ISCA Archive
http://www.isca-speech.org/archive

2nd International Conference on
Spoken Language Processing (ICSLP 92)
Banff, Alberta, Canada
October 12-16, 1992

MUSCLE FORCES IN VOWEL VOCAL TRACT FORMATION

Katherine S. Harris,† Eric Vatikiotis-Bateson,‡ and Peter J. Alfonso†††

Haskins Laboratories, New Haven, Connecticut, †City University of New York,
‡ATR Research Laboratories, Kyoto, Japan, †††University of Illinois, Champaign

10.21437/ICSLP.1992-267

ABSTRACT

This experiment is an extension of earlier work, [1, 2] aimed at understanding the relationship of the principal muscle forces underlying vowel tongue shape. The experiment was done in two parts. A single male talker, TB, was recorded for both speech samples. Stimuli for both parts of the experiment were multiple tokens of disyllables of the form / pVp/, with the vowels /EE, IH, AY, EH, AE, A, AW, OOH, UW/. For Part One of the experiment, electromyographic recordings were made from the principal extrinsic muscles of the tongue using hooked wire electrodes. For Part Two, x-ray microbeam recordings of tongue body, lip, and jaw were made on the same subject for the same inventory.

In general, tongue X and Y position correlated for front pellet and front vowels, but the situation is more complicated for rear pellets. While the correlations of extrinsic tongue muscle activity with tongue pellet positions were as one might expect, no vowel in the set can be considered as a lax equivalent of any other, as Wood's [3] model suggests.

INTRODUCTION

Although there is a tradition of classifying vowels along the dimensions of high vs. low, and back vs. front, it is well known that this classification does not represent the vowel relationships in a conventional geometrical space. However, it is, of course, commonly supposed that the vowels of a given language have some internal structure. While the dimensions of vowel space can be extracted by a purely computational procedure, [4], there is considerable difference of opinion as to why a particular organization holds. Proposals range from the suggestion that certain universals of vowel systems may be a consequence of the properties of speech perception to Wood's proposal that the dimensions are set by the anatomy of the muscles of the oral region. Thus, the vectors of force generated by the muscles must set the basic constraints on the dimensions for the vowels. While the action of the muscles of the tongue have been modelled at least twice, the empirical information necessary for refinement of muscle-based vowel dimension models is lacking. The present experiments are a modest beginning towards filling this gap, and are a continuation of previous work [1, 2].

Methods

The experiment was done in two parts. The talker for both was TB, an author of the previous papers in this series. The speech consisted of isolated utterances of the form / pVp/, produced in short, quasi-random sequences. In all cases, multiple tokens were averaged with respect to a common lineup point at vowel release. For Part One of the experiment, there were ten tokens of each type. For Part Two of the experiment, failures of the x-ray microbeam tracking procedure reduced the number of tokens to about five for each vowel.

For Part One of the experiment hooked wire electromyographic recordings were made from the anterior and posterior genioglossus (GGA and GGP), geniohyoid (GH), mylohyoid(MH), hyoglossus (HG), styloglossus (SG), orbicularis oris superior (OO), and cricothyroid (CT) muscles. The latter muscle was recorded for purposes irrelevant to the present experiment and will not be discussed further here. Since the lip Y position correlates highly with OO activity, OO recordings have been used primarily to line up results. The results of Part One have been previously reported, although we reanalyzed the data, changing the lineup point for averaging slightly, with small differences in the numerical values of the muscle trajectories reported here and previously. The size of the signal recorded from different muscles varied widely. For this reason, the EMG values have been normalized to a value that represents the difference between the maximum observed for any utterance and the noise baseline, transformed into percent.

For Part Two of the experiment, x-ray microbeam data were taken on the same subject for the same inventory. Pellets were placed on mid- and rear-tongue, lower lip and jaw and their trajectories traced, using a system in use until recently at the Institute of Logopedics and Phoniatrics at the University of Tokyo[5].

Results

Results are shown in two forms; in tabular form, in Tables 1 and 2, and in Figures 1 through 8. The tables show positional information on the front of the tongue and rear tongue and jaw for the nine vowels. Position measures were made at the point of maximum jaw opening. This has the effect of maximizing opening for /EE/, /UW/ and /AY/, all of which are heavily diphthongized for this subject.

Table 1. *A summary of the positions in arbitrary units, of the front (F) and rear (R) tongue and jaw (J) in the horizontal (X) and vertical (Y) directions.*

Vowel	FX	FY	RX	RY	JY
EE	564	-296	63	76	-875
AY	523	-305	-5	37	-909
IH	435	-404	-53	-194	-887
EH	420	-424	-58	-228	-917
AE	308	-520	-155	-254	-922
A	191	-594	-267	-443	-926
AW	157	-653	-288	-402	-917
U	190	-619	-217	-358	-875
UW	158	-617	-226	-96	-871

Although electromyographic measures were made of the mylohyoid muscle, the records were quite noisy, and hence, have not been included. It is clear, however, that front and back vowels differ in that mylohyoid is active in raising the floor of the mouth for the front vowels, and decreases activity for low back vowels.

The results for jaw opening, as measured by the Y position of the jaw pellet, are much as we might expect. Jaw opening is smallest for /EE/, and largest for /A/. The vowel /AY/ has a larger jaw opening than /IH/ or /EH/, using the criteria described above.

Overall, as described very crudely by the two pellet positions, the front vowels form a continuous series, in that x and y pellet positions form a series of almost parallel lines—that is, the jaw and the tongue muscles work together so that the tongue moves position but does not change shape very much. The two vowels /IH/ and /EH/ are extremely similar, so that is hard to see that for this speaker, there is a privileged relationship between /EE/and /IH/. Of course, the situation would unquestionably be more complicated if we examined, for example, the amount of tongue body grooving. The four back vowels, /A/, /AW/ /OOH/ and /UW/ behave quite differently, in that x and y positions are not even approximately parallel.

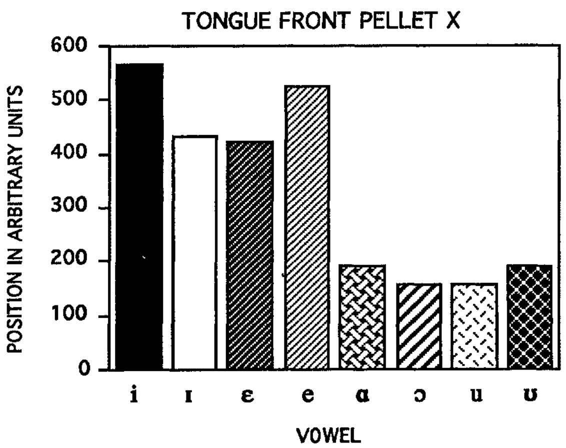

Fig. 1. Extreme positions for front tongue pellet in the X dimension, for all vowels.

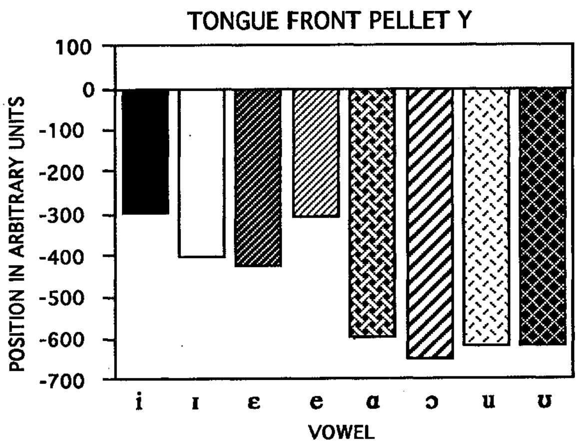

Fig. 2. Extreme positions for front tongue pellet in the Y dimension, for all vowels.

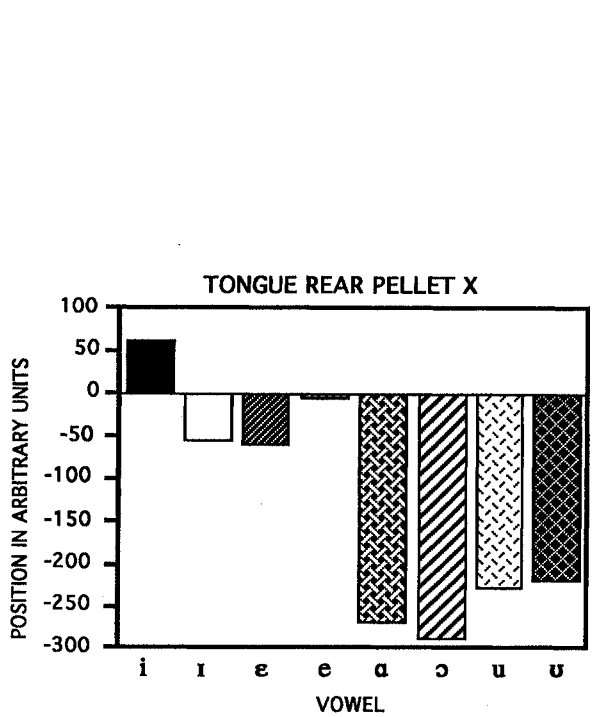

Fig. 3. Extreme positions for rear tongue pellet in the X dimension, for all vowels.

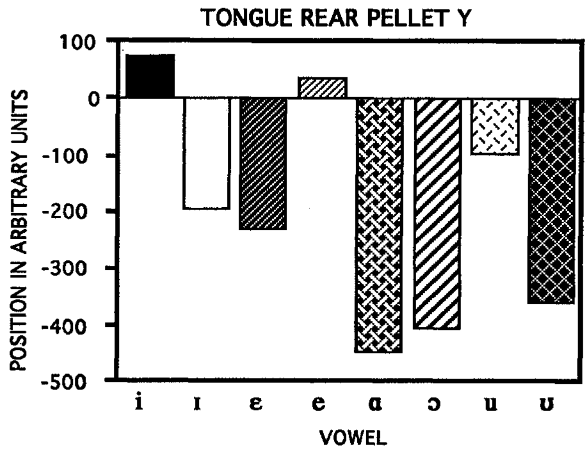

Fig. 4. Extreme positions for rear tongue pellet in the Y dimension, for all vowels.

When we turn to the muscle data, (see Table 2 and Figures 5 through 8) we see that the smooth relationships, even among the front vowels, seen in the movement data disappear.

Table 2. *Peak values of muscle activity in microvolts converted into a percentage of the total range for each muscle, for all vowels.*

Vowel	GGA	GGP	GH	HG	SG
EE	96	100	51	17	17
AY	71	7	69	47	11
IH	59	11	52	42	21
EH	63	6	69	33	21
AE	100	7	100	44	10
A	37	5	35	100	40
AW	22	4	22	68	100
U	28	10	17	51	36
UW	43	57	45	50	78

This point is seen more clearly if we compare muscle activity with tongue position. The two high front vowels, /EE/ and /AY/ are adjacent in front tongue pellet position, but very different in muscle activity pattern (Table 2, Figures 1 through 4). Apparently, the jaw position causes the difference in pellet position, rather than tongue muscle activity. The vowels /EE/ and /AY/ (Figure 5) are quite different particularly in GGP activity, although /IH/ and /EH/ (Figure 6) are very similar. It should be particularly noted that /AE/, which has essentially the same GGA activity as /EE/, has a very different spatial position. Apparently, the jaw repositions the tongue together with the tongue muscles in the series /EE/, /EH/ /IH/, but GGA and GH take over to move the tongue frontwards when the tongue position is low.

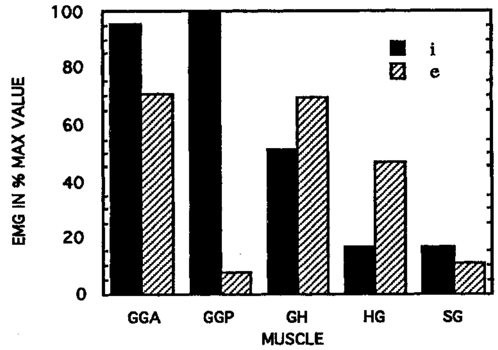

Fig. 5. ***Peak values of muscle activity in microvolts converted into a percentage of the total range for each muscle, for high front vowels.***

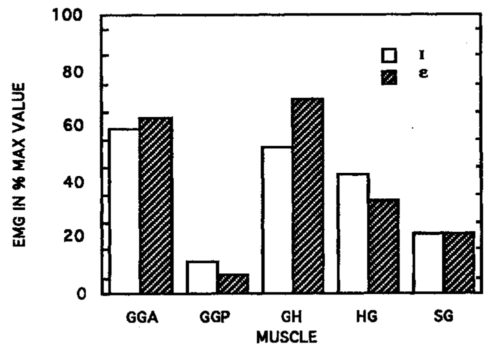

Fig. 6. ***Peak values of muscle activity in microvolts converted into a percentage of the total range for each muscle, for low front vowels.***

Of course, given the well-known reciprocity between jaw and tongue activity in shaping the front vocal tract, it may be that the solution for balancing tongue and jaw is special to this speaker.

The vowels /A/ and /AW/ are compared in Figure 7. The three muscles which raise and front the tongue are all relatively inactive, while the HG and SG are both active, with SG pulling the tongue relatively back for /AW/ while HG pulls the tongue relatively down for /A/. Figure 8 compares /UW/ and /OOH/. Again, there does not seem to be a close relationship of pattern for the two vowels. The vowel /UW/ has its greatest activity for SG.

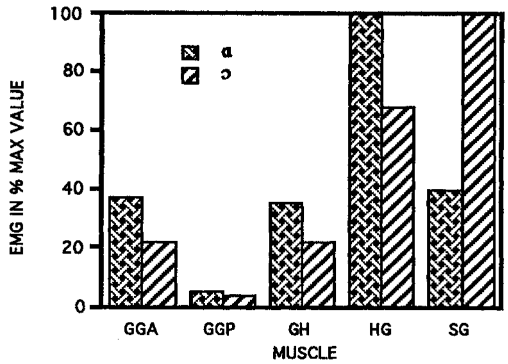

Fig. 7. ***Peak values of muscle activity in microvolts converted into a percentage of the total range for each muscle, for low back vowels.***

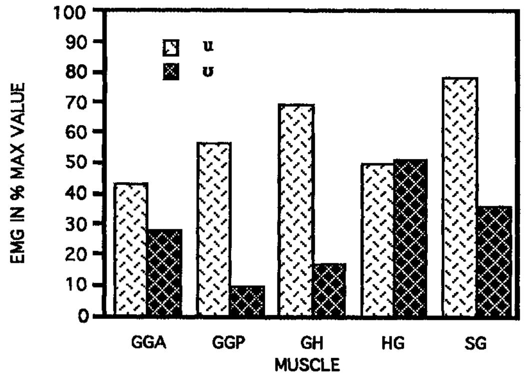

Fig. 8. ***Peak values of muscle activity in microvolts converted into a percentage of the total range for each muscle, for high back vowels.***

Overall, we can see that muscle activity for the vowels conforms reasonable well to the idea of trajectories of pull up and front, up and back, and down and back. However, no vowel appears to be a "lax" copy of any other, nor does any muscle have exclusive control of the position of any vowel.

REFERENCES

[1] P. J. Alfonso and T. Baer. "Dynamics of vowel articulation," *Language and Speech, 25,* 151-173, 1982.

[2 T. Baer, P. Alfonso, and K. Honda. "Electromyography of the tongue muscles during vowels in /epVp/ environment," *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, 22,* 7-19, 1988.

[3] S. Wood. "A radiographic study of constriction location for vowels," *Journal of Phonetics, 7,* 25-43, 1979.

[4] R. Harshman, P. Ladefoged, and L. Goldstein. "Factor analysis of tongue shapes," *Journal of the Acoustical Society of America, 62,* 693-707, 1977.

[5] S. Kiritani, K. Itoh, and O. Fujimura. "Tongue-pellet tracking by a computer-controlled x-ray microbeam system," *Journal of the Acoustical Society of America, 57,* 516-520, 1975.