respectively, with a vertical separation of approximately 1 cm when the lips were in a closed position. For the tongue, the first receiver was placed as far back as the subject could tolerate, and the second one close to the tongue tip; next, an attempt was made to space the other two receivers evenly between the first and the second. Two additional receivers placed on the nose and the upper incisors were used for the correction of head movements. The receivers on the lips, the incisors, and the nose were attached using Iso-Dent (Ellman International). For the tongue receivers, Ketac-Bond (ESPE) was used. 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 during 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 articulatory movement signals (induced voltages from the receiver coils) were sampled at 625 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 17 Hz; this was done forwards and backwards to maintain phase. 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. A measure of lip opening was obtained by subtracting the vertical position of the lower lip receiver from that of the upper lip receiver. 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 9.5 kHz, and sampled at 20 kHz.

To define the onset of the closing movement of the lips for the stop consonant, the second derivative of the derived lip opening signal was used. Using a zero crossing in the first derivative of the lip opening signal was difficult when the first vowel was /u/ that included lip rounding. Here, the rounding gesture made the lip opening change continuously and a zero crossing would not appear in the first derivative at a point in time close to the oral closure. Thus, the onset of the lip closure for the stop consonant was defined as the minimum in the second derivative of the lip opening signal prior to the oral closure, cf. Fig. 1. This point was defined algorithmically. Figure 1 presents the acoustic, lip opening, and tongue body signals for one production of the sequence “api” by subject VG. Since the interpretation of a second derivative is not always straightforward, the lip opening signal and its first derivative are also included in Fig. 1. We should add that the actual lip opening is at zero throughout the oral closure. The change in the lip opening signal during the closure is due to the fact that it represents the vertical distance between the receivers on the upper and lower lips, and these receivers move during the closure (cf. Löfqvist and Gracco, 1997; Westbury and Hashi, 1997). The tongue body receiver was used for analyzing tongue kinematics. Its tangential

velocity  $\left[ v = \sqrt{\dot{x}^2 + \dot{y}^2} \right]$  was calculated. Tongue movement onsets and offsets were identified algorithmically from the tangential velocity signal as minima during the first and second vowels. Their identification is also shown in Fig. 1. We should note that at these points in time, the horizontal and vertical velocity of the tongue is not necessarily zero. The magnitude of the tongue movement trajectory from vowel to vowel was obtained by summing the Euclidean distances between successive samples from movement onset to offset. In addition, the Euclidean distance between the tongue body receiver at movement onset and offset was also measured and used to assess the extent to which the movement trajectory was a straight line or a curved path. This was done by calculating the ratio between the tongue movement measured as a trajectory and as a straight line. A ratio of 1 indicates that the movement trajectory follows a straight line path while a ratio greater than 1 shows that the trajectory is a curved path.

The onset and release of the oral closure were identified in waveform and spectrogram displays of the acoustic signal. The onset of the closure was identified by the decrease in the amplitude of the acoustic waveform, and by the disappearance of spectral energy at higher frequencies. The release was identified by its burst. The onset of regular glottal vibrations for the second vowel was also marked. Measurements of closure duration and of voice onset time were made. All the labeling was made by the first author. Analyses of variance and *t*-tests were used to assess the influence of vowel context and consonant voicing on timing and movement parameters. A *p* value of  $\leq 0.05$  was adopted as significant.

The kinematic signals represent the movements of receivers placed at the midline of the lips, the jaw, and the tongue. When presenting the results, we will use the terms “tongue body receiver” and “tongue body” interchangeably, while acknowledging that we are only examining the movements of a single point. Thus, we make no claims about asymmetrical movements of the left and right sides of the lips, or the tongue. The tongue and lower lip signals contain the contribution of the jaw. These are the appropriate movements to examine when the focus of the analysis is on the lower lip and the tongue as end effectors. It is reasonable to assume that a speaker has joint control of different articulators during speech production to produce the desired results.

## II. RESULTS

Figure 2 shows receiver movement trajectories from the first to the second vowel for all subjects and sequences with a voiced stop /b/. The whole trajectory from vowel to vowel is shown as well as the part of the trajectory that occurred during the oral closure for the consonant. The

letters at the onset/offset of each trajectory identify the positions for the respective vowels. A tracing of the outline of the hard palate is also shown for identification purposes. Since this outline was obtained by having the subject move the tongue tip receiver from the alveolar ridge and as far backwards as possible, these tracings do not necessarily give the true outline of the palate in the posterior region. In these figures, a few facts should be noted. The tongue receivers can move in both straight-line and curved paths. An analysis of the shape of the trajectories for the tongue body receiver was made by calculating the ratio between the movement amplitudes measured as the actual path length and as the Euclidean distance between the receiver positions at movement onset and offset. For all subjects, this ratio tended to be highest for the sequences with /i/ and /u/ as the two vowels, with values of 1.2 or higher. Only in one case, the sequence /abi/ for subject LK, did the tongue body receiver move in a straight line. The trajectory for a pair of VCV sequences with the same vowels but in different positions, such as /abi/ and /iba/, do not show paths that are mirror images of each other. This is due, in part, to the fact that the context for the first and second vowels differ due to the carrier phrase used, and also to the stress pattern. The trajectories of the four tongue receivers show both similarities and differences. For example, subject LK shows very similar trajectories for the tongue tip, tongue blade, and tongue body receivers, but a different pattern for the tongue rear receiver. In particular, for the sequence /ibu/ of her productions, the tongue rear receiver is moving backward while the other three tongue receivers are moving backward and downward. Similarly, in the sequence /ubi/ the rear tongue receiver is moving backward and up while the other tongue receivers are moving forward and up. Similar differences can also be found in the productions of subject DR. On the other hand, subject AL shows more of a similarity in the trajectories of the tongue receivers.

The first analysis focused on the onset of the tongue body movement from the first to the second vowel relative to the oral closure for the stop consonants. The results are shown in Fig. 3. This figure plots frequency distributions of the interval between the onset of the tongue body movement and the acoustically defined oral closure, collapsed across consonant voicing and vowel contexts ( $n=120$ ). The data have been grouped into 25-ms bins. It is evident from these results that they do not agree with the findings of Gay (1977). Only in 7% of the productions for subject LK did the tongue movement start after the oral closure for the consonant. For the remaining three subjects, the tongue movement always started before the closure. The interval between the tongue movement onset and the oral closure ranged up to 175 ms for subjects LK and DR, whereas it was more narrowly distributed for the two remaining subjects, VG and AL. A separate  $t$ -test for each subject showed that stop consonant voicing had no reliable influence on the interval between the onset of the tongue movement and the lip closure (pooled across vowel contexts) for subjects LK, VG, and AL ( $t_{118}=-0.17, -1.78, \text{ and } -1.42$ , respectively). Voicing had a reliable effect for DR ( $t_{118}=-3.87$ ), but the difference was only 7 ms. An analysis of variance performed for each subject showed that the quality of the first vowel had a reliable influence on the interval between the onset of the tongue movement and the oral closure ( $F_{2,114}=15.69, 26.84, 123.89, 303.70$ , for subjects LK, DR, VG, and AL, with corresponding  $\eta^2$  values of 0.19, 0.32, 0.66, and 0.83). However, the pattern varied between subjects. For subjects DR, VG, and AL, the interval decreased in the order /i/ > /a/ > /u/; the average values in ms were 95, 75, and 70 for DR, 65, 40, and 25 for VG, and 95, 70, and 45 for AL. For subject LK, this interval was always shortest when the first vowel was /a/, 15 ms compared to 60 and 70 ms for /i/ and /u/.

The next analysis examined the phasing between the onset of the tongue body movement and the onset of the lip closing movement. The results are summarized in Fig. 4 for all subjects and sequences. A first thing to note in this figure is that subject DR always started the tongue movement before the lip movement. For the other three subjects, the pattern varies, although they all show the tongue movement leading when the first vowel is /i/. For subjects VG and AL, the lip movement tended to lead in the other vowel contexts. The same thing was true for

subject LK, except that in the sequences /upa/ and /uba/ the tongue movement started approximately 100 ms before the lip movement. With the exception of these two sequences for subject LK, the interval between the tongue and lip movement onsets was less than 50 ms for all subjects and sequences. A *t*-test for each subject's productions showed that stop consonant voicing had no reliable effect on the interval between tongue and lip movement onsets for subjects LK, VG, and AL ( $t_{118}=0.67, -1.55, \text{ and } -1.16$ ). For subject DR, there was a very small but reliable difference of 10 ms due to voicing ( $t_{118}=-2.53$ ).

The magnitude of the tongue movement trajectory between the first and second vowels is shown in Fig. 5. Overall, the magnitude was larger for subject AL than for the other subjects. For DR and VG, the trajectory was usually less than 1.5 cm. There was no reliable influence of consonant voicing on the trajectory for subjects LK, DR, and VG ( $t_{118}=0.47, 0.25, 1.15$ ). Subject AL always had a longer trajectory when the consonant was voiced ( $t_{118}=-4.18$ ). The average difference was 0.3 cm. The effect of vowel context did not appear to be consistent between subjects. The only apparent pattern was that subjects LK, DR, and VG showed small trajectories for the sequences /upi/ and /ubi/.

There is, however, another influence of vowel context on the magnitude of the tongue movement that is evident in Fig. 5. Mirror vowel sequences do not show the same magnitude of the tongue movement. For example, it is always the case that the tongue movement is larger in the sequences /ipa, iba/ than in the sequences /api, abi/. Similarly, the magnitude is larger in the sequences /ipu, ibu/ than in /upi, ubi/. Finally, the magnitude is larger for /upa, uba/ than for /apu, abu/ for subjects LK and DR. The reason for these differences can be found in Fig. 6. This figure plots the tongue body receiver positions for the three vowels /i, u, a/ when they occur as the first and second vowel in the VCV sequence. The measurements were made at the minimum tangential velocity shown in Fig. 1. Two findings emerge in this figure. First, the filled symbols (circles and triangles) for the sequences with /a/ and /u/ as the first vowel are higher and more forward in the plots than the unfilled symbols for the same vowels when they occur as the second vowel. Second, there is much less difference in the tongue body position for the vowel /i/ when it is the first or the second vowel in the sequence. Thus, if we compare the sequences /ipa, iba/ with the sequences /api, abi/, we see that the position of the tongue body for the vowel /a/ is lower and further back in /ipa, iba/ than it is in /api, abi/, see the unfilled and filled triangles. The same is true for the sequences /ipu, ibu/ and /upi, ubi/, where the tongue body position for the vowel /u/ is different when it occurs as the second and first vowels (see the unfilled and filled circles). Thus, the tongue body position for /a/ and /u/ differs when these vowels occur as the first and second vowel of the sequence. However, the filled and unfilled squares associated with the vowel /i/ occurring as the first and second vowel overlap for subjects VG and AL, and show a relatively small difference for subjects LK and DR. The variations in tongue body position for the vowels /a, u/ account for the difference in the magnitude of the tongue movement in sequences with mirror vowels. The movement trajectories shown in Fig. 2 also show similar differences in the tongue body position for the same vowel when it is the first and the second vowel in the sequence.

We next turn to the relationship between the phasing of lip and tongue movements and the duration and magnitude of the tongue movement trajectory. One hypothesis proposed in the Introduction is that the tongue movement from the first to the second vowel starts earlier relative to the onset of the lip closing movement as the magnitude of the tongue movement trajectory increases. The relevant results are presented in Fig. 7 showing scatter plots of the interval between the onsets of the lip closing movement and the tongue body movement versus the tongue body movement trajectory. Overall there is a positive correlation between the onset and the magnitude of the tongue movement. That is, the onset of the tongue movement tended to start earlier relative to the lips when the tongue body movement is large. The correlation coefficients for all productions, pooled across vowel contexts and consonant voicing ( $n=120$ ),

were 0.69, 0.43, 0.53, and 0.50 for subjects LK, DR, VG, and AL, respectively. In contrast, there was no significant correlation between the onset of the tongue movement in relation to the onset of the lip closing movement and the duration of the oral closure for any subject.

If the tongue movement starts well in advance of the lip closing movement, and hence well in advance of the oral closure for the consonant, an extra vowel segment might be perceived. One might thus hypothesize that in those cases where the tongue movement started well before the oral closure, the movement up to the oral closure would be rather small, so as not to cause any perceptual effects. That is, one would expect a negative correlation between the onset of the tongue movement relative to the oral closure and the magnitude of the tongue movement displacement from its onset to the oral closure. Such a negative correlation was only found for one subject, VG, however, and it was not significant ( $r=-0.15$ ). For the other three subjects, the correlations were positive but small ( $r=0.1$  (ns), 0.38, and 0.27, for LK, DR, and AL, respectively).

Two of the subjects, LK and DR, showed a positive correlation between the tongue movement trajectory and its duration, with  $r=0.81$  and 0.73. For the two other subjects, these correlation coefficients were very low,  $-0.10$  (ns) and 0.29 for VG and AL, respectively. These results are shown in Fig. 8. The duration of the tongue movement did not correlate with the duration of the oral closure for subjects LK and DR. For subject VG, there was positive correlation of 0.41, whereas for subject AL there was a reliable negative correlation of  $-0.2$ .

The final analysis of tongue–lip phasing focused on how much of the tongue movement occurred during four different parts of the VCV sequence. These parts were (1) from the onset of the tongue movement to the oral closure for the stop; (2) during the oral closure; (3) between the stop release and the onset of the second vowel; and (4) during the second vowel. For each of these parts, the tongue movement was calculated as a percentage of the total movement trajectory. The results are shown in Fig. 9. This figure shows the cumulative percentage of the tongue movement during the four parts of the sequence, shown in their order of occurrence from bottom to top in the figure. That is, the relative movement from the onset of the tongue movement to the oral closure for the stop is shown at the bottom. Overall, the black portion of each bar is the largest. This indicates that most of the tongue movement took place during the oral closure for the stop. For subjects LK, VG, and AL 55%, 60%, and 57% of the total trajectory took place during the closure; for subject DR, the corresponding value was 35%.

As might be expected, consonant voicing had a very robust effect on the relative tongue movement trajectory between the release of the consonant and the onset of the vowel. It was much larger for the voiceless stops: 18%–50% vs 5%–12% for LK, 24%–45% vs 4%–9% for DR, 12%–29% vs 3%–13% for VG, and 10%–21% vs 4%–7% for AL. This was related to the difference in voice onset time between voiced and voiceless stops, which was also significant for all subject. However, the duration of the oral closure did not differ as a function of consonant voicing for any subject. Figure 9 also suggests that the relative tongue movement during the first vowel up to the oral closure tends to be larger for voiced than for voiceless stops. A test performed for each subject showed this difference to be reliable for subjects LK, DR, and VG ( $t_{118}=13.5$ , 8.63, and 31.4), but not for subject AL ( $t_{118}=0.69$ ). The values for the voiced and voiceless contexts were 27% vs 5%, 35% vs 14%, and 16% vs 10% for subjects LK, DR, and VG, while for subject AL the value was the same, 9%, for both contexts. Apart from these general influences of stop consonant voicing, there was no consistent pattern across speakers for vowel context.

Finally, the tongue body position and velocity at the onset of the movement and at the oral closure for the stop was examined for the influence of the upcoming vowel. As can be expected, this influence increased between the first and the second point in time. Although the effect of



the second vowel was in many cases reliable at both points in time, an examination of the  $\eta^2$  values showed them to increase from the first to the second point. This was particularly the case for the velocities, since at the instant of oral closure the tongue was moving in different directions depending on the identity of the second vowel. It proved difficult to assess possible differences in the timing of the horizontal and vertical component of the tongue body movement, however. The reason was that in many cases the tongue body was moving very slowly, which made the identification of zero crossings for movement onset unreliable. In those instances where it was possible to reliably identify the movement onset, the interval between the onset of the vertical and horizontal movement of the tongue body was 50 ms or less, with no clear pattern of one component consistently leading the other.

### III. DISCUSSION

The present results show that the tongue movement from the first to the second vowel in a sequence of vowel-bilabial stop-vowel most often started before oral closure for the stop consonant. In only a small percentage of the productions of one subject (LK) did the tongue movement start after the consonant closure (Fig. 3). This is contrary to the results presented by Gay (1977). Subjects differed in the details of their patterns of lip-tongue phasing. For two of the subjects, LK and DR, the tongue movement could start up to 175 ms before the closure for the consonant. The timing of the tongue movement with respect to the closing movement of the lips for the consonant also differed between subjects. One subject, DR, always started the tongue movement before the onset of the lip movement, irrespective of the nature of the first and second vowels (Fig. 4). For the other three subjects, the pattern of lip-tongue phasing varied. For example, all three of them, LK, VG, and AL, always had the tongue movement leading the lip movement when the first vowel was /i/. In the remaining sequences, subjects VG and AL always showed the lips leading the tongue, whereas subject LK had the tongue leading the lips in the sequences where the first vowel was /u/ and the second vowel was /a/. Generally, the interval between the tongue and lip movement onsets was 50 ms or less. Taken together these results suggest that there is a temporal window during which the tongue movement from the first to the second vowel can start. However, the size and location of this window is different for different subjects. In this respect, the present results are compatible with the proposal by Abry and Lallouache (1995) that anticipatory coarticulation should be modeled as a continuous variable that has to be adjusted for individual subjects.

It is thus apparent from the present results that speakers differ in their patterns of interarticulator programming. For example, one would draw a quite different conclusion about the phasing of the lip and tongue movements if one only studied the results for subject DR in Fig. 4, since she always had the tongue movement leading the lip movement. Similarly, by only focusing on the subset of the linguistic material where the first vowel was /i/, one would arrive at a different conclusion than if all the vowel contexts were examined.

In spite of this variability between subjects, there were several common patterns. The results shown in Fig. 5 indicate that the tongue movement trajectory was larger in the sequences /ipa, iba/ than in /api, abi/. Differences in the tongue body movement onset and offset positions for the same vowel occurring in the first or second position can account for these differences, as shown in Fig. 6. It would appear that the stress pattern of the word is mostly responsible for this difference. The tongue body positions are most different for the vowels /a, u/ when they occur as the first and second vowel in the VCV sequence but very small, or non-existent, for the vowel /i/. Most likely, the increased duration of the second vowel due to the stress pattern can explain the differential variability in tongue position for /a, u/ compared to /i/. In addition, the high front tongue position for the diphthong /ai/ in the carrier sentence might contribute to the small difference in tongue body position for the vowel /i/.

For all of the subjects, there was a clear tendency for the onset of the tongue movement to occur earlier with respect to the onset of the lip closing movement as the tongue movement trajectory increased (see Fig. 7). This is similar to the finding of Alfonso and Baer (1982) for the vertical and horizontal components of the tongue dorsum movement. As explained above, it proved difficult in the present experiment to consistently identify movement onsets from zero crossings in the horizontal and vertical velocity signals of the tongue body. When a movement onset could be identified, the interval between the onset of the vertical and horizontal component was 50 ms or less.

Although the onset of the tongue body movement is free to vary within certain limits, one might argue that the onset of the lip closing movement would be more constrained. This was indeed the case. Obviously, the onset of the lip closing movement must precede the oral closure and it never occurred more than 100 ms before the acoustically defined oral closure. An examination of the relationship between the onsets of the lip closing movement and the tongue movement relative to the oral closure, pooled across consonant voicing and vowel contexts, showed that they tended to covary for three of the subjects, DR, VG, and AL ( $r=0.47$ ,  $0.25$ , and  $0.44$ ), but not for subject LK ( $r=0.01$ , ns). If the tongue movement starts well before the oral closure, one might expect that the speech could be compromised, because an “extra” vowel might be perceived. Similarly, an extensive tongue movement following the release of the stop could result in an “extra” vowel after the consonant. A perceptual study by Carré *et al.* (1996), using an acoustic model of the vocal tract, showed this to be the case. Possibly, the finding in the present study that most of the tongue movement trajectory took place during the stop closure is related to avoiding such effects. On the other hand, the specific hypothesis that the amount of tongue movement during the first vowel should be small when the movement started well before the stop closure was not supported by the data. For three of the subjects, the relative magnitude of the tongue movement during the first vowel was larger when the consonant was voiced than when it was unvoiced. This may be related to the commonly observed longer acoustic vowel duration before voiced consonants. Interestingly, subject AL did not show any such difference. Possibly, the reason the three native speakers of American English showed such an effect of consonant voicing is that such voicing conditioned differences of vowel duration appear to be larger in American English than in most other languages (e.g., Chen, 1970). We should add further that, in instances where anticipatory tongue movements have been observed to span large temporal intervals (e.g., Magen, 1997), the intervening vowels have been the schwa which may not have a clear articulatory specification.

In summary, the present study has shown that the timing of lip and tongue movements in the production of VCV sequences with a bilabial stop consonant is influenced by the magnitude of the tongue movement trajectory. Overall, the results are compatible with the hypothesis that there is a temporal window before the oral closure for the stop during which the tongue movement can start. The results are also generally compatible with the idea that extensive tongue movements before and after the stop closure are avoided so as not to create perceptual effects. Thus, more than 50% of the tongue movement trajectory between the two vowels occurred during the stop closure. The measurement techniques developed for this study will also be useful in the future to address the influences of other variables on the programming of lip and tongue movements, such as stress and prosodic boundaries.

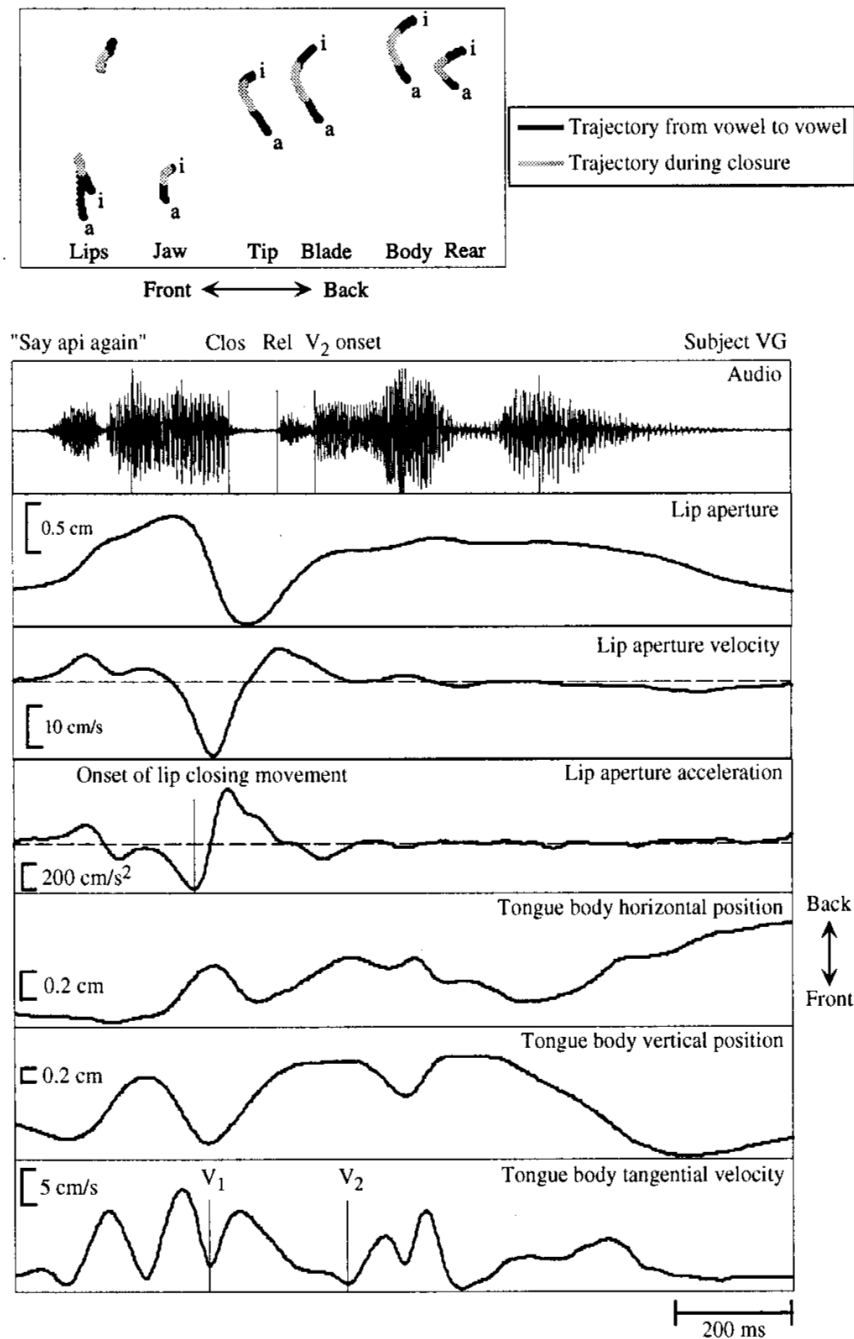
## Acknowledgments

We appreciate the comments from Winifred Strange and three reviewers on an earlier version of this manuscript. This work was supported by Grants No. DC-00865 and DC-00595 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

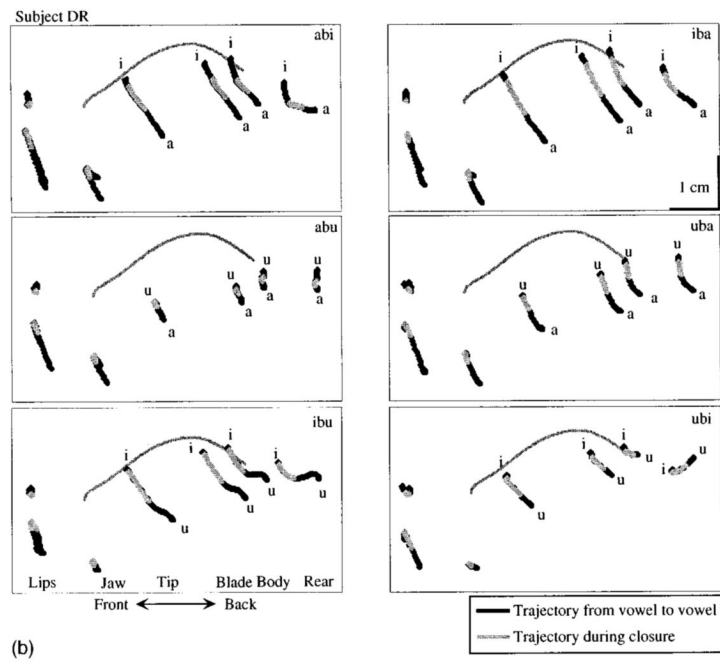
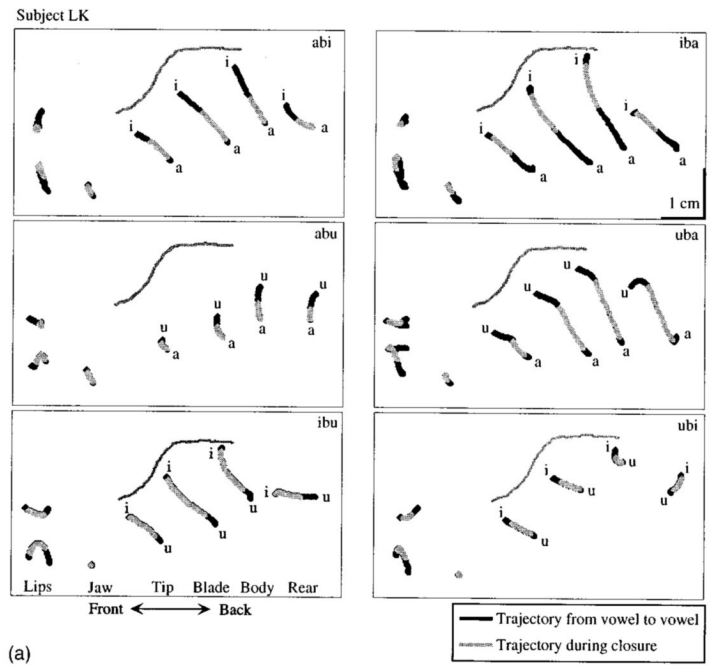
## References

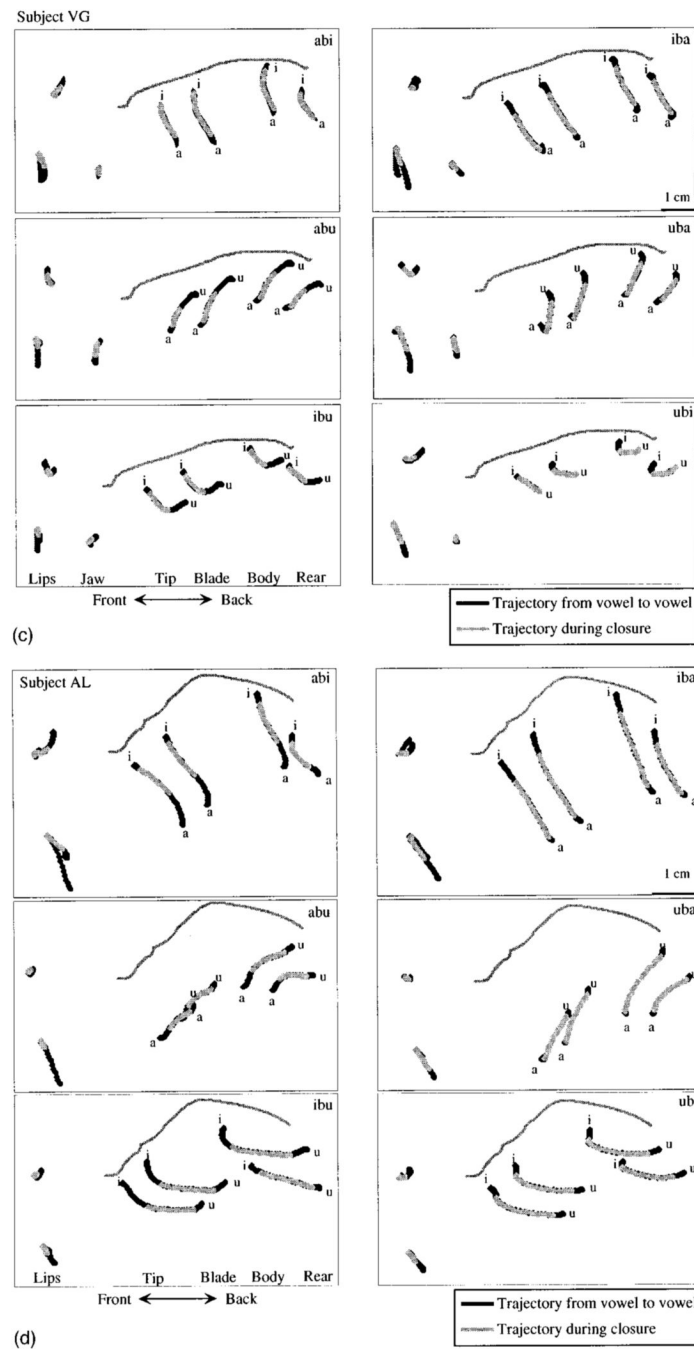
- Abry C, Lallouache T. Le MEM: un modèle d'anticipation paramétrable par locuteur, Données sur l'arrondissement en français. *Bull. Comm. Parlée* 1995;3:85–99.
- Alfonso P, Baer T. Dynamics of vowel articulation. *Language and Speech* 1982;25:151–173.
- Bell-Berti F, Harris KS. Anticipatory coarticulation: Some implications from a study of lip rounding. *J. Acoust. Soc. Am* 1979;65:1268–1270. [PubMed: 458048]
- Bell-Berti F, Harris KS. A temporal model of speech production. *Phonetica* 1981;38:9–20. [PubMed: 7267724]
- Bell-Berti F, Harris KS. Temporal patterns of coarticulation: Lip rounding. *J. Acoust. Soc. Am* 1982;71:449–454. [PubMed: 7069059]
- Bell-Berti F, Krakow RA. Anticipatory velar lowering: A coproduction account. *J. Acoust. Soc. Am* 1991;90:112–123. [PubMed: 1880279]
- Benguerel AP, Cowan HA. Coarticulation of upper lip protrusion in French. *Phonetica* 1974;30:41–55. [PubMed: 4851063]
- Carré, R.; Chennoukh, S.; Jospa, P.; Maeda, S. The ears are not sensitive to certain coarticulatory variations: Results from VCV synthesis/perceptual experiments; 1st ESCA ETRW on Speech Production Modeling and 4th Speech Production Seminar; Autrans. 1996; p. 13-16.
- Chen M. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 1970;22:129–159.
- Clumeck HA. Patterns of soft palate movement in six languages. *J. Phonetics* 1976;4:337–351.
- Gay TJ. A cinefluorographic study of vowel production. *J. Phon* 1974;2:255–266.
- Gay TJ. Articulatory movements in VCV sequences. *J. Acoust. Soc. Am* 1977;62:183–193. [PubMed: 893816]
- Gracco VL, Löfqvist A. Speech motor coordination and control: Evidence from lip, jaw, and laryngeal movements. *J. Neurosci* 1994;14:6585–6597. [PubMed: 7965062]
- Houde, R. A study of tongue body motion during selected consonant sounds. *Speech Communications Research Laboratory; Santa Barbara: 1968. SCRL Monograph 2*
- Koenig LK, Löfqvist A, Gracco VL, McGowan RS. Articulatory activity and aerodynamic variation during voiceless consonant production. *J. Acoust. Soc. Am* 1995;97:3401.
- Kollia HB, Gracco VL, Harris KS. Articulatory organization of mandibular, labial, and velar movements during speech. *J. Acoust. Soc. Am* 1995;98:1313–1324. [PubMed: 7560504]
- Krakow, RA. Unpublished doctoral dissertation. Yale University; 1989. The articulatory organization of syllables: A kinematic analysis of labial and velar gestures.
- Löfqvist A. Interarticulator programming in stop production. *J. Phon* 1980;8:475–490.
- Löfqvist, A. Laryngeal mechanisms and interarticulator timing in voiceless consonant production. In: Bell-Berti, F.; Raphael, L., editors. *Producing Speech: Contemporary Issues*. For Katherine Safford Harris. AIP Press; Woodbury, NY: 1995. p. 99-116.
- Löfqvist A, Gracco VL. Lip and jaw kinematics in bilabial stop consonant production. *J. Speech Hear. Lang. Res* 1997;40:877–893.
- Löfqvist A, McGowan RS. Influence of consonantal environment on voice source aerodynamics. *J. Phon* 1992;20:93–110.
- Löfqvist A, Yoshioka H. Interarticulator programming in obstruent production. *Phonetica* 1981;38:21–34. [PubMed: 7267720]
- Löfqvist A, Yoshioka H. Intra-segmental timing: Laryngeal coordination in voiceless consonant production. *Speech Commun* 1984;3:279–289.
- Löfqvist A, Koenig LL, McGowan RS. Vocal tract aerodynamics in /aCa/ utterances: Measurements. *Speech Commun* 1995;16:49–56.
- Lubker J, McAllister R, Lindblom B. On the notion of interarticulator programming. *J. Phon* 1977;5:213–226.
- Lubker JF. Temporal aspects of speech production: Anticipatory labial coarticulation. *Phonetica* 1981;38:51–65. [PubMed: 7267722]
- Magen H. The extent of vowel-to-vowel coarticulation in English. *J. Phon* 1997;25:187–205.

- McGowan RS, Koenig LL, Löfqvist A. Vocal tract aerodynamics in /aCa/ utterances: Simulations. *Speech Commun* 1995;16:67–88.
- Öhman S. Coarticulation in VCV utterances: Spectrographic measurements. *J. Acoust. Soc. Am* 1966;39:151–168. [PubMed: 5904529]
- Perkell J, Matthies M. Temporal measures of anticipatory labial coarticulation for the vowel /u/: Within- and cross-subject variability. *J. Acoust. Soc. Am* 1992;91:2911–2925. [PubMed: 1629484]
- Perkell J, Cohen M, Svirsky M, Matthies M, Garabiéta I, Jackson M. Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *J. Acoust. Soc. Am* 1992;92:3078–3096. [PubMed: 1474223]
- Rubin, PER.; Löfqvist, A. HADES (Haskins Analysis Display and Experiment System). Haskins Labs. Status Rep. *Speech Res.* 1996. available at [www.haskins.yale.edu/HASKINS/SR/sr.html](http://www.haskins.yale.edu/HASKINS/SR/sr.html)
- Sussman HM, Westbury JR. The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation. *J. Speech Hear. Res* 1981;24:16–24. [PubMed: 7253623]
- Westbury J. On coordinate systems and the representation of articulatory movements. *J. Acoust. Soc. Am* 1994;95:2271–2273. [PubMed: 8201122]
- Westbury J, Hashi M. Lip-pellet positions during vowels and labial consonants. *J. Phon* 1997;25:405–419.



**FIG. 1.** Acoustic and movement signals recorded during the production of the utterance “Say api again” by subject VG. In addition to the derived lip aperture signal, its first and second derivatives are shown. The labels in the audio signal correspond to the onset and release of the oral closure for the consonant, and the onset of the second vowel in /api/. The label in the lip aperture acceleration signal marks the onset of the lip closing movement for the consonant. The labels in the tangential velocity signal of the tongue body identify the onset and offset of the tongue movement from the first to the second vowel in /api/. The top panel shows the receiver trajectories from the first to the second vowel and also the part of the trajectory made during the oral closure for the consonant.

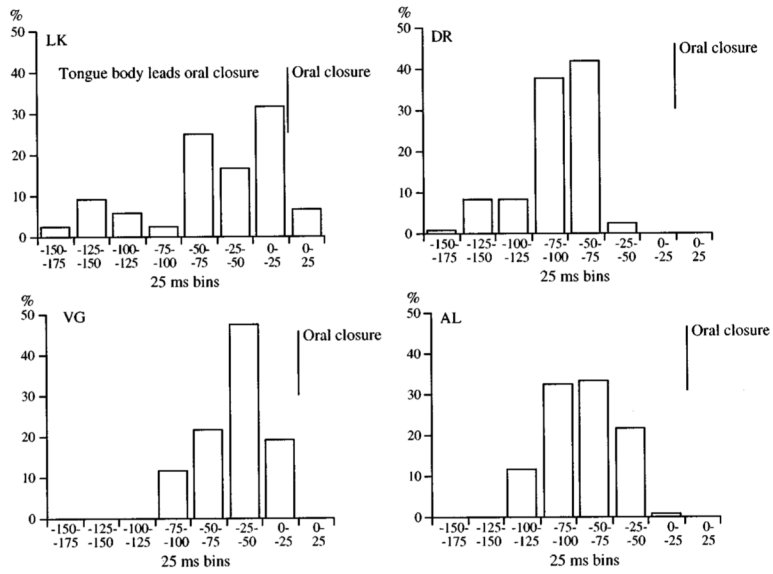


**FIG. 2.**

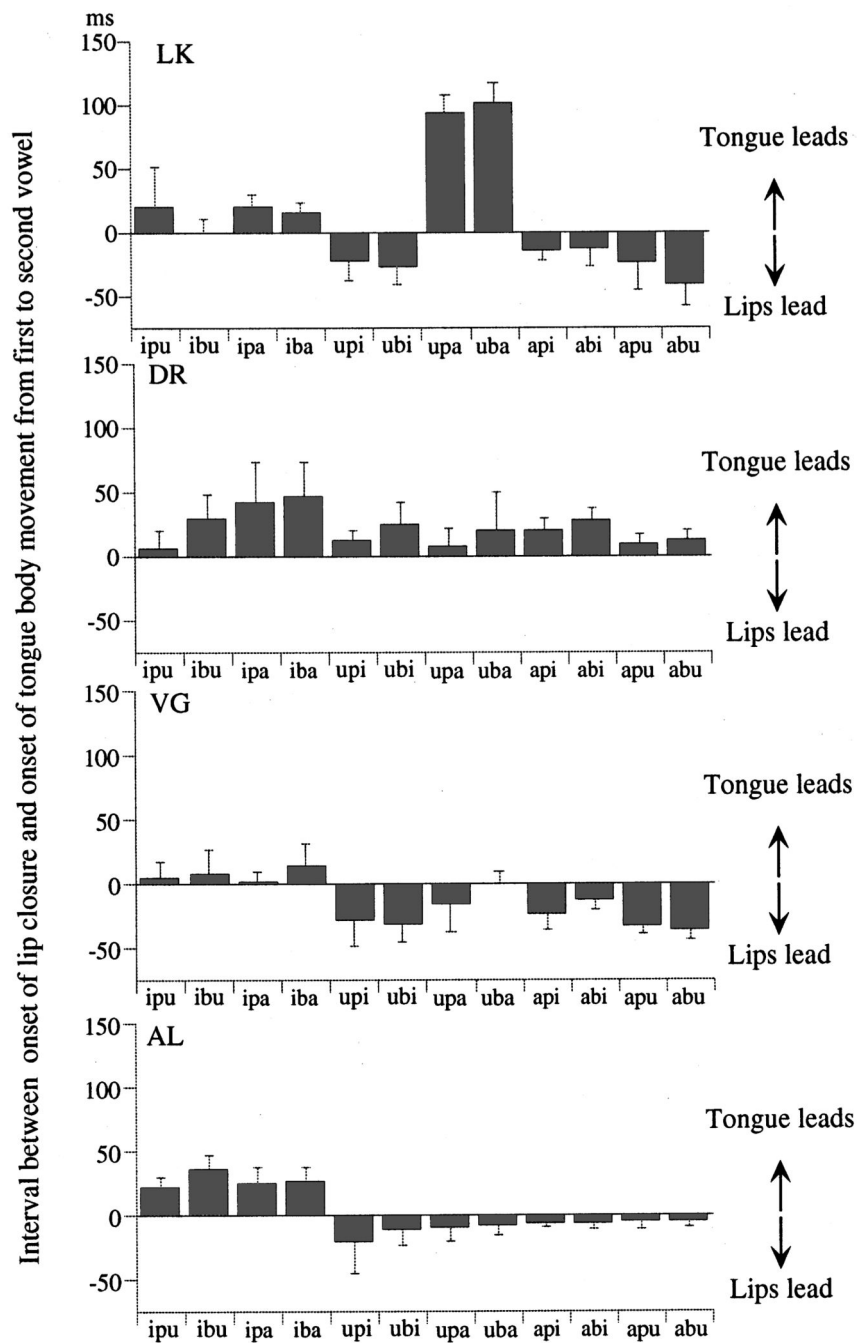
(a) Plots of receiver trajectories during single productions of sequences with a voiced bilabial stop /b/ and asymmetric vowel contexts for subject LK. The trajectory from the first to the second vowel is shown and also the part of the trajectory that occurred during the oral closure for the consonant. (b) Plots of receiver trajectories during single productions of sequences with a voiced bilabial stop /b/ and asymmetric vowel contexts for subject DR. The trajectory from the first to the second vowel is shown and also the part of the trajectory that occurred during the oral closure for the consonant. (b) Plots of receiver trajectories during single productions of sequences with a voiced bilabial stop /b/ and asymmetric vowel contexts for subject VG. The trajectory from the first to the second vowel is shown and also the part of the trajectory

that occurred during the oral closure for the consonant. (d) Plots of receiver trajectories during single productions of sequences with a voiced bilabial stop /b/ and asymmetric vowel contexts for subject AL. The trajectory from the first to the second vowel is shown and also the part of the trajectory that occurred during the oral closure for the consonant.



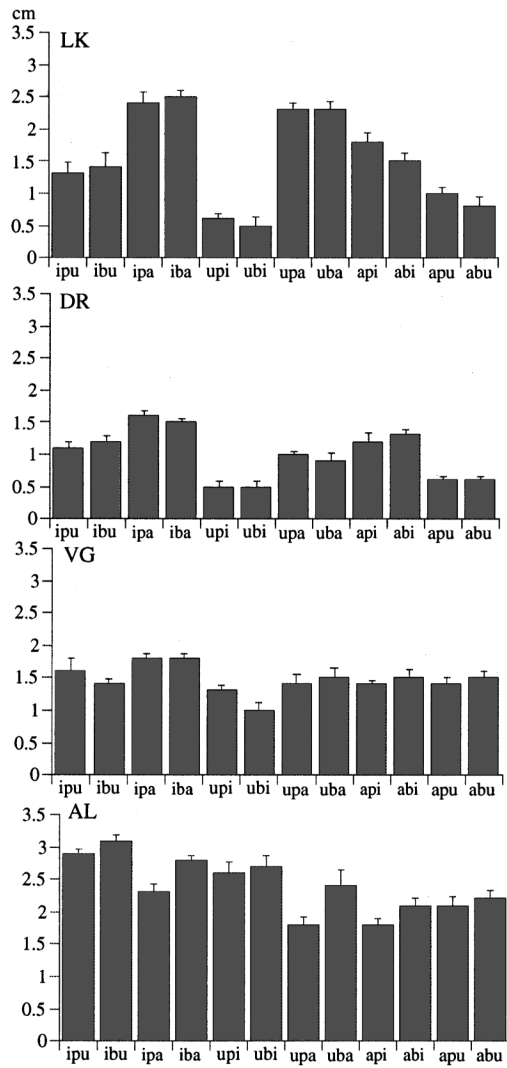


**FIG. 3.** Frequency distributions of the interval between the onset of the tongue movement from the first to the second vowel and the oral closure for the consonant. The bin size is 25 ms.

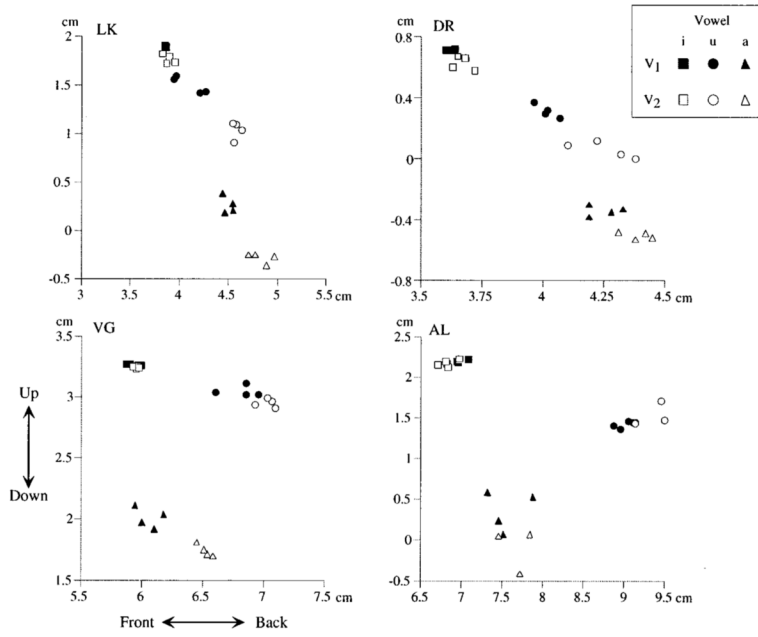


**FIG. 4.** The interval between the onset of the tongue movement from the first to the second vowel and the onset of the lip closing movement for the consonant. Means and standard deviations are shown.

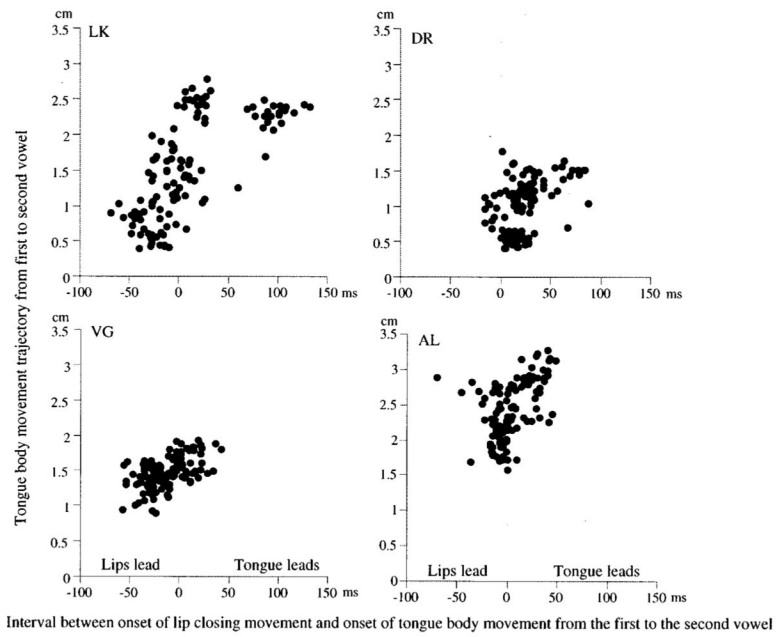
Tongue body movement trajectory from first to second vowel



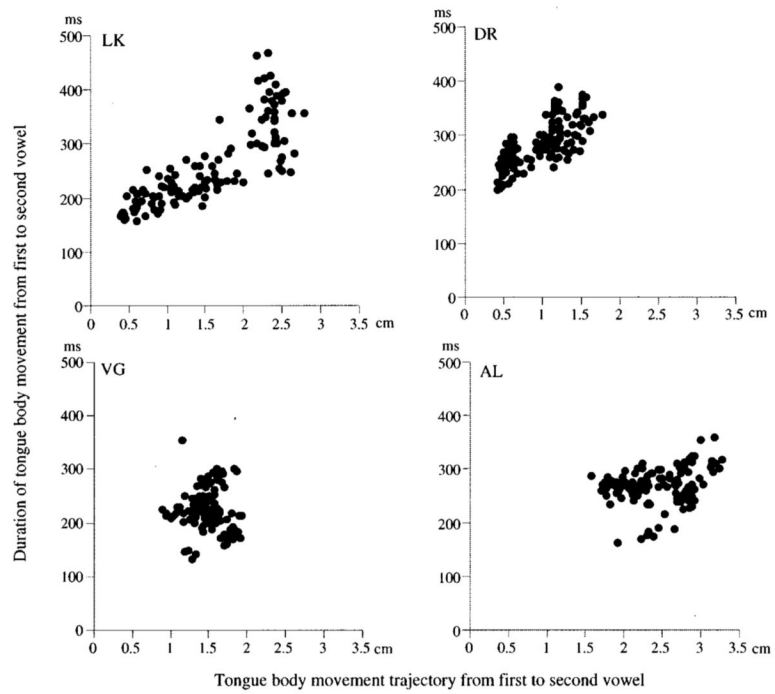
**FIG. 5.** The magnitude of the tongue movement trajectory from the first to the second vowel. Means and standard deviations are shown.



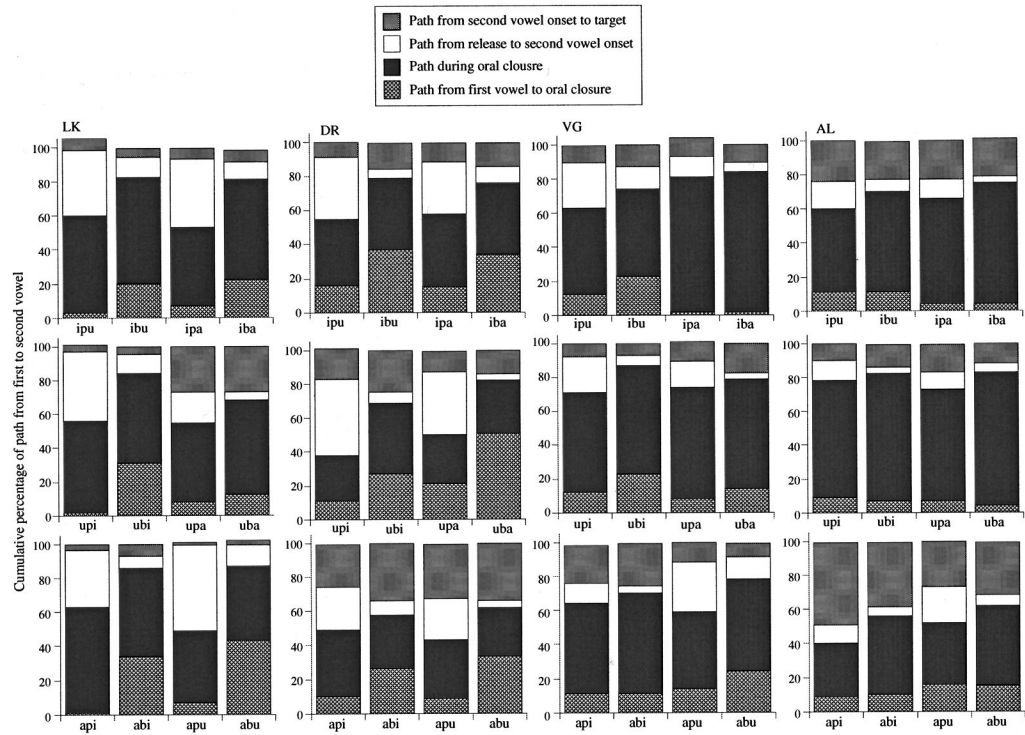
**FIG. 6.** Plot of the average tongue body positions for the three different vowels when they occur as the first (filled symbols) or second (unfilled symbols) vowel in the VCV sequence.



**FIG. 7.** Plot of the interval between the onsets of the tongue body and lip closing movements versus the magnitude of the tongue movement trajectory from the first to the second vowel.



**FIG. 8.** Plot of the magnitude and duration of the tongue movement trajectory from the first to the second vowel.



**FIG. 9.** The cumulative percentage of the tongue movement trajectory during four parts of the vowel–consonant–vowel sequence, shown in their order of occurrence from bottom to top: (1) from the onset of the tongue movement to the oral closure for the stop; (2) during the oral closure; (3) between the stop release and the onset of the second vowel; and (4) during the second vowel.