Reprinted from

DYNAMIC ASPECTS OF SPEECH PRODUCTION

Current Results, Emerging Problems, and New Instrumentation

edited by
MASAYUKI SAWASHIMA and FRANKLIN S.COOPER

Gay

UNIVERSITY OF TOKYO PRESS, 1977

CINEFLUOROGRAPHIC AND ELECTROMYOGRAPHIC STUDIES OF ARTICULATORY ORGANIZATION

Thomas Gay

Department of Oral Biology, University of Connecticut Health Center; and Haskins Laboratories

This paper summarizes the results of several experiments that used the techniques of cinefluorography and electromyography to study the organization of speech gestures. As such, it does not represent a comprehensive review of current speech production theory, but rather is directed towards a discussion of several specific issues that are best studied by these techniques: the dynamics of articulatory movements and the motor command structure that underlies those movements.

Although speech is usually described in terms of a string of invariant segmental units (phonemes), the act of speaking imposes on this string a complex encoding. This is a consequence of the series of events that comprises the speech production chain: the conversions of motor command-to-muscle contraction, muscle contraction-to-vocal tract shape, and vocal tract shape-to-acoustic signal. The result of this encoding is observed as variation both in the production of a given phone and in its acoustic representation. This paper will be concerned with allophonic variation as it appears at the articulatory level, specifically, variations which arise from changes in phonetic context, and variations which arise from changes in the suprasegmental structure of the string, particularly changes in speech rate.

THE ORGANIZATION OF SEGMENTAL GESTURES

Coarticulation is usually defined as allophonic variation of a given phone due to changes in its phonetic environment. The production of a phone can be conditioned by a phone that either precedes it (left-to-right or carryover effects) or follows it (right-to-left or anticipatory effects).

Anticipatory coarticulation effects are essentially timing effects: movements toward some parts of a feature target of a given segment begin before others. Kozhevnikov and Chistovich (1965) studied the anticipation of lip rounding which occurs when a rounded vowel follows a consonant and suggested that the forward extent of this anticipatory gesture was limited by the position of the syllable boundary. Daniloff and Moll(1968), however, showed that lip rounding can begin ahead of the syllable boundary, and across as many as four consonants preceding 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 imbedded in sentence frames, using lateral view x-ray motion pictures. Onset of lip rounding usually began with the first consonant of the utterance. Another type of anticipatory coarticulation was shown to exist by Öhman (1966). In a spectrographic study of coarticulation in VCV sequences, Öhman concluded that the variability observed in transitional movements to the consonant could be predicted by the formant frequencies of the second vowel. This led Öhman to conclude further 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 those toward the consonant. 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 (1966) 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 two recent studies, both cinefluorographic (Gay, 1976) and electromyographic (Gay, 1974a), evidence were used to argue against the ubiquity of anticipatory coarticulation in speech.

In the cinefluorographic experiment, conventional high speed (60 fps) lateral view x-ray films were recorded from two subjects who produced various VCV syllables that contained the vowels /i,a,u/ and the consonants /p,t,k/, in all possible combinations. Articulatory movements were tracked by recording the positions, frame-by-frame, of 2.5 mm diameter lead pellets that were attached to the upper and lower lips, jaw, and several locations along the surface of the tongue relative to a reference pellet attached at the embrasure of the upper central incisors. These data will be used to explore the question of whether, in a VCV, an intervening consonant constrains movements of the articulators, in particular the tongue body and lips, from one vowel to another: is the movement from one vowel to another essentially diphthongal or is it somehow locked to the consonant (Öhman's model); and does the lip rounding gesture for the postvocalic rounded vowel begin ahead of the intervocalic consonant (Henke's model)?

The dynamic properties of articulatory movements in a VCV sequence are illustrated in Fig.1 for an utterance where the intervocalic consonant is /p/. This figure shows the movement tracks of the tongue body, lips, and jaw in 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 illustrates the general finding that the intervocalic consonant affects the timing of the movements of the tongue body 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 was found to be a salient feature in the production of all VCV utterances. The only variability in the timing effect appears in the delay time between consonant closure and tongue body movement. While the lag was usually of the order of 30 msec, it varied anywhere from 10-60 msec. This figure also shows that the movements of the tongue body, because they begin ahead of those for the jaw, are probably independent of jaw movements towards the vowel. As is also evident in this figure, upper lip contributions to lip closure were negligible. Finally, this subject showed a pattern of lip closure that was often characterized by continued compression throughout the closure period.

It was also found that consonant constraints on vowel-to-vowel movements were as evident in the front-back dimension as in the height dimension, and the same rules that apply to /p/ also apply when the intervocalic consonant is either /t/ or /k/. 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 2 shows the movement tracks for the jaw and four tongue pellets duri the production of /iti/ for Subject FSC. Instead of the ton 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 body let appear to shadow movements of the tip while the posteri tongue body pellet moves in the opposite direction (lower), probably in a facilitory gesture. However, since the tongu body gesture (when it does appear) does not reach a specifi repeatable location, it is not interpreted as being target rected.

The discontinuity of the vowel target in an utterance w the same two vowels are separated by a consonant is also e dent at the EMG level (Gay, 1974a). The average EMG activ of the genioglossus muscle for the sequences /ipi/ and /it is illustrated in Fig.3. The genioglossus muscle, which

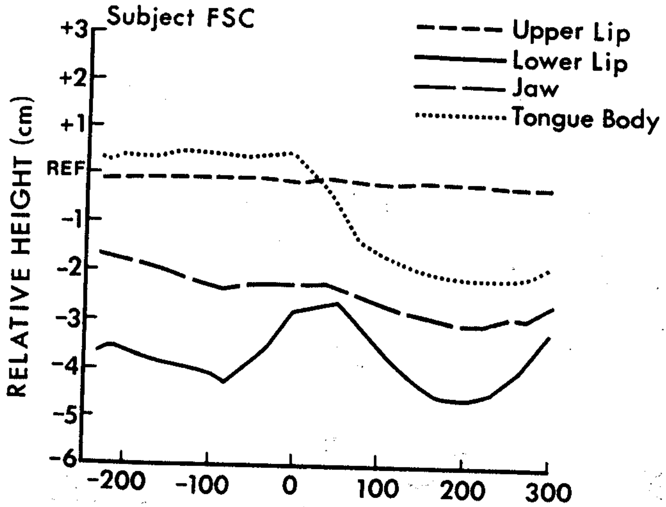

Fig.1. Movement tracks for the sequence /ipa/. Relative tongue height (ordinate) is plotted as a function of time (abscissa) in msec.

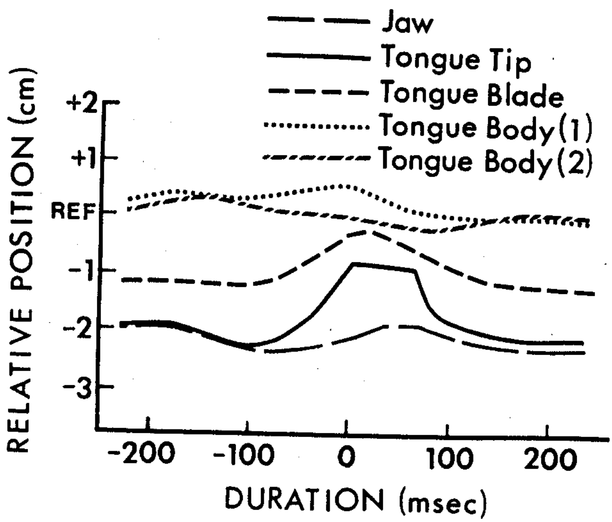

Fig.2. Movement tracks for the sequence /iti/, Subject FSC.

prises the bulk of the tongue body, is primarily responsible for the protruding and bunching associated with the vowel /i/. This figure shows three separate peaks associated with the utterance. The first peak corresponds to the initial /k/ while

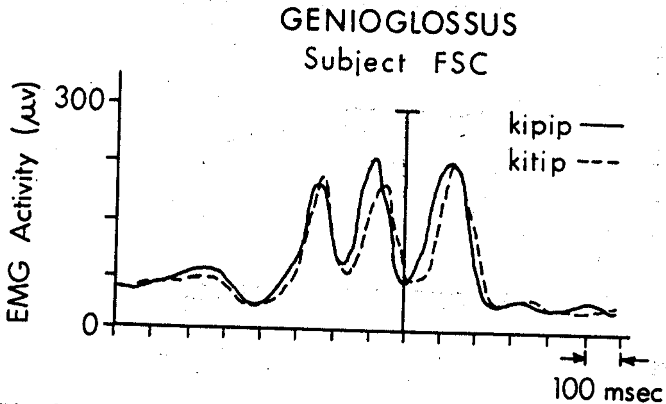

Fig.3. Average genioglossus muscle activity for the utterances /ipi/ and /iti/, Subject FSC. The lineup point is the time of voicing offset of the first vowel.

the second and third correspond to the first and second vowels. Of particular interest is the deep trough that separates the two vowel peaks. The presence of a trough, which signifies a cessation of genioglossus activity, suggests that the two vowels, although phonetically identical, are organized as two separate events. If the movement of the tongue body during the production of the consonant as observed in the x-ray data (ref. Fig.2) was the result of other articulatory influences, tongue tip or jaw movements for example, positional constancy would exist at the EMG level in the form of one broad genioglossus peak across the entire utterance. However, the existence of two distinct peaks separated by a deep trough suggests that the intervocalic consonant has more than a passive effect on tongue body movement at that time.

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 post-vocalic vowel always begin between the time of consonant closure and consonant release. Constraints of the intervocalic consonant are also evident at the EMG level in the form of a separate muscle peak for each syllable.

The data from both the cinefluorographic and electromyographic experiments, in showing consonant constraints on vowel movement in a VCV utterance, argue against Öhman's (1966) model that suggests vowel-to-vowel movement is essentially diphthongal. If Öhman's model were correct, tongue body movements toward the second vowel would begin at about the time of onset of closing for the consonant. However, movement toward the second vowel begins much later, some 10-16 msec after closure for the consonant has already been completed. This would suggest that either the tongue body itself attains a target during consonant production or, more likely, that the intervocalic consonant and the following vowel are linked in a basic gesture. The very short lag time between consonant closure and movements toward the second vowel suggest the latter possibility.

In addition to placing constraints on the movements of the tongue body from one vowel to another, an intervocalic consonant also affects the onset of lip rounding for a rounded second vowel. These constraints are evident in our data at both the articulatory and EMG levels. Lateral view x-rays can provide an indication of lip rounding in the form of degree of lip protrusion. In those cases where a rounded vowel appears in a post-consonantal position, the rounding gesture, like tongue body movements, does not begin until after closure fo the intervocalic consonant is completed. This is true even for the most sensitive case, namely, two rounded vowels sepa rated by a close consonant. Figure 4 shows the movement tracks of lower lip height, lower lip protrusion, and tongue tip height plotted against the same baseline for the syllabl /utu/, produced by Subject GNS. In this example, it appears that the rounding feature of the first vowel is not continuo through the consonant. Rather, what appears to be an additi al (although small) closing and protruding gesture is superi posed on the rounding pattern. This discontinuity of round- ing during the consonant is also evident in the EMG data (Fi 5) which show a trough in orbicularis oris muscle activity during the production of the same syllable. Both sets of da argue against the Daniloff and Moll (1968) anticipatory effe

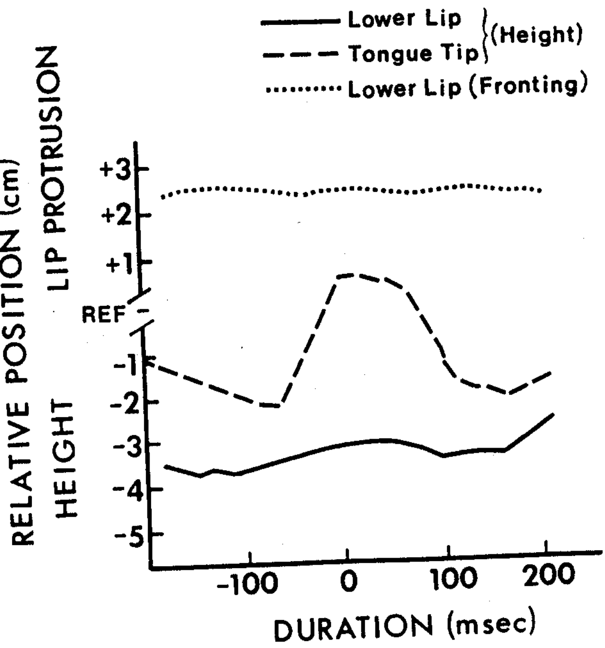

Fig.4. Movement tracks (tongue height and lip protrusio for the utterance /utu/.

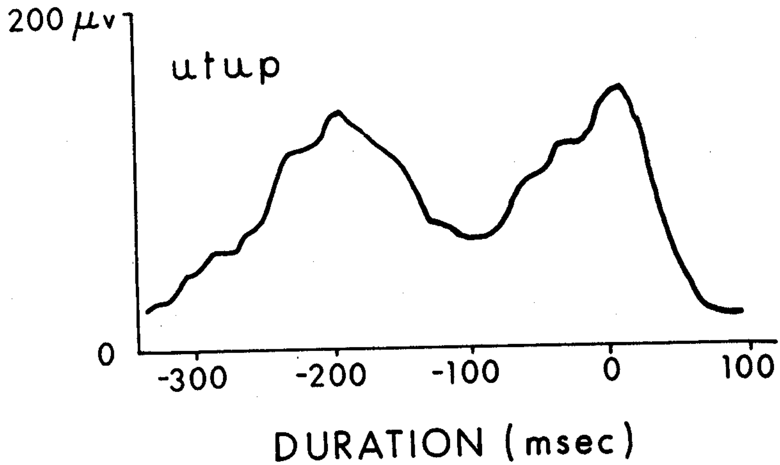

Fig.5. Average orbicularis oris muscle activity for the utterance /utu/.

An alternative interpretation of Daniloff and Moll's result is that the early onset of lip rounding corresponded to a closing or protruding gesture of one or more of the intervening consonants in the utterance (for example, the /n/, /s/, /t/, or /r/ in the word "construe") or some special property of the cluster itself. This explanation is compatible with the EMG data of Bell-Berti and Harris (1976), which show that the beginning of orbicularis oris muscle activity in utterances containing the syllables /stru/ and /stri/ occurs at the same time.

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 shows effects on tongue body movements toward and the lip rounding gesture for the second vowel at both the articulatory and EMG levels. Anticipatory movements toward the second vowel begin during the closure period of the intervocalic consonant, suggesting that the CV component of the VCV sequence might be produced as a basic unit.

Unlike anticipatory coarticulation effects which are essentially timing effects, carryover coarticulation effects are positional 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 mechanical 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, the pervasiveness of these effects is somewhat in doubt. In a study of the production of thirty-six CVC monosyllables, MacNeilage and DeClerk (1969) found that some aspect of the production of every phone was 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 (1970) 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. (1974) and Gay (1974b). 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 post-consonantal positions, remained quite stable (within 2-3 mm) across changes in the consonant and trans-consonantal vowel. Although target stability for /a/ was also the rule rather than the exception, some individual differences did appear. However, the articulatory variability, when it did appear, did not correlate with any acoustic variability.

Similar results were also reported in a more recent cinefluorographic study (Gay, 1976). Carryover effects of an intervocalic consonant on the following vowel again appeared only for the open vowel /a/, and were reflected in differences in jaw, and consequently tongue body, height. However, carryover effects of a preceding consonant on the attainment of target positions for the vowels /i/ and /u/ were minimal, with the tongue body targets of both vowels falling within a range of 2.5 mm for one subject and 3 mm for the other. This lack of variability is illustrated in Fig.6. 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

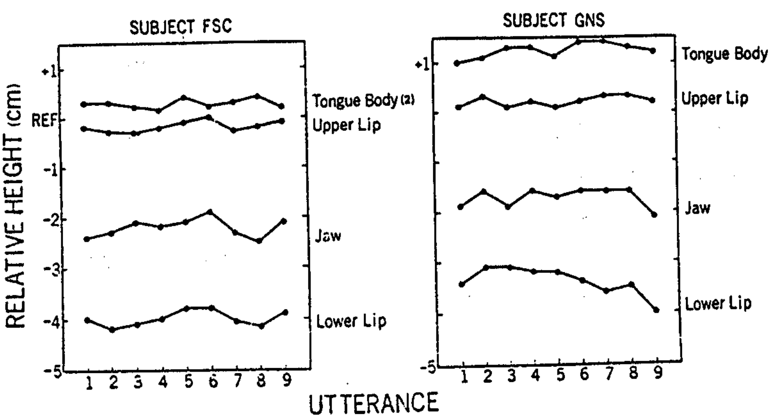

Fig.6. Coordinate positions (height) of upper lip, lower lip, jaw, and tongue body corresponding to the target positions of the vowel /i/.

containing the vowel /i/ in final position. As is evident, variability of tongue body target positions is minimal. Lower lip and jaw positions, on the other hand, vary within a larger 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 (1976), who showed that mouth opening for /i/ remained relatively constant because of trade-offs between lower lip and jaw displacements. This type of equivalence was not evident in the present data for either /i/ or /u/.

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. Stability of tongue body targets for vowels (at least /i/ and /u/) is 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. 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/ is 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, are probably independent ones, with jaw opening being a facilitory gesture and an unmarked phonetic feature.

SUPRASEGMENTAL ORGANIZATION: THE CONTROL OF SPEECH RATE

In the preceding section, variability in the production of a phone due to changes in phonetic environment was discussed. In this section, questions concerning allophonic variation that arises from a different source, the suprasegmental feature of speaking rate, will be explored.

Experiments on the effects of speaking rate and stress have been concerned primarily with the question of whether all such effects can be attributed solely to changes in the timing of commands to the articulators. The classic experiments on the effects of stress and speaking rate on vowels were conducted by Lindblom (1963, 1964). By inferring changes in articulator positions from sound spectrograms, Lindblom found a positive correlation between vowel reduction, or "undershoot," and either decreased stress or increased speaking rate. The failure of the articulators to reach the vowel "target" was attributed to the close temporal succession of motor commands, and so to insufficient time to complete the component gestures. In addition, Lindblom's speaking rate data showed that the rate of target-directed articulator movement remained constant across changes in duration. This supported the concept of undershoot as being a time-based phenomenon; also, it implied a simple model to account for the effects of stress and speaking rate, i.e., a cut-off of the commands rather than a complete reorganization of the gesture.

However, in two separate experiments, one cinefluorographic (Gay et al., 1974) and one electromyographic (Gay and Ushijima, 1974), it was found that changes in speaking rate are brought about by a complex re-programming of the input to the speech string. For example, in a combined electromyographic-cinefluorographic study of speaking rate control (Gay et al., 1974), it was found that an increase in speaking rate was accompanied by a decrease in vowel duration and, for the most part, articulatory undershoot. However, the degree of articulatory undershoot varied with both the individual subject and phonetic identity of the vowel. These differences are illustrated in Fig.7. This figure shows tongue body and jaw movement for the sequence /api/, for two different speaking rates. It is evident that the degree of undershoot is considerably greater for the open vowel /a/ than for the close vowel /i/.

More interesting than the existence of articulatory undershoot for fast speech are the underlying muscle action patterns that control those movements. The EMG data provide a fairly complete account of this control mechanism. These data show that lip muscle activity (orbicularis oris) associated with

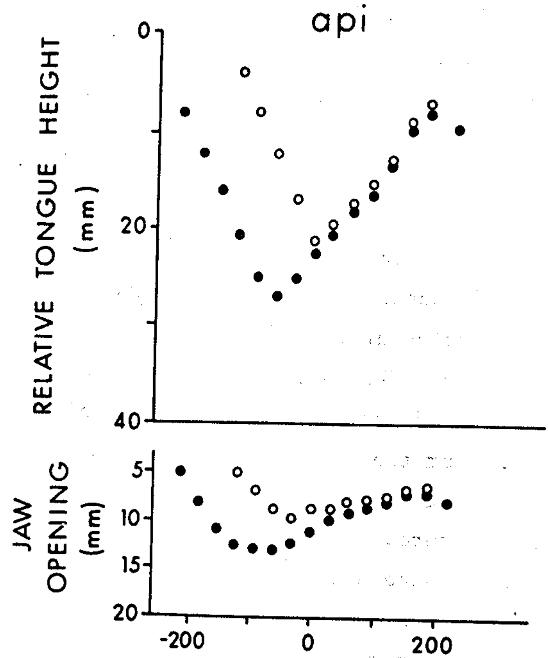

Fig.7. Articulatory movements for two speaking rates. The filled circles correspond to the slow rate and the unfilled circles to the fast rate.

labial consonant production and tongue tip muscle activity (superior longitudinal) associated with lingual consonant production increase for fast speech while genioglossus muscle activity for tongue body movement during vowel production decreases during fast speech. These results are illustrated for the sequence /ipip/ in Fig.8.

The first finding implies an increase in articulatory effort and an increase in the speed of articulatory movement: the production of both /p/ and /t/ requires a complete occlusion of the vocal tract, which must be produced more quickly and with greater effort during fast speech. The reduction in EMG activity for the vowel during fast speech, on the other hand, is compatible with the view that a vowel target has a built-in error or tolerance factor that can absorb the extra demands of

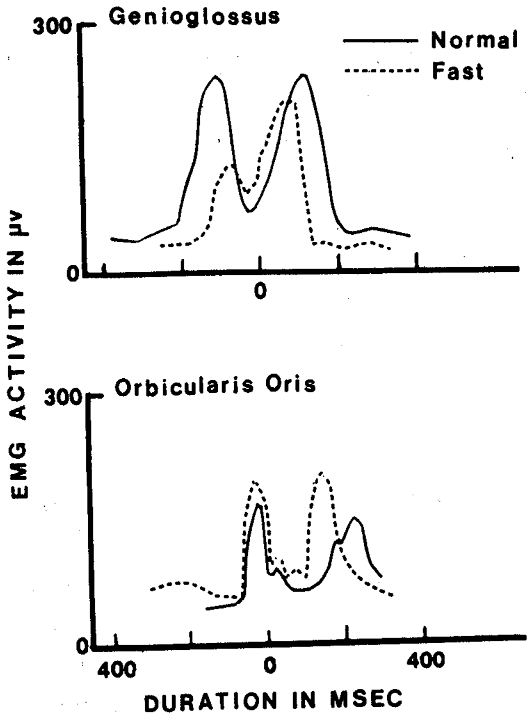

Fig.8. Averaged EMG plots for the genioglossus and orbicularis oris muscles for two different speaking rates. The utterance is /ipip/.

an increase in speech rate.

Two other interesting results appeared in these experiments. One was that lip muscle activity associated with rounding for /u/ also increases with an increase in speaking rate. This implies that the different effects of changes in speaking rate are related either to specific muscle systems or individual phonetic features rather than basic differences in phonetic categories. The other was that, for both subjects, an increase in speaking rate was accompanied by an increase in the frequency levels of both the first and second formants. Thus, since the acoustic triangle was not reduced toward the neutral schwa, articulatory undershoot during fast speech does not produce the same acoustic result as articulatory undershoot during destressed speech.

The most important aspect of the electromyographic speaking rate data is not the direction in which the amplitude of the signals change for consonants and vowels, but rather the fact that they change. Changes in both the timing and amplitude of the EMG signals with changes in speaking rate signified that the control of speech rate requires complex motor programming, and not simply a reordering of the timing of motor instructions (Lindblom, 1963).

SUMMARY

The major points discussed in this paper are as follows: First, our data suggest that anticipatory movements toward the second vowel in a VCV sequence begin during the closure period of the intervocalic consonant. This restricted coarticulatory field includes both the tongue body movement and lip rounding gesture 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 do not extend beyond an immediately neighboring segment. Unlike anticipatory effects, however, the appearance of carryover coarticulation effects depends on the phonetic identity of the particular segment on which these effects might act.

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 fact 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 implies a syllable-sized articulatory unit and argues against the operation of a scan-ahead mechanism at the segmental level. Rather, all features of both elements of the syllable are contained within the boundaries of that unit.

This does not necessarily mean, however, that a scan-ahead mechanism does not operate at another stage of speech production. The complex reorganization of commands accompanying changes in speaking rate suggests that the temporal features of a downstream segment are 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, it may well be that the reverse is true: the segmental input to the speech string is governed 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 scanahead mechanism.

Acknowledgment

The research reported here was supported by grants from the National Institute of Dental Research (DE-01774), National Institute of Neurological and Communicative Diseases and Stroke (NS-10424), and the National Science Foundation (GSOC 740 3725).

REFERENCES

Bell-Berti, F. and Harris, K. S. (1976) An EMG study of coarticulation of lip rounding. A paper presented at a meeting of the Acoustical Society of America, November.

Daniloff, R. G. and Moll, K. L.(1968) Coarticulation of lip-rounding. *J. Speech Hearing Res.*, 11; 707-721.

Gay, T. J. (1974a) Some electromyographic measures of coarticulation in VCV utterances. *SR-44*, 137-145.

Gay, T. J. (1974b) A cinefluorographic study of vowel production. J. Phonetics, 2; 255-266.

Gay, T. (1976) Articulatory movements in VCV sequences. J. Acoust. Soc. Amer., submitted for publication.

Gay, T. and Ushijima, T. (1974) Effect of speaking rate on stop consonant-vowel articulation. A paper presented at the Speech Communication Seminar, Stockholm.

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

Henke, W. (1966) Dynamic Articulatory model of speech production using computer simulation. Ph.D. Thesis, M.I.T.

Hughes, O. M. and Abbs, J. H. (1976) Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. Phonetica, in press.

Kozhevnikov, V. A. and Chistovich, L. A. (1965) Rech', Artikulyatsiya, i. Vospriyatiye. Trans. as Speech: Articulation and Perception. 1966. Washington, D.C.: Joint Publications Research Service, 30, 543.

Lindblom, B. E. F. (1963) Spectrographic study of vowel reduction. J. Acoust. Soc. Amer., 35; 1773-1781.

Lindblom, B. (1964) Articulatory activity in vowels. STL-QPSR, 3.

MacNeilage, P. F. (1970) Motor control of serial ordering of speech. Psychological Review, 77; 182-195.

MacNeilage, P. F. and DeClerk, J. (1969) On the motor control of coarticulation in CVC monosyllables. J. Acoust. Soc. Amer., 45; 1217-1233.

Öhman, S. E. G. (1966) Coarticulation in VCV utterances: Spectrographic measurements. J. Acoust. Soc. Amer., 39; 151-168.

DISCUSSION

Sawashima: What I understood by the data of Tom Gay is that, in a V_1CV_2 sequence, the articulatory transition from V_1 to V_2 is definitely affected by the presence of the intervocalic consonant. The deviation would be the suppression of the muscle activity in one case, or a consonant-specific gesture in the other case.

However, for me, the data do not reveal the fact that the influence, the coarticulatory effect, of V_2 is completely blocked by the intervocalic consonant.

You are opposed to Öhman according to your Fig.1, but your Fig.7 shows clearly, for slow speech particularly, that the tongue movement from /a/ to /i/ does begin prior to the closure of the intervocalic /p/. You said that "the tongue movement from the first vowel to the second vowel does not begin until after closure for the intervocalic consonant is completed," and further, "this was found to be a salient feature in the production of all VCV syllables." But your data appear to me to show that this is not true.

In a Japanese V_1CV_2 sequence, tongue movement from the first vowel to the consonant closure is clearly different when we compare, for example, /aka/ with /aki/ as seen in this slide. Thus, I would like to ask you whether you have consistent findings with your conclusion when you compare /api/ with /apa/, or /ipa/ with /ipi/, etc.

Secondly, I am a little bit puzzled in making time correspondence between EMG patterns and speech segments in Fig.3.

I would like to ask you to make it clear. In this recording, EMG peak associates with /k/ as well as the vowel /i/, and I would like to know the EMG patterns for utterances where /k/ is used as the intervocalic consonant.

Also I would like to ask you whether there was a difference in EMG peak for intervocalic /k/ between /kakap/ and /kakip/. If there was some difference, this also suggests that the tongue movement from the first vowel to the intervocalic consonant is affected by the post-vocalic vowel.

Gay: In response to your first question, movements toward the second vowel cannot be separated from those toward the consonant when going from an open to a closed vowel, for example, from /a/ to /i/. Your comment is correct, though, and I should have qualified that particular statement.

With respect to your second question, the line-up point in Fig.3 is the onset of voicing for the second vowel. Also, where /k/ is the intervocalic consonant, I believe that it and the first vowel are reflected by a single broad EMG peak.

Fujisaki: Although I am very much impressed by the amount of experimental data presented by the author, I feel that they are still not sufficient to support some of the statements made by the author. I should like to just mention that although it may be sufficient to show one counterexample to disprove a hypothesis, a much larger body of evidence is necessary to prove the validity of a new hypothesis such as the one presented by the author on the range of anticipatory and carryover effects. Though I agree with the author that carryover effects cannot entirely be accounted for by a simple inertial smoothing process, we cannot deny the fact that inertial smoothing still constitutes an important factor for which we have a number of results to show that it goes beyond an immediately neighboring segment. What we really need, I believe, is a more quantitative measure to decide whether or not certain effects exist under certain circumstances.

Gay: I fully agree that more data are needed to support any existing model of articulatory organization. Indeed, the purpose of the paper was primarily to show that a substantial amount of contradictory data still exists and that no present model can adequately account for these differences.

Hirano: When a young child learns speech, the minimum unit of speech, I think, is not the phoneme, not the syllable but the word. The extent of the carryover or anticipatory effect should greatly depend on how each individual is accoustomed to uttering a given word. It should also depend on the way he speaks. I appreciate your comments on this.

Gay: This is a very basic question and one which I am usually successful in repressing. However, I believe that we can learn much about both the nature of units and segmental organization by studying the variability that exists in articulatory flow.

Moll: In relation to your Fig.1, you conclude that the tongue does not begin movement to /a/ until the lips have made contact for /p/. Since the tongue position is expressed in terms of a coordinate system fixed on the upper skull and since the jaw is moving downward, adjustments must be being made to keep the tongue position constant. The point is that it would be more meaningful to separate the jaw component from the tongue measure.