Ann. Bull. RILP (1988) No. 22, 7-19

ELECTROMYOGRAPHY OF THE TONGUE MUSCLES DURING VOWELS IN /əpVp/ ENVIRONMENT

Thomas Baer*, Peter J. Alfonso**, and Kiyoshi Honda***

Introduction

Electromyographic recordings from the speech musculature have added much to our understanding of the processes involved in speech production. For organs that have a complex musculature, such as the tongue and larynx, simultaneous recordings from groups of the relevant muscles can particularly help to elucidating the mechanisms of control and their coordination during speech. Thus, for example, the intrinsic musculature of the larynx is reasonably well understood[1]. In contrast, there is surprisingly little data on the coordinated activity of the lingual muscles, especially in view of the tongue's central role as the major speech articulator. In this area, as with the larynx, Sawashima and his colleagues have made important contributions[2]. However, the data that are available for the lingual musculature have generally been recorded from no more than two or three muscles at a time[3,4], or have been collected with surface electrodes so that the contributions of individual muscles remain uncertain[5].

In recent years, there has been some discussion in the phonetics literature regarding distinctive feature systems for classifying vowels[6]. With respect to lingual articulation, three dimensions are usually used: front-back, high-low, and tense-lax. There have been a number of questions about how vowel position according to these dimensions translates to physiological activity, and the appropriateness of the dimensions has come under attack. Wood[7] has suggested replacing the dimensions high-low and front-back, describing the tongue's position in the oral cavity, with a four-way distinction in place where the vocal tract is most constricted, similar to the way consonants are classified. For vowels, the hypothesized distinctive places are palatal, palatovelar, pharyngovelar (high pharyngeal), and low pharyngeal. For the English vowels, this would produce the sets /i I e ɛ/, /u U/, /ɔ o/, and /æ ɑ/, respectively. Part of Wood's argument is that the traditional dimensions, high-low and front-back, do not correspond clearly to the functions of the lingual muscles, but here again, there is insufficient EMG data to clearly refute or support this claim. Many questions have also been raised about the tense-lax dimension. There are diverging thought on how this dimension might be implemented, and even what vowels it applies to. Raphael and Bell-Berti[4], based on limited EMG data from three subjects, have proposed two alternate mechanisms for the pairs /i I/, /e ɛ/, and /u U/: one corresponding to height and the other

* Haskins Laboratories
** Haskins Laboratories and University of Connecticut
*** Kanazawa Institute of Technology

corresponding more closely to the notion of "tension" in distinguishing tense-lax cognates. Wood[7] also includes the pairs /o ɔ/ and /ɑ ə/ and, based on radiographic evidence and model studies, argues for "degree of constriction" or effort in the muscle(s) primarily responsible for each pair. EMG data are needed to evaluate this claim.

The purpose of the present study, then, is to improve understanding of the control of the tongue during vowel production by analyzing simultaneously recorded EMG activity from the largest possible subset of the muscles responsible for tongue movements. To this end, we obtained simultaneous hooked-wire recordings from several extrinsic muscles of the tongue: the anterior and posterior genioglossus (GGA and GGP), styloglossus (SG), and hyoglossus (HG). To our knowledge, there are no other published studies reporting HG recordings with hooked wire electrodes, although recordings of muscle activity in the HG region obtained with surface electrodes have been reported[5]. In addition to these extrinsic muscles, recordings were also obtained from two accessory tongue muscles, the geniohyoid (GH) and mylohyoid (MH), and from the orbicularis oris superior (OOS) and the cricothyroid (CT) muscles. Jaw position was recorded using an opto-electronic movement transducer. EMG, movement, and acoustic signals were analyzed in parallel.

The data reported here have been discussed in preliminary reports[8,9]. Interactions between lingual and laryngeal activity have been reported elsewhere[10], and will not be considered further here. Radiographic data on the tongue have been collected using both lateral-cineradiography[11] and x-ray microbeam pellet tracking for the same subject producing similar utterances. However, in the absence of any simultaneous direct recordings of tongue position or configuration in the present experiment, we use acoustic measures to provide an indirect indication of tongue movements. The first two formant frequencies were extracted from the acoustic signal, and each was then processed in the same manner as the physiological signals. The first and second formant frequencies are used as indicators of vertical and horizontal tongue movements, respectively. Using this lingual EMG, jaw movement, and acoustic information, our goals are to 1) provide basic descriptive information on the functions of the extrinsic and accessory tongue muscles during vowel production, and 2) provide basic information on the organization of speech motor control by examining the relationships among these muscles for the same utterances, and 3) examine the implications of these results on feature systems describing vowel production.

Methods

Subject and Speech Material

The single subject for this study was a speaker of a suburban New York City dialect of American English. The speech material consisted of isolated utterances of the form /əpvp/, where V was one of the 11 vowels /i ɪ e ɛ æ ɑ ɔ o ʊ u ʌ/.

Anatomical Description and Electrode Placement.

Bipolar hooked-wire electrodes were prepared using the techniques described by Honda, Miyata, and Kiritani[12]. The electrode insertion paths are indicated in Figure 1. In the following, we will review the anatomy of the tongue musculature. However, since we are primarily concerned with tongue control, we will not discuss OOS or CT.

Percutaneous insertions from below the jaw were used for GGA and GGP. The genioglossus is the largest muscle of the tongue, its fibers arising from the inner surface of the symphysis of the mandible and radiating toward the root and dorsum of the tongue. Although the anterior and posterior fibers of the genioglossus are part of the same muscle anatomically, previous studies have shown that they can function separately during vowel production[2,3]. Thus, the two parts of the muscle are considered to be functionally distinct.

Per-oral insertions were used for the SG and HG muscles. The two insertions were made into the opposite sides of the tongue. Considering the SG first, its fibers take their origin from the styloid process and course forward and downward to the side of the tongue, where they run longitudinally toward the tip and blend obliquely with the hyoglossus. Contraction of the SG should pull the body of the tongue upward and backward toward the styloid process. Considering the HG, its fibers arise from the hyoid bone and course almost vertically upward to enter the sides of the tongue. Given the anatomical configuration of these fibers, the HG muscle should pull the body of the tongue downward and backward. As noted above, however, EMG behavior of this muscle during speech, as monitored by indwelling electrodes, has not previously been reported.

We consider the MH and GH to be accessory tongue muscles because they contribute indirectly to tongue movements. MH takes its origin on the mylohyoid line of the mandible. Its anteriormost insertions are in the oral floor at the midline and its posterior insertions are on the hyoid bone. The anterior fibers course downward and medially, while the posterior fibers course at a similar angle downward and medially but in a more posterior direction. This muscle thus forms a muscular floor for the tongue body, and its contraction should thus stiffen the floor, which would tend to elevate the tongue body. Finally, the fibers of GH originate on the mental spine of the mandible and course backward and downward to insert on the hyoid bone. Considering its anatomical configuration, this muscle might be expected to move the tongue body forward and slightly upward, through its effect on the hyoid bone, but it may also play a role in jaw lowering. It has also been found to cooperate with the cricothyoid in control of fundamental frequency[13].

EMG Measurements

EMG signals were rectified and integrated over a 5 ms window before sampling. The sampled signals were then smoothed using a 70 ms triangular window, and were ensemble-averaged across the repetitions of each vowel. The alignment point for ensemble averaging was the onset of voicing for the test vowel, as measured from the wideband speech waveform. An audio envelope waveform was obtained by rectifying and integrating the audio waveform as for EMG processing, followed by smoothing with a 70 ms triangular window.

Acoustic Measurements

Tongue position during vowel production was indicated by spectral analysis. The first and second formant frequency trajectories for each production were extracted at 10 ms intervals using LPC analysis. Ensemble average formant trajectories were obtained by aligning the 10 tokens at voice onset, as for EMG processing. Average formant frequency values for each vowel were obtained by pooling across the data points from 20 ms after the lineup point to the major deflection point in the trajectory, which of course varied across vowels. For example, there were 15 data points in the average for the short vowel /I/ and 21 points for the long vowel /e/.

Kinematic Measurements

Relative jaw movement was monitored with the aid of an optoelectronic movement transducer. Infrared light emitting diodes (LEDs) were attached to the undersurface of the chin and to the bridge of the nose. The transucer was used to monitor vertical position of the LEDs, and these signals were sampled and digitized at a rate of 200 per second for each channel. An estimate of jaw movement trajectories was obtained by subtracting the signal for the nose LED, which served as a head reference, from that for the chin LED.

Results

Acoustic

Figure 2 indicates the average F1 and F2 values for each of the 11 vowels produced in this experiment. In this figure, the front vowels and back vowels form two distinct groups, with the front vowels arrayed along a diagonal line and the back vowels clustered around a different, more nearly vertical, line. Among the front group, the tense vowels /i/ and /e/ are produced with relatively high and front positions, while their lax counterparts, /I/ and /ɛ/ are lower and more central. The acoustic data suggest a nearly linear continuum of tongue positions, from the highest and most anterior /i/ to the lowest and least anterior /æ/. Among the back vowels, the tense vowel /u/ is highest and its lax counterpart /U/ is lower and more central. The vowel /ɔ/ has the most posterior position, and /ɑ/ has the lowest position. The "mid-central" vowel /ʌ/ has a relatively low and back position in this space, suggesting a tongue position for this vowel similar to that for the low back vowels.

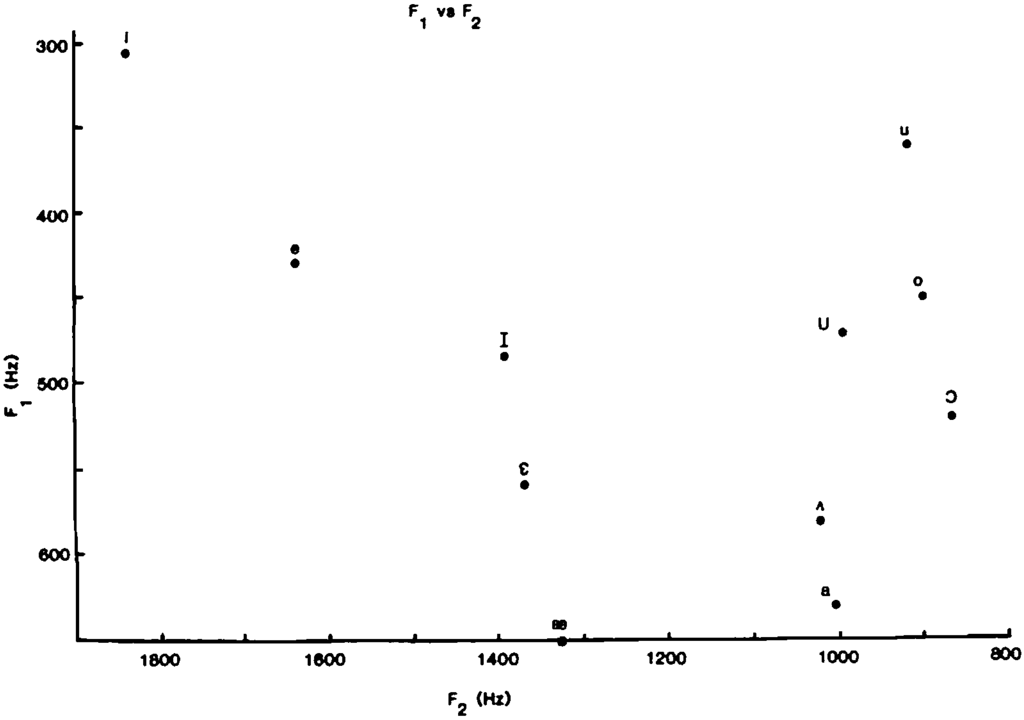

Figure 2

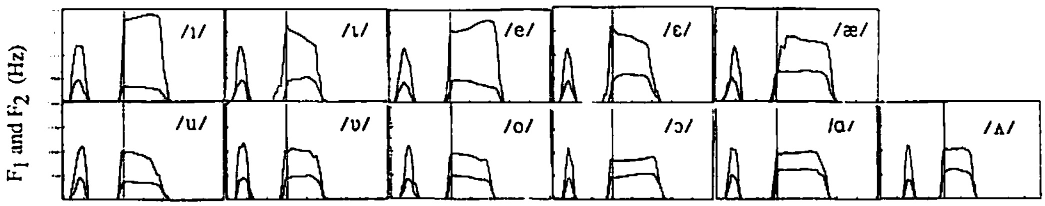

Figure 3

Figure 3 represents the ensemble average first and second formant frequency trajectories for each of the 11 vowels. The front vowels are shown along the top row, the mid and back vowels along the bottom row. Frequency is shown along the ordinate. The tick marks indicate 500 Hz intervals and extend to 2000 Hz. Time is shown along the abscissa at 100 ms intervals. The heavy vertical line represents the onset of voicing for the vowels. Thus, the formant trajectories indicate that the onset of the syllable initial schwa occurred approximately 250 ms before the onset of the vowel, and that the duration of the vowels varied from about 200 to 250 ms as a function of vowel type. These formant trajectories form the basis for the following discussion of horizontal and vertical tongue movements. We infer horizontal and vertical tongue movements in the discussion that follows from In the discussion that follows, we make reference to these formant trajectories to infer horizontal and vertical tongue movements.

EMG and Jaw Position

Figure 4 shows the ensemble averages for acoustic, EMG, and jaw movement channels. Each row represents a channel and each column represents a vowel type. The columns are organized into three groups, representing the front vowels, the back vowels, and /ʌ/. Tick marks along the abscissa represent 100 ms intervals. The heavy vertical line represents the onset of voicing for the vowel, which served as the lineup point for each token. Audio envelopes appear on the first row, and can be used as a guide to segment timing. Note that the time domain shown in Figure 4 corresponds to that shown in Figure 3.

Vertical jaw position waveforms are shown on the bottom row of Fig. 4. Since jaw position has such a major effect on tongue position, we will deal with these waveforms first. Each shows a pair of peaks, corresponding to the two /p/'s in the frame utterance. The level of the intervening valley reflects vowel height. Among the front vowels, /i/ is highest and /æ/ is lowest, as expected. Jaw height is about the same for the tense-lax pair /i I/ and also for the tense-lax pair /e ɛ/. Among the back vowels, tense-lax vowels /u ʊ/ share the highest jaw position, and /ɑ/ has the lowest position. The position for /ʌ/ is similar to that for /o/ and /ɔ/.

Next, considering the patterns of tongue EMG activity, it seems that the differences between activity patterns for the front vowels as a group versus the back vowels as a group are greater than the differences among the members of either group. This overall impression results not only from differences in the level of activity of a number of muscles, but also from qualitative differences in the patterns of activity of others, most notably the HG. Only the GGP shows a very inconsistent pattern of activity within groups, and this pattern is consistent with vowel height variations. It can be noted, in addition, that the overall pattern of EMG activity for /ʌ/ is more like that for

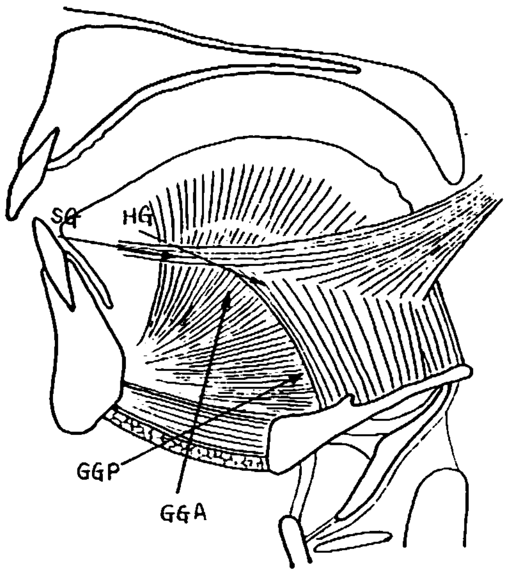

Figure 1

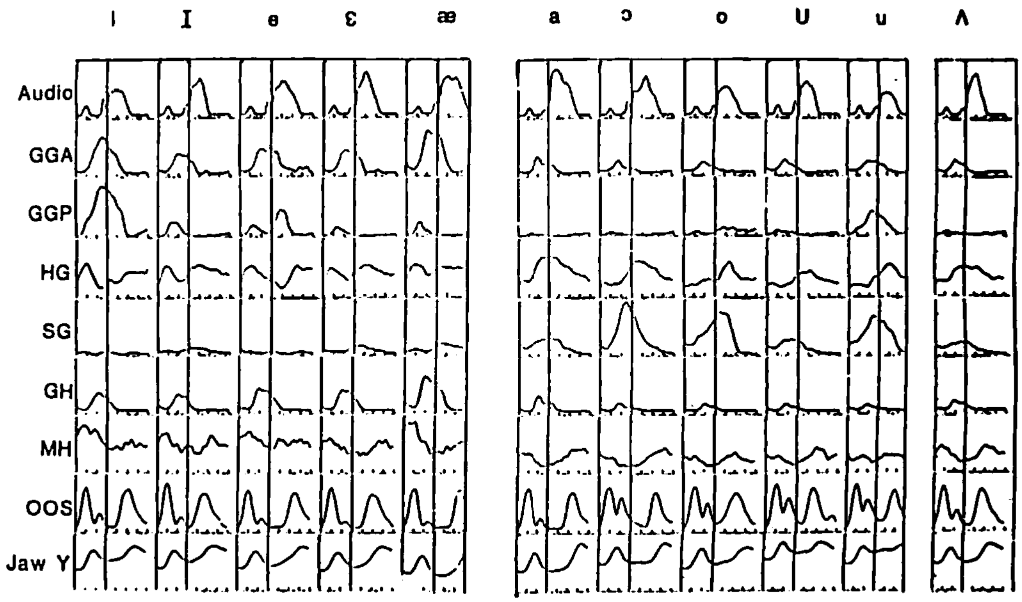

Figure 4

the back vowels, in agreement with the acoustic data. In the following, then, we consider each of the muscles individually.

GGA, shown on the second row, is much more active for front vowels than for back vowels. This result is consistent with those of previous studies, showing that GGA is associated with tongue fronting. Within the front group, peak GGA level is associated roughly with vowel height, except that highest peak activity is observed for the low-front vowel, /æ/. GGA exerts a downward force on the tongue and, in the dorsal region, there is also a forward component. If the tongue is stiffened transversely, or constrained laterally by the hard structures surrounding the oral cavity, and the oral floor is relatively stiff, then downward displacement of lingual tissue must result in anterior or posterior extension, due to the incompressibility of the tongue mass and the compliance of its surface tissues. When GGP or other muscles prevent posterior extension, anterior extension must be produced. Grooving may also be expected to occur. Both the anterior movement and the grooving make the GGA an important muscle for /i/ production. For /æ/, perhaps the lowering component is most significant, since the other muscle that pulls the tongue down, HG, would also have a backward component, and does not contribute to /æ/.

GGP, shown in the third row, is most clearly related to vowel height, as expected, in that it shows most activity for /i/ and /u/. By far, the greatest activity is observed for /i/. For the other front vowels, GGP show peaks of activity during the schwa voicing and stop occlusion. The function of this activity is not clear. In addition, for /e/ there is a second peak, apparently associated with diphthongization of this vowel. GGP shows a low level of activity for the lax vowels, /I ɛ U/.

Relative activity of the GGA and GGP support the notion that GGA and GGP are effectively different muscles with different functions. In particular, GGP pulls the tongue root forward, raising the dorsum, while GGA pulls the dorsum forward and down. To summarize the evidence, both are active during /i/, where they contribute jointly to fronting and raising, and where the GGA is antagonistic to the GGP in their influence on the coronal surface near the midline. Only the GGA is active during /æ/, which requires fronting and lowering, and only the GGP is active during /u/, which requires raising and backing. For /e/, acoustic data shown in Figure 3 suggest that this vowel starts at a relatively front position, and that raising movements then persist almost until occlusion for the final stop. GGA activity can be seen to begin 200 ms before the acoustic onset of the vowel, thus contributing to tongue fronting before voicing onset. Vowel-related activity of GGP begins about 50 ms before the acoustic onset and continues to increase to a maximum at about 100 ms after the lineup point, thus correlating with the tongue raising gesture. There is about 200 ms latency between peak EMG activity for the GGA and GGP.

HG and SG activity are represented on the next two rows of

Fig. 4. These two muscles show greater vowel-related activity for the back vowels than for the front vowels. Among the back vowels, HG shows more activity for the low ones and SG shows more activity for the high ones.

HG displays qualitatively different patterns for the utterances containing front vowels and back vowels. During the front-vowel utterances, a peak in HG activity occurs at about the time of lip closure for the initial stop. Its timing is about the same as the initial peak of OOS activity. Activity is then suppressed during the vowel, to a varying extent related to vowel height, and it increases again, so that it is relatively high during the final stop. The magnitude and timing of HG suppression is related in a rough way to the magnitude and timing of GGP activity, suggesting that these muscles function as antagonists. Maximum HG suppression occurs during /i/ and /e/, where GGP activation is greatest, and least HG suppression occurs for /æ/, where GGP activation is among the least. For /i/ and /e/, the peak of HG is almost simultaneous with the dip in GGP. There is a consistent latency between the prevocalic peaks in HG and GGP, with the HG always leading by about 50-100 msec. This pattern of activity may be associated with a gesture to lower and then raise the tongue during the stop closure[14,15].

For the back vowels, prevocalic HG activity is low, rather than high as for the front vowels. HG is activated during the vocalic period for all the back vowels. The peak is broadest and highest for the lowest back vowel, /ɑ/, consistent with Perkell's suggestion that HG is the major agonist for /ɑ/ articulation. The peak is lowest for the lax high back vowel, /ʊ/. For /o/ and /u/, the peak is skewed to the right, possibly associated with diphthongization movements for these vowels as suggested by Figure 3.

SG activty is negligible or small for all the front vowels. There is some activity for the two lowest front vowels, /ɛ/ and /æ/. The timing of the broad peaks for these vowels is later than those for GGA and GH, the muscles most clearly associated with these vowels. The activity pattern thus seems to be associated with adjustments in tongue position toward the end of the vocalic segment.

SG is active for all the back vowels, showing greatest activity for the high ones, /ɔ o u/, as expected. Thus, SG clearly exerts a backward and upward force while HG exerts a backward and downward force. There is clearly coordination between the SG and HG to control vertical tongue position for the back vowels. For the four high vowels, /ɔ o u ʊ/, the SG peak occurs before the HG peak suggesting that tongue backing and raising are not time-locked. This hypothesis is somewhat consistent with the vowel trajectories shown in Figure 3, although the acoustic effect of lip rounding makes this correspondence less clear in the case of the high back vowels. Alternatively, this pattern of activity may be to stabilize tongue position as the jaw goes down and then comes back up for stop closure. In any case, the two muscles appear to act synergistically to move the tongue backward for these vowels and to control its vertical position.

Finally, GH and MH muscles are shown on the sixth and seventh rows, respectively. Both show generally higher activity during front vowel utterances than during back vowel utterances.

GH activity indicates a single peak of activity associated temporally with the vowel. Activity for each of the front vowels is greater than that for any of the back vowels. This suggests that GH participates in fronting, probably by pulling the hyoid bone forward. For most of the back vowels, activity is negligible, but there is clearly a peak for /ɑ/. Highest activity occurs for /æ/. Thus, for both the front and back vowels sets, greatest activity occurs for the vowel with the lowest jaw position. This suggests, then, that GH also contributes to jaw lowering.

The activity pattern of the MH varies with vowel identity, with overall level higher for the front vowels. For all vowels, there is at least some decrease in activity near the lineup point - that is, in the period associated with the vowel. Activity during this period, although reduced for all vowels, is still somewhat higher for the front vowels. The highest activity is associated with the prevocalic period for the front vowels, especially /æ/ and /i/. The peak of this prevocalic activity occurs at about the same time as the initial peak of the OOS, suggesting that it is associated with the initial stop. For the front vowels, the timing of this peak is also similar to that for HG. However, the onset of the increase occurs earlier than that for either OOS or HG. The beginning of the decrease in MH activity coincides approximately with the increase of GGA and GH activity. We are uncertain about the function of this MH activity, but it may be associated with stiffening of the oral floor during bilabial occlusion in the environment of front vowels. We do not have a suitable explanation for the high level of activity preceding /i/ and /æ/. There may be some relationship with the high level of GGA activity for these vowels.

Discussion

Probably the most striking aspect of the results is the quantal difference in the neuromuscular organizations for front- and back-vowel utterances. There were distinctive differences in level of activity for GGA, HG, SG, and MH, such that for each of these muscles, a vowel could be classified as either front or back using a single criterion of EMG level. For both HG and MH, not only average level of vowel-related activity but also overall pattern of activity through an utterance was different for front versus back vowels. (Note that, for this speaker, /ʌ/ appears to be a back vowel, according to these criteria.) Although results of a single subject study must be treated with caution, this study seems to lend support to the physiological reality of "front-back" as a distinctive feature for vowels.

Within the front and back classes, vertical level of the jaw was correlated with vowel height, and three levels could be distinguished, corresponding to high, mid, and low vowels. GGP showed nearly quantal behavior across vowels, demonstrating activity mostly for the high tense vowels /i/, /u/. (It was also active for the latter part of /e/, which diphthongizes toward /i/.) SG showed high activity for /u/, but no more activity for /U/ than for /ɑ/. Thus, for the lax vowels, it seemed likely that height was controlled mostly by the jaw, and the tense vowels /i/, /u/, and /e/ were characterized by greater activity in their major agonists (and hence, more extreme articulations) than their lax counterparts. This result for the SG is different than that of Raphael and Bell-Berti[4], who found no difference in SG activity for their tense-lax pairs. The patterns of muscle activity for the other vowels do not separate them into tense-lax classes in any obvious way. The pair /o ɔ/ differ more in timing of EMG activity than in amount. Both HG and SG have later peaks for /o/ than for /ɔ/.

We note that there is EMG activity associated with tongue movements prior to and following the period associated with the vowel. This activity shows systematic variation with vowel identity, and thus bears further consideration. In particular, we expect that the pre-vocalic HG and MH activity in the front vowel environment affected vowel-related tongue activity for these and possibly other muscles, and that this effect would be quite different in other stop environments.

Acknowledgement

This work is supported by NIH Grants NS-13870 and NS-13617 awarded to Haskins Laboratories.

References

1) Sawashima, M. and Hirose, H.: Laryngeal gestures in speech production. In P.F. MacNeilage (Ed.), The Production of Speech, New York: Springer-Verlag, 11-38, 1983.
2) Miyawaki, K., Hirose, H., Ushijima, T., and Sawashima, M.: A preliminary report on the electromyographic study of the activity of the lingual muscles. Ann. Bull. RILP, 9, 91-106, 1975.
3) Smith, T.: A phonetic study of the function of the extrinsic tongue muscles. Working Papers in Phonetics (UCLA) 18, 1971.
4) Raphael, L.J. and Bell-Berti, F.: Tongue musculature and the feature of tension in English vowels, Phonetica, 32, 61-73, 1975.
5) MacNeilage, P. F. and Sholes, G. N.: An electromyographic study of the tongue during vowel production. J. Speech Hear. Res., 7, 209-232, 1964.
6) Fischer-Jorgensen, E.: Some basic vowel features, their articulatory correlates, and their explanatory power in

phonology. In V. A. Fromkin (Ed.), Phonetic Linguistics, San Diego: Academic Press, 79-99, 1985.

7) Wood, S.: X-ray and model studies of vowel articulation, Working Papers (Lund U. Dept. Linguistics), 23, 1-192, 1982.

8) Alfonso, P.J., Baer, T., and Honda, K.: Lingual control and vowel articulation. Proc. 11th ICA, 3-10, 1983.

9) Alfonso, P.J., Honda, K., and Baer, T.: Coordinated tongue muscle activity during /əpvp/ utterances. In M.P.R. Van den Broecke and A. Cohen, (Eds.), Proc. Tenth Int. Cong. Phon. Sci., IIB, Dordrecht: Foris Publications, 390-394, 1984.

10) Honda, K.: Variability analysis of laryngeal muscle activities. In I. R. Titze and R. C. Scherer (Eds.), Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control, Denver: Denver Center for the Performing Arts, 127-137, 1983.

11) Alfonso, P.J. and Horiguchi, S.: Vowel-related lingual articulation in /əCVC/ syllables as a function of stop contrast., Proc. Eleventh Int. Cong. Phon. Sci.,II, 41-44, 1987.

12) Honda, K., Miyata, H., and Kiritani, S.: Electrical characteristics and preparation technique of hooked-wire electrodes for EMG recording, Ann. Bull. RILP, 17, 13-22, 1983.

13) Honda, K.: Relationship between pitch control and vowel articulation. In D. M. Bless and J. H. Abbs (Eds.), Vocal Fold Physiology: Contemporary Research and Clinical Issues, San Diego: College Hill Press, 286-297, 1983.

14) Houde: R.A.: A study of tongue body motion during selected speech sounds. Ph.D. Thesis, University of Michigan, 1967.

15) Alfonso, P.J. and Baer, T.: Dynamics of vowel articulation. Language and Speech, 25, 151-173, 1982.