Journal of Phonetics (1975) 3, 25–33

The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English

Lawrence J. Raphael

Haskins Laboratories, New Haven, Conn., U.S.A. and
Herbert H. Lehman College of the City University of New York, U.S.A.

Received 28th October 1974

Abstract: A series of two electromyographic experiments was designed to determine the nature of the muscular activity underlying the articulation of CVC syllables in which identical vowels differed in duration because of the voicing characteristic of the consonant which followed them. Results indicate that the most reasonable hypothesis explaining the durational differences is the one which posits a sustention of muscular activity in the articulatory gesture of the vowel preceding voiced consonants, relative to the gesture for vowels preceding voiceless consonants. It is noted that the acoustically determined differences between vowels, the differences between the durations of the muscular-articulatory gestures for the vowels, and the temporal displacement of the final consonant peaks generally show remarkably similar values.

The differences between the durations of vowels preceding voiced and voiceless consonants in English is well documented in the phonetic literature (Kenyon, 1951; Gimson, 1962; Locke & Heffner, 1940; Peterson & Lehiste, 1960; House, 1961). Investigators have described and/or commented on the perceptual consequences of these differences (Denes, 1955; Jakobson, Fant & Halle, 1967; Noll, 1960; Raphael, 1972), and have theorized as to whether the variation in vowel duration is a physiologically mandated behavior, one that is learned, or, to some extent, both (Delattre, 1962; Elert, 1964; Zimmerman & Sapon, 1958; Chen, 1970; House, 1961; Peterson & Lehiste, 1960).

Little, however, has been discovered or written about the physiological activity which must underly durational differences, no matter what their cause. The present study was undertaken in order to specify the muscular activity governing the articulatory gestures for vowels preceding both voiced and voiceless consonants. The studies referred to above, whether based on impressionistic evidence or on the analysis of acoustic records, suggest three hypotheses for the mechanism which renders vowels longer in duration before voiced than before voiceless consonants. The first, and perhaps the simplest, posits a greater duration of muscular activity for a vowel preceding a voiced consonant than for one preceding a voiceless consonant. Under this hypothesis final consonant articulations of either voicing type would be more or less identical, with the same time of onset relative to the offset of the preceding vowel.

The second hypothesis posits muscular activity of the same duration for vowels in both the voiced and voiceless environments. The durational difference would then be effected by a difference in the timing of the onset of muscular activity of the following consonants in relation to the offset of preceding vowel activity: relatively earlier in the voiceless case and relatively later in the voiced case. Such differences have been found in lip and jaw movements for stops (Kozhevnikov *et al.*, 1965; Ohala *et al.*, 1968; Kim & MacNeilage, 1972; Leanderson & Lindblom, 1972; Chen, 1970), although the magnitude of these differences does not appear to be great enough to account for the durational difference between English vowels (MacNeilage, 1973).

The third hypothesis merges the first two and posits differences in both the duration of muscular activity for vowels and in the relative timing of the onset of the muscular activity for the following consonants.

Experiment I

Procedure

A series of real-word, minimal pair, CVC, test utterances was constructed in which the articulation of the initial consonants and vowels would be essentially controlled by a muscle or set of muscles different from and independent of that which controlled the articulations of the final consonants. For example, in the minimal pair *leaf–leave*, it was assumed that the initial consonant and the vowel would be controlled by lingual muscles, whereas the final consonant would be under the control of labial muscles. (This proved to be the case for a pair such as *leaf–leave*, but for other pairs, namely those containing back, lip-rounded vowels, the separation of muscle function was not as clear as had been desired. For pairs such as *bought–bawd* and *moat–mowed* there was genioglossus activity for the vowel as well as for the final consonant. A more anterior electrode placement might reduce or eliminate this activity. Thus, some of the data concerning onset of muscle activity for the final consonant was obscured, although by no means completely). Of the six minimal pairs used, three were of a labial-to-lingual configuration (*mowed–moat*, *bawd–bought*, *moos–moose*), and three of a lingual-to-labial configuration (*leave–leaf*, *thieve–thief*, *lab–lap*).

There were two subjects in the experiment. Both read the words in isolation from a series of 10 randomized lists. At least 12 and as many as 19 tokens of each type were used to produce the averaged EMG curves.

The muscles explored for labial articulation were the orbicularis oris and the depressor anguli oris. Lingual articulation was investigated by recording EMG signals from the genioglossus muscle.

Concentric needle electrodes of standard Disa type were used for insertions into the muscles. Both the EMG output and a voice trace were recorded on magnetic tape for subsequent computer processing. The onset of voicing was used as a reference line-up (zero) point in the data manipulation.

Results and discussion

The only effect consistently found was that of greater duration of muscular activity in the articulation of vowels preceding voiced consonants. Figure 1 shows the most common manifestation of this effect:[1] The peaks associated with the vowel articulation occur almost

[1]In order to maximize, visually, the temporal relationships between EMG traces, the peaks for all vowels and consonants have been equated in height, regardless of their actual microvolt values. These values are, however, recorded. In Fig. 1, the microvolt values on the left-hand ordinate are those

simultaneously in both the voiced and voiceless cases; there is a sustention of muscular activity in the voiced case, relative to the voiceless case; the onsets of the muscular activity for the following consonants occur at approximately the same time relative to the offset of muscular activity for the preceding vowel; the onset durations and slopes for the muscular activity associated with the following consonants are generally equivalent. Certainly there is no durational difference between the onsets of consonant activity (relative to the preceding vowel) on the order of the durational differences between vowels as determined from acoustic records. For the utterances and subject of Fig. 1, the average vowel duration for *thief* was 150 ms, and that for *thieve* 360 ms. The durational difference between the muscular activity underlying the vowel articulations, with reference to the time each EMG curve reaches its base line, is on the order of 220 ms, quite close to the 210 ms difference between the durations of the vowels in the acoustic measurement.

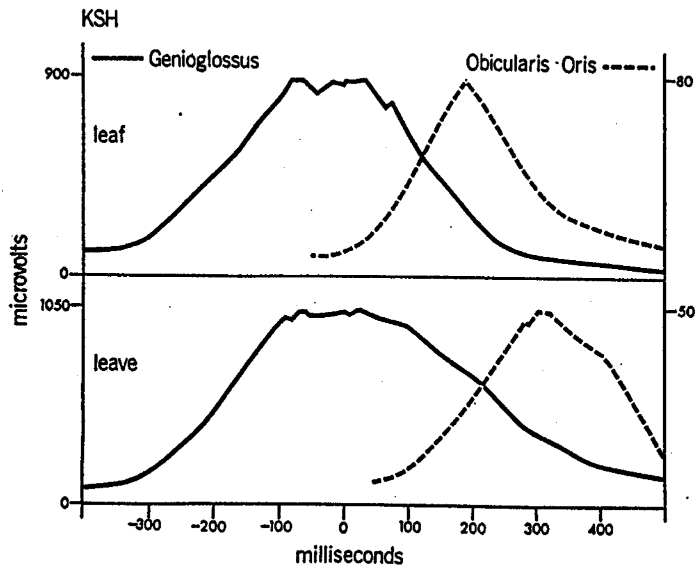

Figure 1 Typical EMG activity for two syllables contrasting in the voicing characteristic of the final consonant.

This sustention of muscular activity following the peak for the vowel was found for both subjects in most cases. Figure 2 shows typical examples of the averaged durational differences between the EMG signals caused by the sustention. Figure 3 displays the temporal displacement of the terminal consonant peaks associated with the vowels of Fig. 2. It may be observed that these final-consonant peaks were displaced from each other by time values approximately equal to both the durational differences between the EMG signals for the preceding vowels and their acoustically determined durational differences. The data for the

of the vowel (genioglossus) gesture, and those on the right-hand ordinate are those of the final consonant (orbicularis oris) gesture. In Figs. 2–5, the microvolt values for the vowel preceding the voiceless consonant and for the voiceless consonant itself are shown on the left-hand ordinate; those for the vowel preceding the voiced consonant and for the voiced consonant itself, on the right-hand ordinate. The actual values shown are peak values for vowel or consonant.

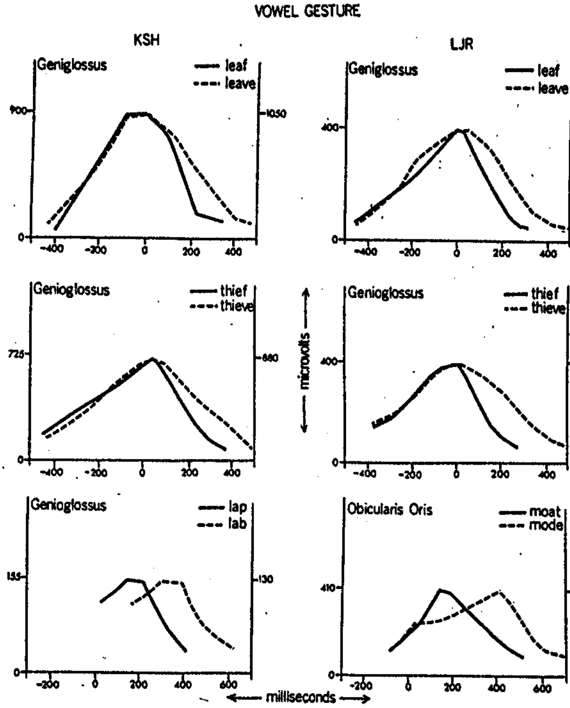

Figure 2 Paired EMG signals for identical vowels, one preceding a voiced and the other a voiceless, syllable-final consonant.

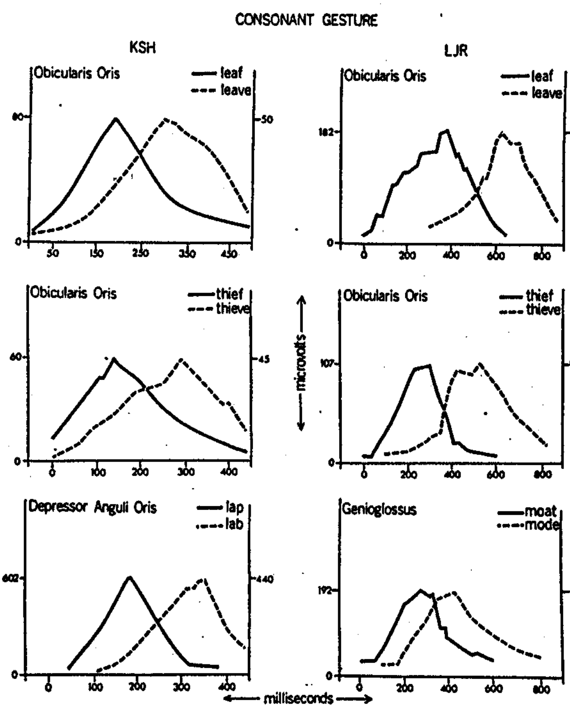

Figure 3 Paired EMG signals for voiced/voiceless syllable-final consonants.

Table 1 Comparison of vowel duration differences as determined spectrographically and by EMG measurements with temporal displacement of final consonant EMG peaks

	Duration differences (ms)		
	Vowel (Acoustic)	Vowel (EMG)	Consonant peak displacement (ms) (EMG)
Subject LJR			
leaf–leave	260	230	210
bought–bawd	155	110	115
moose–moos	180	160	80
moat–mowed	190	185	135
thief–thieve	300	290	170
lap–lab	175	180	165
Subject KSH			
leaf–leave	175	185	250
bought–bawd	90	100	90
moose–moos	100	120	110
moat–mowed	125	160	150
thief–thieve	210	220	200
lap–lab	105	85	130

vowel duration differences (both acoustic and EMG) and the temporal displacement of the EMG peaks of the final consonants are summarized in Table I.

One other articulatory strategy can be found underlying the durational differences between vowels: in two cases there is a delay in the onset and peaking of muscular activity for the vowel preceding the voiced consonant (Figs 2, 3, bottom graphs). This, in turn, causes a delay in the peak of muscular activity for the vowel, so that even though the slopes of the off-sets of EMG activity are virtually identical in both voiced and voiceless cases, the separation of the final voiced and voiceless consonant peaks is still achieved in this case as the result of the difference in timing and duration of vowel articulation. It may be possible, since the initial consonants in these utterances are semi-vocalic in nature, that part of the durational difference usually carried by the vowel is absorbed by the preceding consonant. Although the acoustic records do not reveal any consistent differences between the durations of these initial consonants, the EMG signals for the initial /l/ and /m/ in the voiced syllables show a slower onset and later peak than do those in the voiceless syllables. Thus the data do not provide a consistent explanation of this effect. Further, there remains the question of why one subject shows the effect for /l/ and not for /m/, and the other for /m/, but not for /l/.

Experiment II

Procedure

The minimal-pair utterances of the second experiment were disyllables which began with schwa, followed by /p/. The interconsonantal vowel was variously /i ɪ e ɛ æ a ʌ ɔ o ʊ u/. The final consonant was either /p/ or /b/. Utterances of these types provided negligible coarticulation effects, for the muscles investigated, between consonant and vowel, or between the vowel and the final consonant. Of the two subjects who provided data in this experiment, one had also participated in Experiment I. A third subject (who had also provided data in Experiment I) read an alternate list of utterances ending in /k/ or /g/ from which only final consonant EMG data was obtained.

Data was obtained from the orbicularis oris muscle for the labial stop consonants and from the mylohyoid muscle for the alternate utterances ending in velar stops. Vowel data was obtained from the genioglossus and inferior longitudinal muscles. Electrode insertions were made as described by Hirose (1971) for the obicularis oris, mylohyoid, and genioglossus muscles. The insertion to the inferior longitudinal muscle was made at the lower tongue surface near the back of the anterior third, approximately 1 cm from the lateral margin and roughly parallel to the lower surface of the tongue at a depth of approximately 5 mm.

The EMG signals were stored on magnetic tape for subsequent data processing. The onset of voicing of the interconsonantal vowel was used as a reference line-up (zero) point for the data manipulation and displays.

Results and discussion

As in Experiment I, there was a greater duration of muscular activity in the articulations of vowels proceding voiced consonants than in those preceding voiced consonants. For the utterances used in this experiment, both subjects showed this main effect for all muscles and for all vowels for which data was obtained. Figure 4 (a), (b) displays the genioglossus data for the two subjects of this experiment. The similarity of the vowel sustention effects, as evidenced in the EMG curves, among subject LJR in Experiment I (Fig. 2) and in Experiment II, and subject FBB in Experiment II is readily apparent. The similarity is

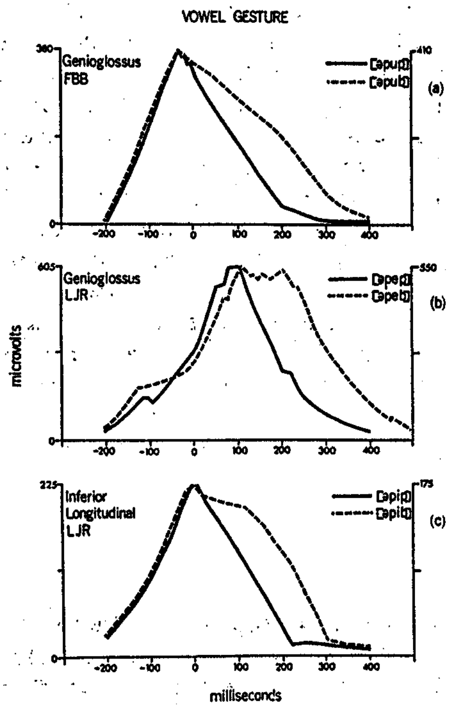

Figure 4 Paired EMG signals for identical vowels, one preceding a voiced and the other a voiceless, syllable-final consonant.

Table II **Comparison of vowel duration differences as determined spectrographically and by EMG measurements with temporal displacement of final consonant EMG peaks (for syllables ending in /p/ *vs* /b/)**

	Duration differences (ms)		Consonant peak displacement (ms) (EMG)
	Vowel (Acoustic)	Vowel (EMG)	
Subject LJR			
/i/	135	125	145
/ɪ/	60	70	65
/e/	130	140	120
/ɛ/	55	55	65
/æ/	85	75	100
/ʌ/	30	40	45
/a/	85	80	105
/ɔ/	115	105	100
/o/	155	165	150
/ʊ/	45	55	60
/u/	110	115	110
Subject FBB			
/i/	150	140	150
/ɪ/	85	95	75
/e/	170	175	150
/ɛ/	50	50	45
/æ/	130	—	145
/ʌ/	65	75	80
/a/	180	—	195
/ɔ/	110	—	120
/o/	205	190	185
/ʊ/	65	60	70
/u/	145	150	140

Table III **Comparison of vowel duration differences as determined spectrographically with temporal displacement of final consonant EMG peaks (for syllables ending in /p/ *vs* /b/)**

	Vowel duration difference (Acoustic—ms)	Consonant peak displacement (EMG—ms)
Subject KSH		
/i/	135	130
/ɪ/	45	45
/e/	130	140
/ɛ/	60	70
/æ/	155	145
/ʌ/	45	50
/a/	185	165
/ɔ/	115	100
/o/	135	125
/ʊ/	45	60
/u/	115	105

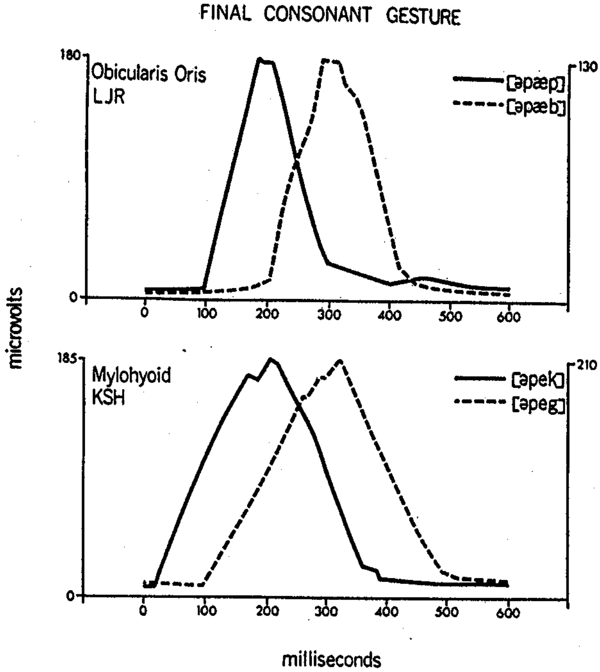

Figure 5 Paired EMG signals for voiced/voiceless syllable-final consonants.

further reflected in the EMG curve for the inferior longitudinal muscle for subject LJR [Fig. 4 (c)].

The displacement of the final consonant peaks is illustrated in Fig. 5 for one of the subjects of this experiment and for the subject who read the alternate list of utterances ending in /k/ or /g/. As in the first experiment, the acoustic durational differences, as determined from spectrograms, the EMG durational differences, and the temporal displacement values of the EMG peaks for the final consonants show remarkably similar values (Tables II and III).

Conclusion

The data presented here provide strong confirmation for the first hypothesis presented above. That is, the acoustically measured durational differences long observed between vowels preceding voiced and voiceless consonants are primarily controlled physiologically by motor commands to the muscles governing the articulators which are active in the formation of vowels. The timing of these commands is generally such that after the peak of the articulatory-muscular activity has been reached, the articulators are maintained (although not statically) in shapes and positions appropriate for vowels somewhat longer when they precede voiced consonants.

This research was supported by NIDR grant DE-01774.

References

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.

Delattre, P. C. (1962). Some factors of vowel duration and their cross-linguistic validity. *Journal of the Acoustical Society of America* 34, 1141–1143.

Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America* **27**, 761–764.
Elert, C-C. (1964) *Phonologic Studies of Quantity in Swedish*. Uppsala: Almqvist and Wiksells.
Gimson, A. C. (1962). *An Introduction to the Pronunciation of English*. London: Edward Arnold Ltd.
Hirose, H. (1971). Electromyography of the articulatory muscles: current instrumentation and technique. *Haskins Laboratories Status Report on Speech Research* SR-25/26, 73–86.
House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America* **33**, 1174–1178.
Jakobson, R. & Halle, M. (1967). Tenseness and laxness, supplement to *Preliminaries to Speech Analysis*. Cambridge, Mass.: MIT Press.
Kenyon, J. S. (1951). *American Pronunciation*, 10th ed. Ann Arbor: George Wahr Publishing Co.
Kozhenikov *et al.* (1965). *Speech: Articulation and Perception*. Washington, D.C.: United States Department of Commerce Clearinghouse for Scientific and Technical Information.
Leanderson, R. & Lindblom, B. E. F. (1972). Muscle activation for labial speech gestures. *Acta Otolaryngologica* **73**, 362–373.
Locke, W. N. & Heffner, R-M. S. (1940). Notes on the length of vowels II. *American Speech* **15**, 74–79.
MacNeilage, P. (1972). Speech physiology. A chapter in (J. Gilbert, Ed.) *Speech and Cortical Functioning*. New York: Academic Press, pp. 1–72.
Noll, J. D. (1960). The perceptual significance of certain acoustical correlates of consonant voicing contrasts. Unpublished Ph.D. dissertation, University of Iowa.
Ohala, J., Hiki, S., Hubler, S. & Harshman, R. (1968). Transducing jaw and lip movements in speech. Paper presented at the 76th meeting of the Acoustical Society of America. Abstract of above published (1969) in *Journal of the Acoustical Society of America* **45**, 324.
Peterson, G. E. & Lehiste, I. (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* **32**, 693–703.
Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of voicing of word-final consonants in American English. *Journal of the Acoustical Society of America* **51**, 1296–1303.
Zimmerman, S. A. & Sapon, S. M. (1958). Note on vowel duration seen cross-linguistically. *Journal of the Acoustical Society of America* **30**, 152–153.