Articulatory movements in VCV sequences[a)]

Thomas Gay

University of Connecticut Health Center, Farmington, Connecticut 06032
and Haskins Laboratories, Incorporated, New Haven, Connecticut 06511
(Received 13 October 1976; revised 22 February 1977)

The purpose of this experiment was to study both the timing and positional properties of articulatory movements in VCV utterances. Conventional cinefluorographic techniques were used to track the movements of the upper lip, lower lip, jaw, tongue tip, and tongue body of two speakers who read randomized lists of VCV utterances containing the vowels /i,a,u/ and the consonants /p,t,k/, in all possible combinations. Results showed that the timing of articulatory movements in a VCV sequence are constrained by the intervocalic consonant, even if the gesture for the consonant is not a contradictory one. Anticipatory movements toward the second vowel always begin during the closure period of the intervocalic consonant. The appearance of carry over coarticulation effects depends on the phonetic identity of the particular segment or degree of involvement of the articulator. Carryover effects, like anticipatory effects, did not extend beyond an immediately adjacent segment. These findings suggest that the rules governing the segmental input to a speech string might be simpler than present models suggest.

PACS numbers: 43.70.Gr, 43.70.Ve, 43.70.Bk

INTRODUCTION

The purpose of this paper is to explore a number of questions related to the properties of articulatory movements in VCV utterances. The experiment was motivated by the fact that there exist in the literature contradictory reports concerning the nature and extent of various coarticulatory phenomena. While the traditional view and the earlier papers of Öhman[1] and of Daniloff and Moll,[2] for example, hold that coarticulation is inherent in the programming of speech sequences, and that its effects can extend across various structural boundaries, other more recent studies[3–5] suggest that the rules governing coarticulation (both anticipatory and carryover) might be somewhat simpler than previously believed.

Anticipatory coarticulation effects are essentially timing effects: movements toward some parts of a feature target of a given segment begin before others. In a study of anticipatory lip rounding, Kozhevnikov and Chistovich[6] found that the onset of the rounding gesture for the vowel /u/ placed in a CCV sequence occurred at the beginning of the sequence. Daniloff and Moll,[2] in extending the observations of Kozhevnikov and Chistovich, showed that lip rounding for /u/ can begin across as many as four segments ahead of the vowel. In their experiment, anticipation of lip rounding for the vowel /u/ was studied for a number of mono- and disyllabic single and two word utterances embedded in sentence frames using lateral view cinefluorography. Onset of lip rounding usually began during the closure phase of the first consonant in the sequence, and was not affected by the position of word or syllable boundaries within the sequence. Another type of anticipatory coarticulation was shown to exist by Öhman. In a spectrographic study of coarticulation in VCV sequences, Öhman showed that the variability observed in transition movements to the consonant could be predicted by the formant frequencies of the second vowel. This led Öhman to conclude that vowel-to-vowel movement in a VCV is essentially diphthongal with the consonant simply superimposed on the basic gesture; in other words, movements toward the second vowel begin independently from and at about the same time as those toward the consonant. In other studies, Moll and Daniloff[7] showed that velopharyngeal opening for a nasal consonant can begin two vowels in advance of the consonant, and McClean[8] showed that, in a CVVN sequence, velar opening for the final nasal begins ahead of the syllable boundary unless the two vowels are separated by a marked junctural boundary. These studies, among others, suggest that articulatory encoding is a complex phenomenon whose effects can spread across several adjacent segments. Most support, either explicitly or implicitly, Henke's[9] articulatory model that proposes the operation of a mechanism that scans future segmental inputs, or features thereof, and sends commands for the immediate attainment of those feature targets that would not interfere with the attainment of immediately intervening articulations.

However, in several recent studies, both electromyographic[4,10] and acoustic[5,12] evidence were used to argue against the ubiquity of anticipatory coarticulation effects in speech. In an experiment by Gay,[4] EMG recordings were obtained from the genioglossus and orbicularis oris muscles of two subjects during the production of various VCV syllables. In those utterances where the genioglossus muscle was involved in the production of both the first and second vowels (as in /upi/ or /itu/) or where the first and second vowels were the same (as in /ipi/ or /utu/), a cessation of activity occurred for the genioglossus muscle during the time of consonant production. In other words, each vowel in the sequence (even in a symmetrical VCV) was marked by a separate muscle pulse. This finding was interpreted as reflecting a discontinuity in vowel-to-vowel movement, and thus, a contradiction of Öhman's[1] diphthongal movement hypothesis. Another finding of this experiment was the presence of a trough in the orbicularis oris envelope during the production of an alveolar or velar consonant that separated two rounded vowels. This finding was not consistent with others that showed a considerably earlier onset of the lip rounding gesture.[2,6,11] In another EMG experiment, Ushijima and

Hirose[10] showed that in a CVVN sequence, lowering of the velum in anticipation of the final nasal was restricted by the syllable boundary. While these results were obtained from Japanese, they nonetheless argue against a general model of anticipatory velar lowering.

In an experiment performed by Bell-Berti and Harris,[5] spectrographic measurements were made from 18 utterance types that consisted of the vowels /i, a, u/ in CVC combinations with the consonants /p, t, k/. The data showed that the effects of the terminal consonant on the midpoint of the stressed vowel were not nearly as large as those of the initial consonant; in other words, the carryover effect of the initial consonant on the vowel is considerably greater than the anticipatory effect of the second consonant. The same results were also obtained independently by Ohde and Sharf[12]; in a variety of CVC sequences, carryover consonant effects on vowel targets were likewise greater than anticipatory effects.

Carryover coarticulation effects are essentially mechanical effects and exist in the form of variability in target (or target feature) positions as a function of changes in phonetic context. Carryover effects have traditionally been attributed to hysteresis or inertial effects and, in general, have been studied less extensively than anticipatory effects. Although carryover effects have been shown to exist at both the EMG and articulatory levels,[13–15] the pervasiveness of these effects is somewhat in doubt. In a study of the production of 36 CVC monosyllables, MacNeilage and DeClerk[13] found that some aspect of the production of every phone was always influenced by a preceding phone and almost always influenced by a following phone. In particular, the size of the EMG signal would be different depending on the identity of the adjacent vowel or consonant. In countering the argument that a motor command representation of the phone shows less variability than an articulatory target representation, MacNeilage[16] later proposed that the observed EMG variability reflected a complex motor strategy, the underlying goal of which is a relatively invariant articulatory end. The concept of an articulatory based target system as proposed by MacNeilage was further supported, at least for vowels by the cinefluorographic data of Gay *et al.*[17] and Gay.[3] In the latter study, lateral view x-ray motion pictures were obtained from two speakers who produced the vowels /i, a, u/ in a variety of VCV contexts. The results of this experiment showed that for both subjects, the target positions for both /i/ and /u/, in both pre- and postconsonantal positions, remained quite stable (within 2–3 mm) across changes in the consonant and transconsonantal vowel. Finally, a careful examination of Öhman's[1] acoustic data shows that carryover effects of the first vowel or the intervocalic consonant on the formant frequencies of the second vowel were virtually nonexistent: formant frequencies fell within a 50–60-Hz range regardless of the identity of the preceding phones. However, in contrast to the studies cited above, carryover effects have been shown to exist at the articulatory level. Both Sussman, MacNeilage, and Hanson[14] and Gay,[15] for example, have produced data showing jaw position during consonant and vowel production to be sensitive to the degree of jaw opening of an adjacent phone. Thus, although evidence exists to support an articulatory target formulation, no present theory specifies the rules governing failure to achieve a particular target.

The divergent research results of the last ten years, whether arising from differences in interpretation or the utilization of different experimental techniques, nevertheless serve to point out that a number of important questions concerning the dynamic properties of speech gestures remain unanswered. In this experiment, both the timing and positional properties of articulatory movements were studied in VCV utterances using conventional pellet tracking and spectrographic techniques, in an attempt to provide answers to some of these questions. The format of the experiment was designed to explore questions related to two particular issues: (1) the constraints an intervening consonant might place on the movements of the articulators, especially the tongue body, from one vowel to another (is the movement from vowel essentially diphthongal or is it locked somehow to the intervocalic consonant?) and (2) the extent of carryover coarticulation effects throughout the syllable (are such effects limited to phonetically unmarked features such as jaw position or do they extend to other properties of both vowel and consonant production?).

I. METHOD

A. Subjects and speech material

Subjects were two adult males, both native speakers of American English. The speech material consisted of CVCVC strings where the initial and final consonants remained constant (/k/ and /p/, respectively) and the medial VCV sequences contained the vowels /i, a, u/ and the consonants /p, t, k/, in all possible combinations. Each of the 27 utterances was placed in the carrier phrase, "Say——again," and random ordered into a master list.

B. Data recording

Lateral view x-ray films were recorded with a 16-mm cine camera at a speed of 60 fps. The x-ray generator delivered 1 msec pulses at 120 kV to a 9-in. image intensifier tube. For purposes of tracking articulatory movements, 2.5-mm lead pellets were attached to the upper and lower lips, tongue tip, dorsum, and body (at two locations) of both subjects.[18] In addition, a reference pellet was attached at the embrasure of the upper central incisors. Jaw movements for both subjects were tracked by measuring the distance between the tip of the lower central incisors and the reference pellet. All pellets were attached at the midline using a cyanoacrylate adhesive. The locations of the pellets are shown for both subjects in Fig. 1.

Each subject was positioned in a head holder. The subjects were instructed to read the list at a comfortable speaking rate and with equal stress placed on the two syllables. A brief practice session preceded each

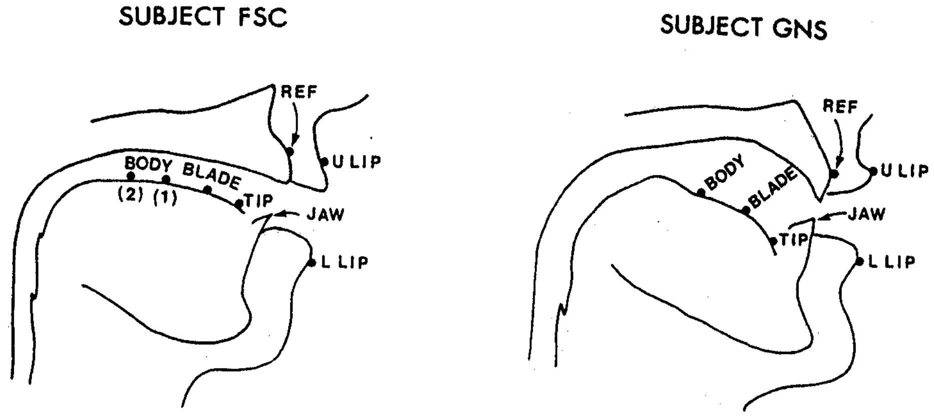

FIG. 1. Locations of pellets for tracking articulatory movements. Jaw movements were measured at tip of lower central incisors.

run. During the x-ray run, the corresponding acoustic signal was also recorded on magnetic tape.

C. Data analysis

A semiautomated system for analyzing the x-ray data was developed for this purpose. It consists essentially of a 16-mm film analyzer (Perceptoscope, Mark III) and digitizing tablet (Summagraphics) that is interfaced to a small laboratory computer (D.E.C., PDP/8E). The film image is projected, frame-by-frame, via an overhead mirror system onto the surface of the digitizing tablet. The position coordinates of each pellet (or other anatomical landmark) are stored in the computer when a hand-held pen is depressed over the pellet location. Sections of the tablet outside the image area are used for control operations, e.g., storing a special skip code or indicating end of utterance. The computer measures the X and Y coordinate positions of each pellet relative to the position of the reference pellet and stores the accumulated data, frame-by-frame-by-utterance, on disk. A second program is used to display the X and Y components separately as a movement track on a large display scope. The resolution of the digitizing tablet is 0.25 mm. By projecting the film twice real size, measurement error is easily reduced to within ± 1 mm. This was the usual maximum real size error obtained from repetitive measurements of selected samples.

One particular problem inherent in x-ray pellet tracking techniques are the obstacles dental fillings present in marking pellet locations. Because of the density of amalgams, the pellets become lost when they enter behind such fillings. Dental restorations interfered with the tracking of the first tongue body pellet of subject FSC and the tongue body and tongue tip pellets of subject GNS, both to varying degrees in different utterances.

Wide-band spectrograms, using a Haskins Laboratories digital spectrograph routine, were made for all utterances. A particular advantage of this routine is a software thresholding feature that can be used to reduce the background noise produced by the x-ray generator. This permitted spectrographic measurements to be made for almost all of the vowel nuclei, although the less intense parts of the signal associated with formant transitions were lost in the noise.

The acoustic recordings of both subjects were analyzed for the purpose of determining whether stress differences appeared for the first and second vowels. Perceived destressing occurred consistently for /a/ in preconsonantal position for subject GNS. Destressing of preconsonantal /a/ was also evident in the spectrographic measures. First and second formant frequencies for /a/, pooled across consonants and vowels, were 640 and 1340 Hz for the initial position, and 810 and 1210 Hz, for the final position. Instances of first vowel destressing for /a/ also occurred for subject FSC, but not consistently. These were the only stress effects that appeared for either subject.

II. RESULTS AND DISCUSSION

A. The timing of articulatory movements

One of the basic questions addressed in this experiment concerns the coordination of articulatory movements throughout a VCV utterance, that is, the relative timing of the movements of the tongue body in relation to those of the lips, jaw, and tongue tip, especially during the production of the intervocalic consonant. The three different consonants appearing in the various utterances were selected on the basis of the varying degrees of involvement of the tongue during their production: complete independence as a primary articulator for /p/, only tongue tip involvement for /t/, and complete involvement of the tongue body as a primary articulator for /k/. As will be shown, however, tongue body movements are either involved in or constrained by each of the three different intervocalic consonants.

Measurements of the relative onsets of articulatory movements in the various VCV sequences are summarized in Fig. 2. This figure shows the ranges of onset times of tongue body, jaw, and primary articulator (either the lower lip, tongue tip, or tongue body for /p, t, k/, respectively) movements, from the first vowel to the intervocalic consonant and from the intervocalic consonant to the second vowel. Onset times are relative to the time of closure for the consonant and are plotted separately for the three consonants. These data provide an overall picture of the relative timing of articulatory movements through the VCV sequence.

For both subjects, the timing of articulatory movements from the first vowel to the consonant were far

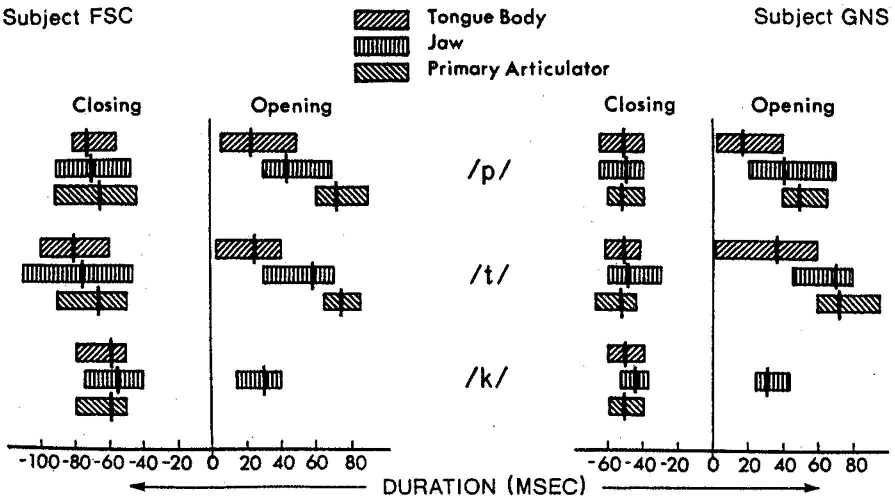

FIG. 2. Ranges of relative onset times of articulatory movement associated with consonant closing and vowel opening. The vertical lines indicate mean values.

more constrained than articulatory movements from the consonant to the second vowel. For closing movements, the onset times of tongue body, jaw, and primary articulator movements fell within the same overall time window. While the window, itself, is rather wide, coordination within the window is much more constrained, with the movements of the tongue body, jaw, and primary articulator beginning within 10–15 msec of each other. The observed overall variability could not be attributed to either the duration of consonant closure or the identity of the first vowel, although there was some tendency for earlier starting times to occur for /a/, probably as a function of greater articulatory displacement. It should also be noted that in a number of instances, notably those sequences where the first vowel is /u/, closing movements of the primary articulatory were not accompanied by corresponding movements of either the tongue body or jaw.[19]

In contrast to the synchronous closing movements from the first vowel to the consonant, opening from the consonant to the second vowel was characterized by a staggered pattern of movements. For both subjects, opening toward the second vowel began with the tongue body, and was followed by the jaw and primary articulator, in that order. Movements of the tongue body began anywhere from 5–50 msec for subject FSC and 5–60 msec for subject GNS after the time of consonant closure. All tongue body movements, however, were underway before the time of consonant release. The onset time of jaw opening also varied within the interval of consonant closure, but usually followed tongue movements and preceded primary articulator movements. The variability of opening onset times, like those for closing, did not correspond to any feature other than a tendency for earlier opening to occur for a following open vowel.

The dynamic properties of articulatory movements in a VCV sequence, and the rules that govern these movements, will be discussed for each consonant category using graphical illustrations produced from the frame-by-frame measurements of the x-ray films. The movements of the tongue body, lips, and jaw for a VCV sequence where the intervocalic consonant is /p/ are illustrated for both subjects in Fig. 3. This figure shows the movement track of the height dimension for the sequence /ipa/. Each track was graphed from discrete points measured every film frame, that is, at approximately 17-msec intervals. Measurements begin during the closure period of the initial /k/ and end at the time of closure for the final /p/; 0 on the abscissa corresponds to the time of consonant closure. This figure il-

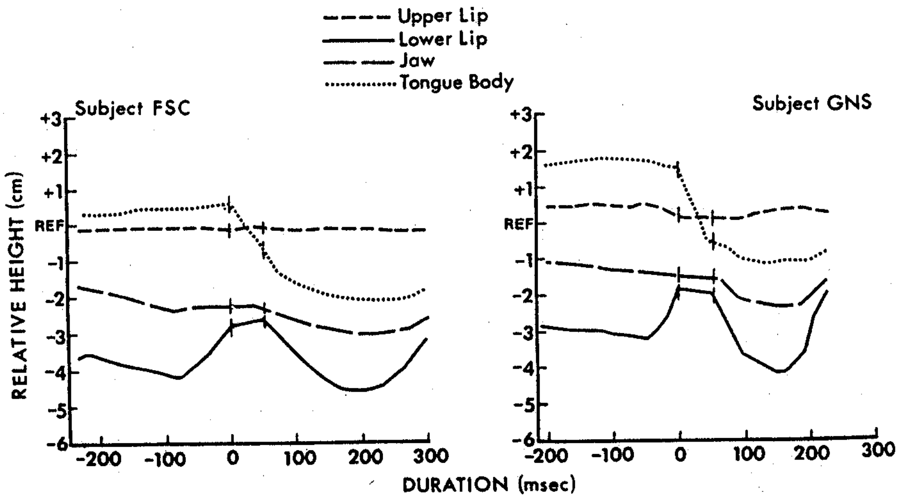

FIG. 3. Movement tracks for utterance /ipa/. 0 on the abscissa, in this and all subsequent figures, corresponds to time of consonant closure; vertical bars indicate the times of consonant closure and consonant release. The tongue body pellet for subject FSC is the second, more posterior, one.

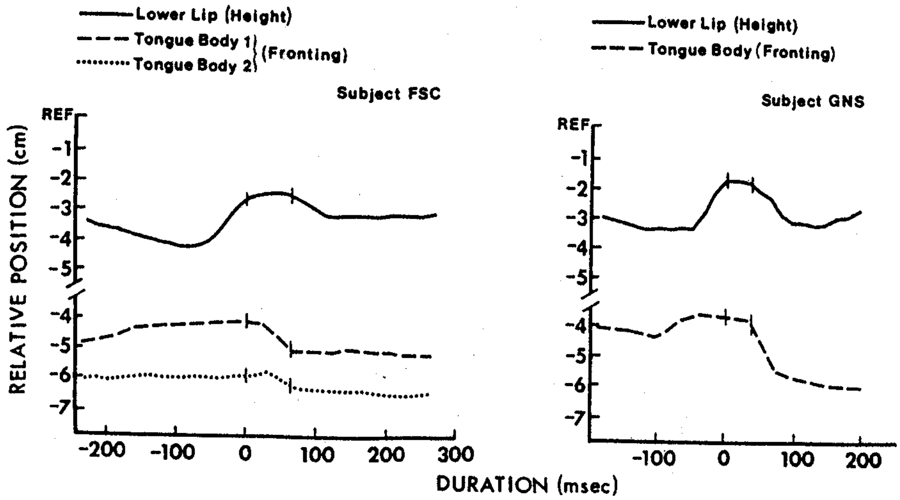

FIG. 4. Movement tracks for utterance /ipu/. Both height and fronting measurements are plotted on the same baseline.

lustrates the constraints the intervocalic consonant places on the timing of tongue body movements from vowel to vowel. The movement of the tongue body from the first vowel to the second vowel does not begin until after closure for the intervocalic consonant is completed. This, of course, was a salient feature in the production of all VCV utterances by both subjects (see Fig. 2). This figure also shows that the movements of the tongue body begin ahead of those for the jaw. The delay time is approximately 40 msec for subject FSC and 60 msec for subject GNS. This delay suggests that tongue body movements toward the vowel are probably independent from jaw movements toward the vowel. This figure also illustrates the variability of jaw movements associated with consonant production. For subject FSC, jaw closing begins at the time of lip closing while jaw opening precedes lip opening. For subject GNS, on the other hand, jaw closing does not accompany lip closing and jaw opening follows lip opening (this pattern is the only exception to the general rule). As is also evident in this figure, upper lip contributions to lip closure were negligible for both subjects. Finally, subject FSC showed a pattern of lip closure that was often characterized by continued compression throughout the closure period.

Consonant constraints on vowel-to-vowel movements are as evident in the front–back dimension as in the height dimension. Figure 4 shows tongue movement in the X dimension plotted against the same baseline as lower lip movement in the Y dimension, both as a function of time, for the sequence /ipu/. Again, it is apparent that tongue movement toward the second vowel does not begin until after consonant closure. The data for subject GNS also show what might be a tongue body gesture associated with the consonant. Such a gesture, however, did not appear regularly in the data nor did the tongue body appear to reach a specific, repeatable target position when such a gesture did appear.

The same rules for tongue body movement associated with /p/ are also evident for utterances where /t/ is the intervocalic consonant (Fig. 5). Here, as before, movements toward the second vowel do not begin until after closure for /t/. Also, this figure shows that the movements of the tongue body, tongue tip, and jaw are again, independent from each other; they all begin moving into the second vowel at different times, with the tongue body leading the jaw and tongue tip, in that order.

Perhaps the best illustration of consonantal constraints on tongue body movements is one where the first and second vowels of the utterance are the same. Figure 6 shows the movement tracks for the jaw and four tongue pellets during the production of /iti/ for sub-

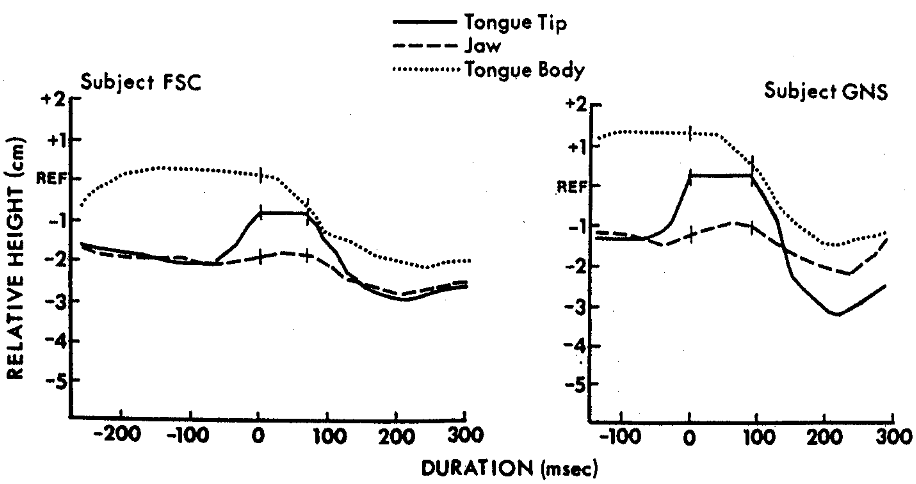

FIG. 5. Movement tracks for utterance /ita/. The tongue body pellet for subject FSC is the second, more posterior, one.

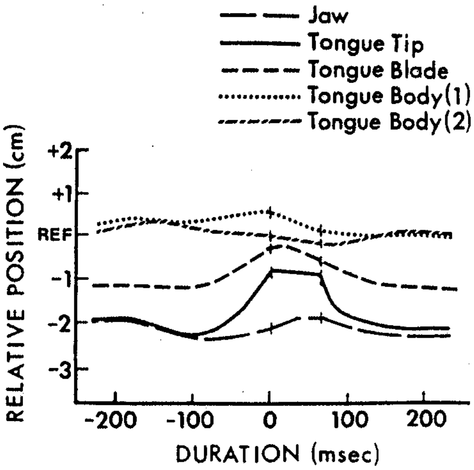

FIG. 6. Movement tracks (height) for utterance /iti/, subject FSC.

ject FSC. Instead of the tongue maintaining the /i/ target during the consonant, the tongue blade and both tongue body pellets show movement throughout the consonant gesture. The blade and anterior tongue pellet appear to shadow movements of the tip while the posterior tongue body pellet moves in the opposite direction (lower). Because the tongue body is displaced at least 5 mm from the vowel target during the time of consonant production, the movement is probably not passive (a pressure perturbation, for example). Rather, it would seem that the gesture is a facilitory one or one that reflects a strategy to modulate the degree of aspiration that might otherwise occur if the postalveolar channel were too constricted.[20] It should also be noted that the present finding agrees with the x-ray data of Kent[21] which also showed tongue body movement in a symmetrical VCV at the time of consonant production.

The most interesting tongue movements are those associated with /k/ production. Figure 7 shows both the height and fronting components of tongue body movement during the production of /aki/, /aka/, and /aku/, for subject FSC. These traces show that the tongue body is in continuous movement throughout the closure phase of the consonant. From the time of /k/ closure, the tongue body continues to move upward and forward for a following /i/ or /a/, and upward and slightly backward for a following /u/. Continuous movement of the tongue body during /k/ production has also been reported in a number of other papers. The data of both Kent[21] and Perkell[22] show elliptical patterns of movement of the tongue body for /k/ in symmetrical /VkV/ and /əkV/ sequences, respectively. A similar pattern exists in the present symmetrical /VkV/ sequences and would emerge from the /aka/ data in Fig. 7 if a composite trace were constructed from the two movement tracks. The present data are also in general agreement with those of Houde[23] who showed that the tongue body was in continuous movement during /k/ in an asymmetrical /VkV/ sequence.

Of particular interest in the present data is the finding that, irrespective of the identity of the second vowel in the sequence, closure for /k/ occurs at approximately the same location in the vocal tract. Tongue movement continues through the consonant, with release occurring at different locations in anticipation of the following vowel (see Fig. 7). While the three movement tracks are within 3 mm of each other, in both dimensions, at closure, they diverge towards release, at which point the differences are 8 mm between /i/ and /a/ in the height dimension, and 10 mm between /i/ and /u/ in the fronting dimension. Thus, consistent with the data for both /p/ and /t/, the data for /k/ show anticipatory movements to occur primarily during the closure phase of the consonant.

For VCV utterances containing either /p/, /t/, or /k/ as the intervocalic consonant, the usual sequence of articulatory events is as follows. Movements of the jaw, tongue body, and primary articulator begin at about the same time, with jaw closing continuing past the time of occlusion for the consonant. Shortly after closure for the consonant occurs, tongue body movement toward the second vowel begins. This movement is followed independently, by jaw opening, and release of the consonant. Articulatory movements for the postvocalic vowel always begin between the time of consonant closure and consonant release.

The data of this experiment, in showing consonant constraints on vowel movement in a VCV utterance, are not consistent with Öhman's[1] hypothesis that vowel-to-vowel movement in a VCV sequence is essentially diphthongal. Öhman's hypothesis is based on the assumption that tongue body movements toward the second vowel begin at about the time of onset of closing for the consonant. However, the present data show that movement toward the second vowel begins much later, some 5–60 msec after closure for the consonant has already been completed. This pattern of movement even occurs for /VpV/ sequences, where the tongue body is not actively involved in the production of the intervocalic consonant. These data suggest that either the tongue body, itself, attains a target during consonant production, or, more likely, that the release of the consonant and the movement toward the vowel are linked in a basic gesture.

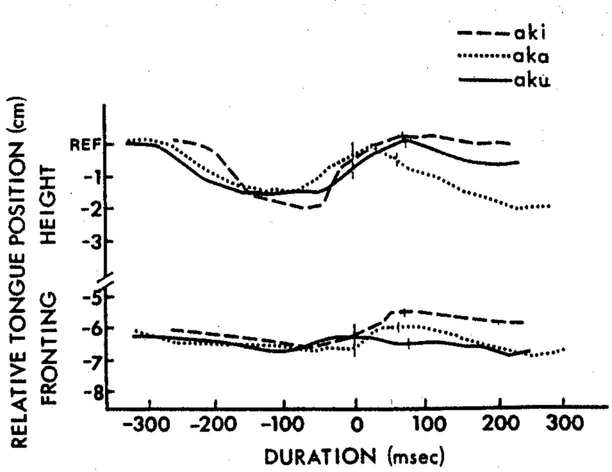

FIG. 7. Movement tracks (height and fronting) for the second tongue body pellet of subject FSC, for the utterances /aki/, /aka/, /aku/.

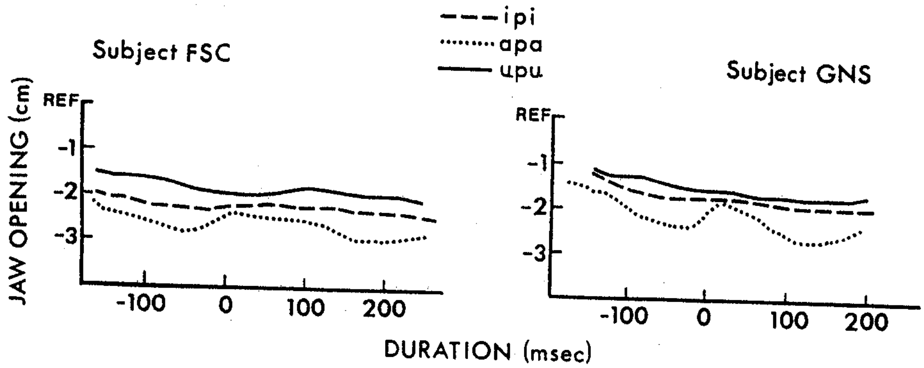

FIG. 8. Movement tracks of jaw opening for /ipi/, /apa/, /upu/, both subjects.

In addition to questions concerning anticipatory movements of the tongue body, it was expected that the data of this experiment could be used to track the onset of lip rounding for a rounded vowel preceded by a variety of different phones. Lateral view x rays can provide an indication of lip rounding in the form of degree of lip protrusion. Unfortunately, however, this measure was not a very sensitive one for the two speakers used in this experiment. The difference in protrusion between the spread vowel /i/ and the rounded vowel /u/ averaged only 5 mm for both speakers. It might be noted though, that in no case did evidence of a protruding gesture appear for the rounded second vowel in any of the VCV utterances until after closing for the intervocalic consonant was completed.

To summarize the data thus far: the relative timing of articulatory movements in a VCV sequence is affected by the intervocalic consonant, even if the gesture for the consonant is not a contradictory one. The intervocalic consonant affects both tongue body and jaw movements toward the second vowel. Anticipatory movements toward the second vowel always begin during the closure period of the intervocalic consonant, suggesting that the CV component of the VCV sequence might be organized as a basic unit.

B. The attainment of articulatory targets

Carryover coarticulation effects were studied in relation to both the influence the first vowel exerts on the position of the intervocalic consonant, and the influence the intervocalic consonant exerts on the attainment of the target for the second vowel.

In contrast to timing measurements, useful positional measurements for /p/ could not be obtained. The important positional information for /p/ appears primarily in the coronal plane; lateral view x rays simply do not reveal this information. However, the present data do show a rather strong vowel effect on jaw position during /p/. Figure 8 illustrates this effect for both subjects. These plots, which agree with the data of Sussman, MacNeilage, and Hanson[14] and of Gay[15] show that the position of the jaw during the production of /p/ is sensitive to the openness of the adjacent vowel: greater jaw opening for the consonant occurred with a more open adjacent vowel. This figure also shows what is presumed to be a stress effect in the data of subject GNS. Jaw opening (and consequently tongue height) for /a/ is reduced in the preconsonantal position.

Carryover effects of the first vowel on the positional properties of /t/ did not appear in either the tongue tip or jaw measurements. Figure 9 illustrates the insensitivity of tongue tip position for /t/ to different preceding vowels, in both the height and fronting dimensions. It is apparent that neither the retrusiveness of /u/ nor the openness of /a/ had any measurable effect on the /t/ target, in either dimensions. The only differences in the three traces appear in the timing of the closing movements. Since the onset of closing is earliest for /a/ and latest for /u/, the differences are presumed to be displacement related. Finally, jaw movements for /t/, unlike those for /p/, were not affected by the openness of the preceding or following vowel.

The most interesting and extensive carryover effects of the first vowel on consonant production appeared in the movement track of the tongue body during /k/ production. This is illustrated in Fig. 10 for the VCV sequence where /i/ is the common second vowel. Here the predicted effect of different first vowels is evident. At the time of closure for /k/, the tongue body is higher and more fronted for /i/, and progressively lower and more retruded for /u/ and /a/. The magnitude of these effects is on the order of 7 mm between /i/ and /a/ in the height dimension and 5 mm between /i/ and /a/ in the fronting dimension. The most interesting feature of this graph, however, is that the carryover effects of the

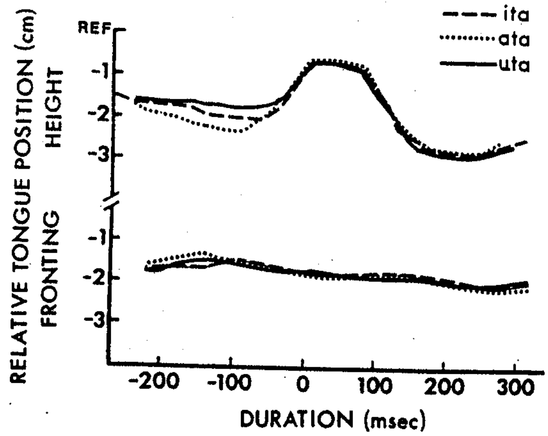

FIG. 9. Movement tracks for tongue tip, subject FSC.

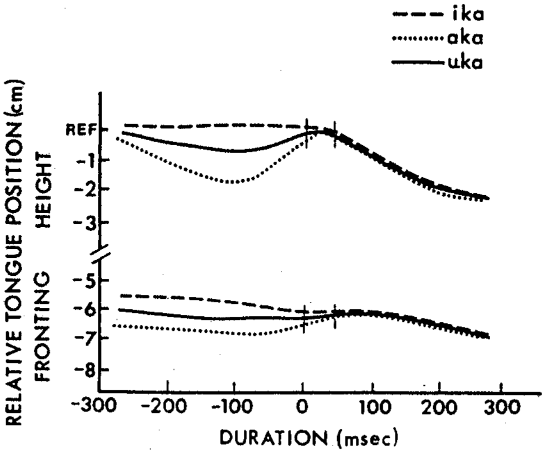

FIG. 10. Movement tracks of second tongue body pellet for /ika/, /aka/, /uka/, subject FSC.

first vowel do not extend far into consonantal closure. On the contrary, the three curves converge before consonant release at about the time movement begins toward the second vowel. The relative invariance of the movement from consonant release toward the second vowel further strengthens the suggestion that the CV transition is produced as an integral unit.

Carryover effects of the first vowel on the production of the intervocalic consonant were variable: they could not be adequately measured for /p/, they did not appear for /t/, but did appear, in a predictable way, for /k/. The jaw effect evident for /p/ is apparently due to the secondary importance of jaw closure in bringing about lip closure for /p/. Although closure for /p/ can have both lower lip and jaw components, the jaw component is probably facilitory and, as such, sensitive to phonetic environment. Likewise, the difference in effects for /t/ and /k/ is presumably related to the differences in degree of involvement of the tongue body during the production of the two consonants.

Carryover effects of the intervocalic consonant on the following vowel appeared only for the open vowel /a/, and were reflected in differences in jaw and, consequently, tongue body height. These effects, which are consistent with those reported by Gay,[3] are illustrated in Fig. 11. This figure shows the differences in tongue body and jaw height for the vowel /a/ when the intervocalic consonant varies from /p/ to /t/. Opening for the vowel is greater when the intervocalic consonant is /p/ as opposed to /t/. The difference in tongue body height for the first vowel is probably due to differences in stress between the two utterances. However, this was not apparent when listening to the tapes. This figure also shows what appears to be tongue body involvement during the production of /t/. The movement track for the tongue body shows greater elevation than that for the jaw during the time of consonant production. This means that the tongue body position during consonant production is not simply being carried passively by the jaw, but rather has underlying it, an active muscle component as well. Although variability in tongue body and jaw opening appeared in the articulatory data for both subjects, similar variability was not reflected in the corresponding acoustic measures. Apparently, the differences in jaw position as measured anteriorly at the incisors either do not correspond to the size of the pharyngeal constriction for /a/, or are much less when the arc of rotation is measured closer to the hinge axis of the jaw.

Carryover effects of a preceding consonant on the production of the vowels /i/ and /u/ were small. These effects are summarized in Fig. 12 and Table I. The figure shows the relative positions of the upper lip, lower lip, jaw, and tongue body at the time the tongue body reached its target (point of maximum displacement) for each of nine utterances containing the vowel /i/ in final position. Table I shows the corresponding values of the first and second formant frequencies at that point in time.

As is evident in the figure, variability of tongue body target positions is minimal (2.5 mm for subject FSC and 3 mm for subject GNS). Lower lip and jaw positions, on the other hand, vary within a large range, approximately 5 mm for subject FSC and 10 mm for subject GNS. Interestingly, lower lip and jaw targets seem to vary independently from tongue body positions but covary for both subjects. This finding contradicts that of Hughes and Abbs[24] who showed that mouth opening for /i/ remained relatively constant because of tradeoffs between lower lip and jaw displacements. This type of equivalence was not evident in the present data for either /i/ or /u/. Differences between the two sets of data might be attributable to differences in either or both the speech material and instrumental methods used in the two experiments.

The acoustic measurements of target formant frequencies showed some variability among the nine utterances (Table I). First formant frequencies were within a range of 40 Hz for both subjects while second formant frequencies fell within a range of 230 Hz for subject FSC and 120 Hz for subject GNS. The measured acoustic variability did not appear to correspond to any observed articulatory variability. For example, utterances 2 and 7 for subject FSC were characterized by

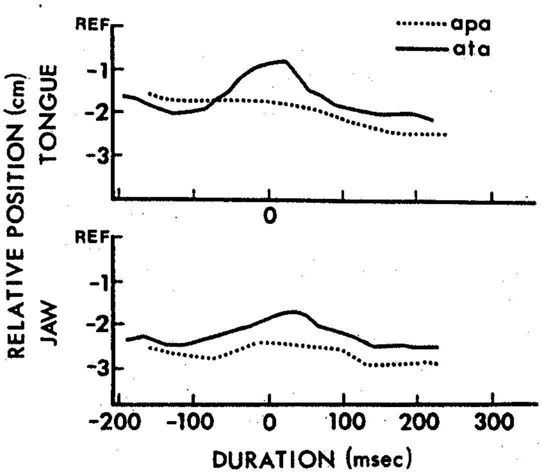

FIG. 11. Movement tracks of tongue body (pellet 2) and jaw height for /apa/ and /ata/, subject FSC.

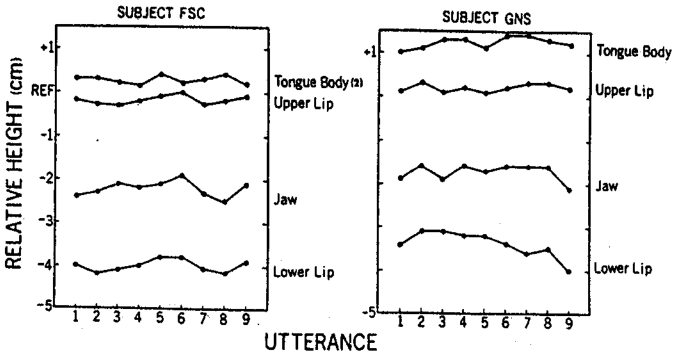

FIG. 12. Coordinate positions of upper lip, lower lip, jaw, and tongue body for the target positions of the vowel /i/, both subjects. Each utterance number corresponds to the utterance in Table I.

similar articulatory target points but quite different formant frequencies. Conversely, utterances 3 and 4, and 1 and 9, were characterized by virtually the same formant frequencies, but different articulatory target points. Either the variability observed fell, for the most part, within the range of measurement error, or more likely, a four-point parametrization tracking procedure of the type used in this experiment is simply inadequate for the purpose of relating differences in articulatory target points to the acoustic output. It might also be noted that acoustic variability for both /u/ and /a/ were, in terms of percentage, within the same range as variability for /i/.

Carryover effects, then, when they do appear, are unlike anticipatory effects in that they depend on the phonetic identity of the particular segment. Like anticipatory effects, however, carryover effects seem to spread no farther than the neighboring phone. These findings support an articulatory based formulation of speech production.[16] For the most part, an articulatory target corresponded as a relatively invariant representation of a phoneme. Articulatory variability, when it did occur, did so only under special circumstances. First, carryover effects for a consonant are reflected mainly in variability of jaw position, and only when the jaw is not primarily involved in the production of the phone, as in /p/. However, when the jaw is more tightly involved in the production of a phone (/t/, for example), degree of jaw opening was not sensitive to that of the adjacent phone. The only other strong carryover effect appeared in tongue body movements for intervocalic /k/. Here, unlike variability in jaw opening, carryover effects on tongue body movements do not seem to be either random in appearance or inertial in origin. Unlike /VpV/ and /VtV/ sequences where the tongue body is usually in a waiting position before it moves toward the second vowel during consonant closure, in a /VkV/ sequence, the tongue body is involved as a primary articulator in the production of the consonant. The movements of the tongue body through /k/ (Fig. 10) seem to be directed, in a straight-line fashion, to a common target position for release of the consonant. The data for /k/ provide a fairly convincing illustration of the limited spreading effects of coarticulation in a VCV sequence. Because of continuous tongue body involvement in the production of CVC syllables containing /k/ as the intervocalic consonant, the elements of these syllables, especially /k/ itself, should be the most sensitive to the spreading of coarticulation effects, in both directions. Yet, the assimilation of carryover effects and the onset of anticipatory movements both occur within the closure period of the consonant, with movements from the same vowel into /k/ (see Fig. 7), or movements toward the same vowel from /k/ (see Fig. 10), not being affected by the articulatory event on the other side of the consonant.

Stability of tongue body targets for vowels (at least /i/ and /u/) was also the rule rather than the exception. The only substantial articulatory variability occurred in jaw displacement, with /a/ showing the greatest effects and /u/ the least. As was mentioned before, however, variability in jaw displacement for /a/, as measured anteriorly at the incisors, might be either exaggerated or irrelevant in relation to variability that might exist in the pharyngeal constriction for /a/. Likewise, the variability of maximum jaw displacement for both /i/ and /u/ seems unrelated to the variability observed in the position of the tongue body for those vowels. Thus, the two features, tongue body height and jaw displacement, might be independent ones, with jaw opening being a facilitory gesture and an unmarked phonetic feature. This formulation suggests a reevaluation of models of vowel articulation that specify jaw position as a primary determiner of tongue height.[25]

III. SUMMARY AND CONCLUSIONS

The major findings produced by this experiment are as follows. First, anticipatory movements toward the second vowel in a vowel–stop consonant–vowel sequence begin during the closure period of the intervocalic consonant. This restricted coarticulatory field includes both tongue body and jaw movements associated with the second vowel. Furthermore, the size of this field is not affected by the identity of the intervocalic consonant. Second, like anticipatory effects, carryover effects did not extend beyond an immediately neighboring segment. Unlike anticipatory effects, however, the appearance of carryover coarticulation effects depended on the phonetic identity of the particular segment on which these effects might act.

TABLE I. First and second formant frequency values (Hz) for the vowel /i/ in nine different VCV utterances. Each utterance number corresponds to that of Fig. 12.

	Subject FSC		Subject GNS	
Utterance	F_1	F_2	F_1	F_2
1. ipi	340	2200	310	2230
2. api	360	2030	320	2250
3. upi	360	2220	300	2160
4. iti	360	2220	330	2200
5. ati	320	2120	340	2210
6. uti	350	1990	320	2120
7. iki	320	2210	320	2270
8. aki	360	2160	320	2160
9. uki	350	2190	320	2250

The implication of these findings is that the rules governing the segmental input to a VCV string might not be as complex as present models suggest. The finding that anticipatory movements begin and primary carryover effects end at about the same time during the closure period of the consonant, suggests that the release of the consonant and movement toward the vowel are organized and produced as an integral articulatory event.

This formulation, which suggests a limited coarticulatory field is not consistent with the operation of a complex scan-ahead mechanism. This does not necessarily mean, however, that a scan-ahead mechanism does not operate on larger units or at another stage of the speech production process. For example, Lindblom and Rapp,[26] Nooteboom and Cohen,[27] and Fromkin,[28] have suggested the existence of an anticipatory mechanism in the temporal formulation of speech sequences. Likewise, the complex reordering of commands accompanying changes in speaking rate[17] also suggests that the temporal features of a downstream segment might be known in advance.

Thus, while it has been traditionally considered that the serial ordering of segments is governed by complex rules whose effects can spread across several adjacent segments, and the temporal control of speech is governed by a simple adjustment of timing of commands to the articulators[29] it may well be that the reverse is true: The segmental input to the speech string is governed primarily by simple rules which act upon syllable sized units, while the temporal formulation of the string requires complex articulatory adjustments based on advance information obtained from a higher level scan-ahead mechanism.

Like most studies of speech organization, especially those using high-speed cinefluorographic techniques, the results of this experiment are based on data obtained from a relatively small subject population and are applicable to the production of only a few phonetic elements, themselves constrained by the artificial format in which they were placed. Thus, the findings of this experiment are obviously far from conclusive, and go only part way toward answering those questions posed at the outset. The present findings can serve, however, as a basis for examining or reexamining a number of questions concerning the organization of segmental gestures. For example, it was shown that a four-point parametrization procedure for relating articulatory targets to acoustic targets is inadequate. In order to resolve the differences between the acoustic data of Öhman[1] and the articulatory data of the present study, formant tracking must be matched to a far more comprehensive multipoint parameterization of the vocal tract. The present results also suggest, without providing convincing evidence, that the onset of anticipatory lip rounding might be conditioned differently in CCCV and VCV sequences; also, they raise further questions about the use of trade-offs between tongue and jaw movements in achieving articulatory targets, and the importance of jaw position in determining tongue height in vowel articulation.

ACKNOWLEDGMENTS

The author wishes to thank Dr. J. Daniel Subtelny, Department of Orthodontics, Eastman Dental Center, Rochester, New York, for use of the cinefluorographic facilities at Eastman, and Ms. Kathleen Kirchmeier for her assistance in the analysis of the data.

[a]This research was supported by grants from the National Institute of Neurological and Communicative Disorders and Stroke (NS-10424), The National Science Foundation (GSOC-7403725), and the National Institute of Dental Research (DE-01774).

[1]S. E. G. Öhman, "Coarticulation in VCV utterances: Spectrographic measurements," J. Acoust. Soc. Am. 39, 151–168 (1966).

[2]R. G. Daniloff and K. L. Moll, "Coarticulation of lip-rounding," J. Speech Hear. Res. 11, 707–721 (1968).

[3]T. J. Gay, "A cinefluorographic study of vowel production," J. Phonetics 2, 255–266 (1974).

[4]T. J. Gay, "Some electromyographic measures of coarticulation in VCV utterances," Haskins Labs. Status Rep. Speech Res. SR-44, 137–145 (1974).

[5]F. Bell-Berti and K. S. Harris, "Some acoustic measures of anticipatory and carryover coarticulation," Haskins Labs. Status Rep. Speech Res. SR-42/43, 297–304 (1975).

[6]V. A. Kozhevnikov and L. A. Chistovich, *Rech', Artikulyatsiya i Vospriyatiye*, Vol. 30, p. 543; English Transl.: *Speech: Articulation and Perception*, Joint Publications Research Service (Washington, DC, 1966).

[7]K. L. Moll and R. G. Daniloff, "Investigation of the timing of velar movements during speech," J. Acoust. Soc. Am. 50, 678–684 (1971).

[8]M. McClean, "Forward coarticulation of velar movement at marked junctural boundaries," J. Speech Hear. Res. 16, 286–296 (1973).

[9]W. Henke, "Dynamic Articulatory Model of Speech Production Using Computer Simulation," Ph.D. thesis (MIT, 1966) (unpublished).

[10]T. Ushijima and H. Hirose, "Velar movement and its motor command," Haskins Labs. Status Rep. Speech Res. SR-41, 207–216 (1975).

[11]A-P. Benguerel and H. A. Cowan, "Coarticulation of upper lip protrusion in French," Phonetica 30, 41–55 (1974).

[12]R. N. Ohde and D. J. Sharf, "Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets," J. Acoust. Soc. Am. 58, 923–924 (1974).

[13]P. F. MacNeilage and J. DeClerk, "On the motor control of coarticulation in CVC monosyllables," J. Acoust. Soc. Am. 45, 1217–1233 (1969).

[14]H. Sussman, P. F. MacNeilage, and R. Hanson, "Labial and mandibular dynamics during the production of bilabial consonants," J. Speech Hear. Res. 16, 397–420 (1973).

[15]T. J. Gay, "Jaw movements during speech: A cinefluorographic investigation," Haskins Labs. Status Rep. Speech Res. SR-39/40, 219–230 (1974).

[16]P. F. MacNeilage, "Motor control of serial ordering of speech," Psychol. Rev. 77, 182–195 (1970).

[17]T. Gay, T. Ushijima, H. Hirose, and F. S. Cooper, "Effect of speaking rate on labial consonant–vowel articulation," J. Phonetics 2, 46–63 (1974).

[18]The second, more posterior, tongue body pellet for subject GNS fell off during the experiment run.

[19]When the intervocalic consonant was either /p/ or /t/, tongue body participation in the consonant gesture depended on the identity of the first vowel; tongue body movements always accompanied primary articulator movements when the first vowel was /a/, sometimes showed movement when the first vowel was /i/, and never showed movement when the first vowel was /u/. For /k/, of course, the tongue body always

showed movement into the consonant. In those cases where tongue body movements did not appear for the consonant, the tongue body simply maintained the target position of the first vowel.

[20]K. N. Stevens (private communication).
[21]R. D. Kent, "A cinefluorographic-spectrographic investigation of the component gestures in lingual articulation," Ph. D. thesis (University of Iowa, 1970) (unpublished).
[22]J. S. Perkell, *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study* (MIT, Cambridge, MA, 1969).
[23]R. A. Houde, "A study of tongue body motion during selected speech sounds," Ph. D. thesis (University of Michigan, 1967) (unpublished).
[24]O. M. Hughes and J. H. Abbs, "Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence," Phonetica 33, 199–201 (1976).
[25]B. E. F. Lindblom and J. Sundberg, "Acoustical consequences of lip, tongue, jaw and larynx movement," J. Acoust. Soc. Am. 50, 1166–1179 (1971).
[26]B. Lindblom and K. Rapp, "Some temporal regularities of spoken Swedish," PILUS (Stockholm University, 1973).
[27]S. G. Nooteboom and A. Cohen, "Anticipation in speech and its implications for perception," in *Proceedings of the Symposium on Dynamic Aspects of Speech Perception* (IPR, Eindhoven, 1975).
[28]V. A. Fromkin, "The non-anomalous nature of anomalous utterances," Language 47, 27–52 (1971).
[29]B. E. F. Lindbloom, "Spectrographic study of vowel reduction," J. Acoust. Soc. Am. 35, 1773–1781 (1963).