VOWEL-RELATED LINGUAL ARTICULATION IN /∂CVC/ SYLLABLES AS A FUNCTION OF STOP CONTRAST

Peter J. Alfonso
Univ. of Connecticut and Haskins Laboratories
Storrs (06268) and New Haven (06511) CT, U. S. A.

Satoshi Horiguchi
Univ. of Tokyo
Tokyo 113, Japan

ABSTRACT

Lateral cineradiographic pellet-tracking of tongue blade and tongue body movements along with formant frequency trajectories show that articulation of the vowels /i/ and /u/ in /∂CVC/ syllables vary by as much as 8-10 mm as a function of the stop consonant environment in which they are produced. The magnitude of the variation is related to the identity of the stop and vowel.

INTRODUCTION

This study represents one of a set of experiments completed or underway at Haskins Laboratories that aim to study the dynamics of vowel articulation. What distinguishes the set of studies from previous work is that the dynamics of vowel articulation is studied by examining data representing the four accessible measurement levels of speech production, namely 1)muscle activity (by hooked-wire electromyography), 2)corresponding movements of the speech structures (primarily by tracking the movements of lead pellets glued to the lips, tongue, and jaw by lateral cineradiography or x-ray microbeam), 3)representative speech acoustic signals, and 4)perceptual testing of selected auditory segments to determine whether or not the underlying articulatory movements provide relevant linguistic information. In the first experiment to use these measurement techniques, a single subject's productions of disyllables of the form /∂pVp/, where V represented one of eleven vowels, were analyzed. As an example of the conclusion that can be drawn from multi-level analysis, the results showed that vowel-related tongue horizontal and vertical movements can have different time constraints in labial environments. When fronting and raising occur together they are necessarily time-locked since they are caused primarily by the same muscle, genioglossus. Backing and raising, on the other hand, can and do occur independently, since they are caused by different muscle groups. While vertical movements for all vowels and horizontal movements for front vowels always began about the moment of implosion for the initial stop, horizontal movements for back vowels began much earlier, even before the acoustic onset of the schwa. Acoustic and perceptual analysis indicated that anticipatory tongue movements were linguistically significant since listeners were able to identify the vowels when presented with only a portion of the schwa segment.

*This work was supported by NIH Grant NS-13617 to Haskins Laboratories. The authors gratefully acknowledge the assistance of Drs. Thomas Baer and Hiroshi Muta in data collection and analysis.

More recently, a second subject has been run following the same procedures but increasing the data base in two ways: 1) the same set of vowels were produced in labial, alveolar, and velar stop environments, and 2) more extensive tongue EMG insertions representing the complete set of lingual extrinsic and accessory muscles. Analysis of EMG data from the second subject in general supports the temporal differentiation in vertical-horizontal tongue movements in the first subject in that the muscles of the tongue appear to be distinctly organized for front versus back vowel production [2]. Biomechanical descriptions of tongue dynamics, such as the relationship between genioglossus and fronting-raising in the first subject, will be enhanced significantly by mapping the EMG data for the second subject onto his x-ray data. While the overall purpose of the combined EMG and x-ray runs is to study the dynamics of vowel articulation in a fashion similar to the initial experiment [1], the paper presented here will focus on the x-ray run. The purpose of the paper is to give a quantitative description at the movement and acoustic levels of the variation in vowel-related tongue articulation that occur as a function of producing the vowels /i/ and /u/ in labial, alveolar, and velar stop environments.

METHODS

An adult male native speaker of American English with a New York City dialect produced two repetitions of /∂CVC/ disyllables where /C/ represents /p/, /t/, or /k/ and /V/ represents one of eleven vowels. The initial and final consonants in each utterance were identical. Lateral cineradiographic films were made at a rate of 60 frames per second while the subject produced isolated syllables at a rate of about one every two seconds. Figure 1 represents a schematic diagram of the midsagittal plane of the vocal tract sketched from a single frame of the x-ray film. The figure shows the location of lead pellets glued to the tongue blade, middle and rear dorsal areas of the tongue surface at the midline, and a jaw pellet attached between the lower central incisors. Measurements of pellet movements were made on a frame-by-frame basis with the aid of

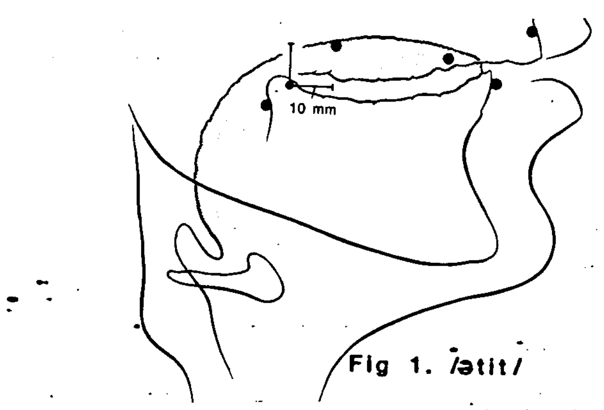


Fig 1. /ətit/

a digitizing tablet. Pellet locations were fixed with respect to two reference positions, the lead pellet attached between the upper central incisors and the point marked by a template reflecting maxillary boundaries. The latter reference location is shown in Figure 1 as the point of origin in the 10 millimeter (mm) grid. Pellet displacement data shown in the following section are displayed in reference to the point of origin. For example, positive vertical and positive horizontal displacement trajectories indicate movements above and to the right of the origin, respectively. Displacement values represent calibrated units in mm.

RESULTS

Considering, first, the dynamics associated with the articulation of /i/, Figure 2 shows vertical movement trajectories for the tongue blade (top panel), middle (mid panel) and rear dorsal (bottom panel) pellets for production of /əCiC/. Within each panel, the solid line represents trajectories for /əpip/, the dashed line represents trajectories for /ətit/, and the dotted line represents trajectories for /əkik/. Except for infrequent instances when pellet locations were not visible and therefore not tracked, each of the trajectories represent the average displacement of two tokens per syllable. Although trajectories represent combined jaw and tongue displacements, they primarily represent tongue displacements since jaw movements were negligible for this speaker (see Table 1). The ordinate represents calibrated units in mm from the origin (see Figures 1 and 4). The distance between vertical markers along the abcissa represents 100 ms intervals. The solid vertical line at zero time represents the initial consonant release. Thus, the time-span of the trajectories is 800 ms, segmented as a 300 ms and 500 ms interval before and after initial consonant release. The acoustic onset of the schwa and the acoustic duration of the vowel varied with the identity of the stop but, in general, schwa onset occurred from about 150 to 200 ms before consonant release and vowel duration varied from 200 to 250 ms. Implosion of the initial consonant varied from 75 to 125 ms before consonant release. Not unexpectedly, the top panel of Figure 2 shows greater tongue blade vertical displacement during the consonantal segment in the alveolar environment compared to the labial and velar environments. Notice, however, that tongue blade trajectories during the vocalic segment appear similar in all stop consonant environments. On the other hand, the relative vertical displacement trajectories for middle and rear dorsal pellets are clearly different from tongue blade trajectories. The middle and bottom panels of Figure 2 show that the corresponding tongue locations were much higher in the velar environment than in the labial and alveolar environments throughout the syllable. That is, the vertical differentiation begins very early, at about initial consonant closure for the rear pellet and at the earliest point of measurement for the middle pellet, and continues throughout the vowel and final consonant.

Figure 3 shows horizontal displacement trajectories for the same three pellets. Notice that fronting is clearly differentiated during the vocalic portion of the syllables and that the velar context produces the most front /i/.

Table I lists jaw and tongue displacement measurements in mm from the origin taken from trajectories shown in Figures 2 and 3 at the moment of vocalic peak acoustic amplitude. Acoustic peak amplitude occurred an average of 125 ms after consonant release for the two tokens of /əpip/ and /ətit/ and 175 ms after release for /əkik/. Also shown are average first and second formant values in Hz. The formant values represent the average of five samples taken at five ms intervals beginning 10 ms before through 10 ms after the moment of peak amplitude. Large differences in tongue shape across the three stop environments during /i/ production are clearly indicated in Table I. For example, the middle dorsal area of the tongue is displaced about 10 mm anteriorly and superiorly in the velar stop environment relative to the labial and alveolar environments. The large effect of velar closure on tongue body positioning for /i/ is reflected in the formant frequency values as well; the lowest first formant and the highest second formant values occur for /əkik/.

Finally, the influence of alveolar and velar stop production on the entire tongue shape during /i/ production is shown in Figures 1 and 4. The figures are schematic diagrams of the midsagittal plane of the vocal tract sketched from x-ray frames corresponding to the temporal interval represented in Table 1. Vocal tract shape at peak vowel amplitude for /i/ during /ətit/ is shown in Figure 1, and the corresponding interval during /əkik/ is shown in Figure 4. A comparison of the figures shows that alveolar and velar stop consonant constrictions produce dramatic differences in the tongue shape for /i/ articulation. Taken together, the data indicate that the dynamics of tongue articulation for /i/ is clearly different in each of the three stop contexts, that the velar context yields the most high and front vowel articulation.

Similar analyses have been made for /əCuC/ syllables, but space limitations rule that they be presented in abbreviated form. For example, Figure 5 shows vertical displacement trajectories of the middle dorsal pellet for /əCuC/. Note that the velar context produces the highest tongue body movements, and that the vertical differentiation begins before the acoustic onset of the schwa.

Figure 2. Vertical Displacement /i/

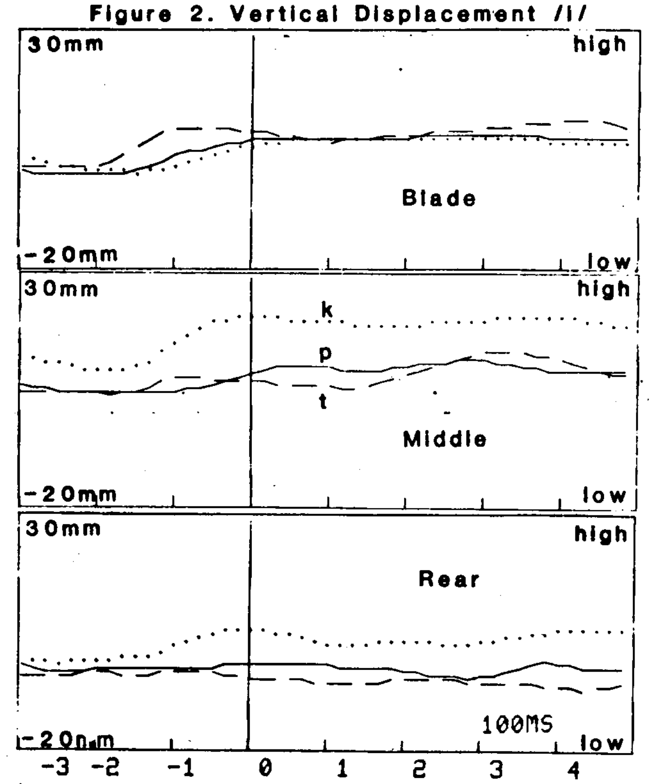
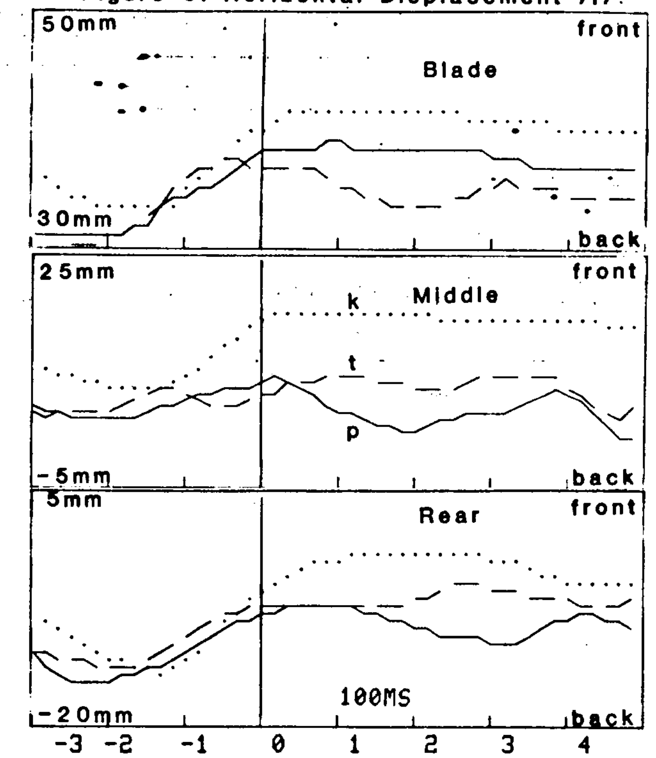


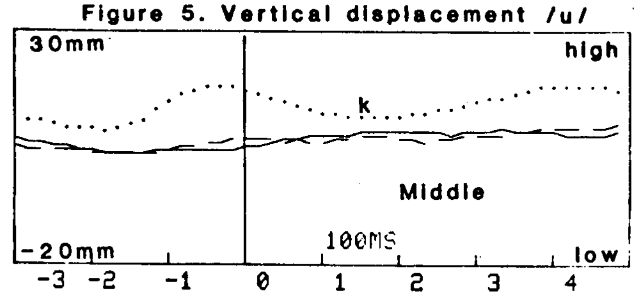
Figure 3. Horizontal Displacement /i/.



A comparison of Figures 2 and 5 shows that the vertical differentiation in vowel articulation is less dramatic in /u/ than in /i/. Figure 6 demonstrates that the alveolar context yields greater tongue blade raising than the labial and velar contexts for /u/ and that the vertical differentiation begins at about acoustic onset of the schwa. Furthermore, a comparison of Figures 2 and 6 shows greater tongue blade vertical differentiation in /u/ than /i/. The tongue blade is lower in labial and velar contexts in /əCuC/ relative to /əCiC/.

Horizontal displacement trajectories yield similar trends. First, the velar stop constriction has less influence on vowel related vertical displacement for /u/ than for /i/ and, second, the relative maximum raising usually cooccurs with relative maximum fronting. For example, the middle dorsal pellet in /əkuk/ is more front and high than in /əpup/ and /ətut/. Finally, Table I also shows displacement and formant frequency values for the vocalic segment in /əCuC/ syllables. The vocalic peak amplitude occurred 140 ms after release for /əpup/ and 200 ms after release for /ətut/ and /əkuk/. Once again, appreciable

Figure 5. Vertical displacement /u/



differences in vowel related tongue displacement as a function of stop constriction location are observed. Although the center frequency of the first and second formant also showed large variation with stop context, the expected relationship between formant frequency and tongue vertical and horizontal displacement is not observed, presumably due to the stop consonant influence on pharyngeal cavity width and the movements of other structures, most likely the lips or larynx.

DISCUSSION

The results show that for this subject the dynamics of vowel articulation for /i/ and /u/ can be altered significantly as a function of the identity of the stop in which they are produced. Assuming that the middle and rear dorsal pellets yield appropriate estimates of tongue body shape, Table I shows that tongue body configurations for /i/ vary by nearly 10 mm, with the velar context producing the most high and front /i/. For /u/, variation in tongue body configuration is about 6 mm, and again the velar context produces the most extreme lingual displacement.

Figure 6. Vertical Displacement /u/

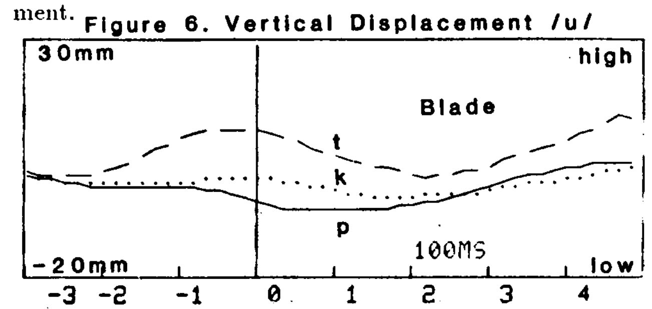


TABLE I.
Average pellet displacement in mm from origin and average first and second formant values in Hz for the vocalic segment of /əCiC/ and /əCuC/. See text for definition of measurement techniques.

	Vertical Displacement				Horizontal Displacement				Formant	
	Jaw	Blade	Middle	Rear	Jaw	Blade	Middle	Rear	F1	F2
əpip	0	7.2	9.6	-2.4	45.6	38.4	9.6	-8.0	310	1799
ətit	1.6	8.0	9.6	-4.8	46.4	33.6	8.0	-6.4	312	1722
əkik	0.8	7.2	19.2	3.2	48.8	41.6	17.6	-4.8	301	1851
əpup	0.8	-5.6	8.8	-0.8	46.4	28.0	4.0	-16.0	382	941
ətut	1.6	0.8	7.2	-1.6	46.4	30.4	4.8	-11.2	351	1194
əkuk	0.8	-2.4	12.8	0.8	48.0	30.4	9.6	-12.0	389	1007

Tongue blade configurations also vary greatly as a function of the stop. Whereas the magnitude of fronting and raising generally covary in tongue body articulation, they do not seem to covary for tongue blade movements. The variation for tongue blade configurations is about 6 mm in the vertical dimension, the differentiation being larger in /u/ than in /i/ with the alveolar context producing the greatest displacement. In the horizontal dimension, tongue blade displacement varies by as much as 8 mm, the velar context producing the most front tongue blade displacement.

Comparisons between movement data reported here with previously published data are not straight-forward since either the method used to measure tongue displacement or the design of the experiments differ significantly. For example, the patterns reported here can be compared with corresponding data taken from a cinefluorographic experiment that similarly measured the effect of the stop consonant on vowel-related vertical displacement [3]. The two experiments are in agreement in that vowels in velar context are produced with greater vertical displacement relative to labial and alveolar contexts. However, the published displays of the cinefluorographic trajectories appear to indicate that the magnitude of the variation during the vocalic segment is no greater than 5 mm. For a number of reasons, a more appropriate comparison of the magnitude of the variation can be made between experiments that employ pellet-tracking. X-ray microbeam pellet-tracking was used to estimate the variation in a large number of repetitions of the vowels /i/, /a/, and /ae/ occurring in a wide variety of consonantal environments [4,5]. Though not specified in the text, the figures indicate that the variation approximates the 8-10 mm maximum variation reported here. Finally, there are many acoustic based studies on the topic (e.g. [6,7]). The first and second format frequency values reported here are in good agreement with these data.

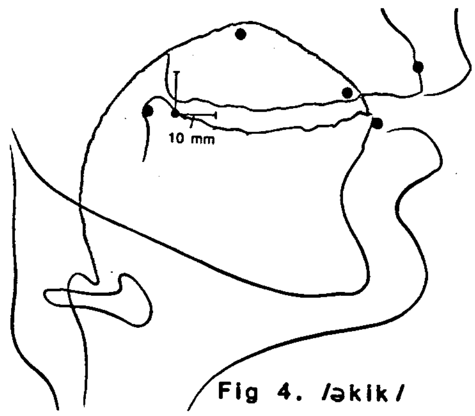


Fig 4. /əkik/

REFERENCES

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

[2].P.J. Alfonso, K. Honda, and T. Baer. Coordinated tongue muscle activity during /∂pVp/ utterances. *Proceedings of the Tenth International Congress of Phonetic Sciences*, IIB:390-394, 1984.

[3].P.F. MacNeilage and J.L. DeClark. On the motor control of coarticulation in CVC monosyllables. *J. Acoust. Soc. Am.*, 45:1217-1233, 1969.

[4].J.S. Perkell and W.L. Nelson. Articulatory targets and speech motor control: A study of vowel production. *Speech Motor Control*, Pergamon, New York, 1982.

[5].J.S. Perkell and W.L. Nelson. Variability in production of the vowels /i/ and /a/. *J. Acoust. Soc. Am.*, 77: 1889-1895, 1985.

[6].K.N. Stevens and A.S. House. Perturbation of vowel articulation by consonantal context: An acoustical study. *J. Speech Hearing Res.*, 6:111-128, 1963.

[7].K.N. Stevens, A.S. House, and A.P. Paul. Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation. *J. Acoust. Soc. Am.*, 40: 123-132, 1966.