The effect of position in utterance on speech segment duration in English

D. Kimbrough Oiler

Departments of Psychology and Linguistics, University of Texas at Austin, Austin, Texas 78712

(Received 18 February 1971; revised 17 March 1972 and 3 July 1972)

The duration of speech segments as a function of position in utterances (initial, medial, final) was studied. In the first experiment seven English speakers read nonsense utterances of the form "say a [bab], say a [babab], say a [babab], etc." Spectrograms were used to determine the duration of speech segments in the readings. Final syllables were found to be longer than nonfinal syllables. Final-syllable vowel increments were approximately 100 msec. Final-syllable consonant increments were less than vowel increments; for instance, absolute final consonant increments were about 20 msec. Also word-initial consonants were found to be lengthened by 20-30 msec over medial consonants. Subsequent experimentation demonstrated with English nonsense words that (1) final-syllable and initial-consonant lengthening occur in utterances with various intonational patterns (imperative, declarative, interrogative); (2) final-syllable lengthening occurs in word-final and phrase-final positions as well as in utterance-final position; and (3) final-syllable and initial-consonant lengthening occur in various kinds of syllables, including syllables with diphthongs, with fricative consonants, with voiceless stops, with consonant clusters, and with no final consonants (i.e., CV syllables). These studies report durational increments of particularly great magnitude for absolute final fricative consonants. Explanations of the lengthening effects are discussed. One theory suggests that lengthening in certain utterance positions is a learned aspect of language which cues listeners concerning the location of boundaries of words, phrases, or sentences. Explanations based on hypothesized properties of the speech production process are also discussed.

Subject Classification: 9.5.

INTRODUCTION

Phoneticians have noted the existence of final-syllable lengthening2 in many languages; for instance, Russian (Zlatoustova), English, Spanish, German, French (Delattre), and Swedish (Lindblom). But the significance of final-syllable lengthening has not yet been adequately assessed. It is quite possible, for instance, that the regularity of the data reflects some inherent property of perceptual or articulatory control processes. Yet few researchers have attempted explanations of final-syllable lengthening.

One prerequisite to such explanation is a careful investigation of position-in-utterance duration effects. It is necessary to answer as precisely and as fully as possible the following questions: (1) What part of the final syllable (vowel or consonants) carries the durational increment? (2) Is the final position the only position in utterances where syllables are lengthened? (3) Do position-in-utterance duration effects vary with stress? (4) How much difference in duration of speech segments occurs as a function of position-in-utterance?

Lindblom has conducted a study which is ideally designed to answer the above questions. His data are based on readings by Swedish subjects of a systematically controlled inventory where the segmental structure of all syllables is comparable. His results show a considerable final-syllable lengthening effect for both vowels and consonants.

In the present study, the first experiment represents a replication and extension of the Lindblom data with English-speaking subjects. In the remaining experiments we attempt to determine: (1) if the lengthening effects occur in all kinds of utterances (declarative, interrogative, imperative); (2) if the lengthening effects occur in word-final and phrase-final positions as well as in utterance-final position; and (3) if the effects occur in various types of syllables.

I. EXPERIMENT 1

A. Methods

Inventory: As in Lindblom’s study, a set of nonsense words of the form [bab], [babab], [babab], etc., were embedded in a carrier phrase. Words ranged from one to five syllables in length with all possible primary-stress patterns accounted for. Thus there were five words which were five syllables in length—one with primary stress on the first syllable, one with primary stress on the
second syllable, and so on. Table I consists of the inventory plus a list of English words with corresponding stress patterns.

Subjects: The subjects were seven male students at the University of Texas at Austin. All were native speakers of American English.

Recording: Subjects were given a few minutes to learn to read the nonsense words with the correct stress patterns. If a subject had difficulty with a stress pattern, he was asked to make use of the example words, and if he still had difficulty, the experimenter would assist by pronouncing the word. Subjects were instructed to maintain constant rate and constant volume. They were asked to pronounce each [b] with a clear plosive burst since the segmentation of the recordings would have been far more difficult with the fricativized [b]'s that are common in rapid American speech. Also, in order to insure the possibility of consistent segmentation of word-final [b]'s, subjects were instructed to pronounce all final [b]'s with a plosive release.

At least four samples of each utterance were recorded from each subject. Three subjects read the inventory in the order presented in Table I, while the other four read a randomized ordering of the utterances. The results revealed no important differences between the data of subjects who read the items in order of increasing word length and those who read in random order.

Segmentation: For each utterance a wide-band spectrogram was made on a Kay Sonagraph. The scale magnifier was set so as to reproduce between 80 and 3600 Hz.

The segmentation of the acoustic signal was carried out by reference to spectrographic evidence of lip closure and lip opening. The great majority of the utterances were produced with clear plosive bursts, and consequently the problem of determining the times of closure and opening was usually easy to solve. Fricativized utterances were discarded and replaced by new sample utterances whenever possible.

Two segmentation procedures were used—one in which the appearance and disappearance of high-frequency spectrographic information (2000-3600 Hz) was the segmentation criterion (two subjects), and one in which the appearance and disappearance of lower-frequency information (80-2000 Hz) was the criterion (five subjects). In the former case the resulting vowels were 10-20 msec shorter and the consonants 10-20 msec longer than in the latter case. None of the durational effects of position-in-utterance or word length appear to have been altered by the change in segmentation procedure. In addition, the results were checked by making oscillograms of 13 of the tape-recorded utterances. Segmentation of the oscillograms produced results that correlated very highly (r=0.98) with the original spectrographic data. Furthermore, a set of 15 of the utterances was segmented by reference to the amplitude displays of the oscillograms. Again the resulting data correlated very highly (r=0.98) with the original spectrographic results. The values for particular segments determined by the oscillographic or amplitude display methods were within 10 msec of the original spectrographically determined values in 88% of the cases. The study's basic data (viz., those concerning final-syllable lengthening and initial consonant lengthening) were not altered by changes in segmentation procedure.

An estimate of measurement error was arrived at by segmenting spectrograms of two of the experimental utterances four separate times. A standard deviation was computed on the measurements of each of the 22 segments in the two nonsense words. An average standard deviation over the 22 segments was then taken. This average was approximately 3 msec which agrees well with the estimates of measurement error arrived at by Shockey et al. (5 msec) and Lehiste (between 2 and 3 msec).7

B. Results

The most striking effect in the data from Experiment 1 is the lengthening of the vowel in the final syllable of utterances. If final-syllable stressed vowels are compared with other stressed vowels, and final-syllable unstressed vowels are compared with other unstressed vowels, it is determined that vowels in final syllables are consistently longer than vowels in other syllables. Over all seven subjects, the average increment (i.e. durational difference between final-syllable vowel and non-final-syllable vowel) is approximately 100 msec for both stressed and unstressed vowels. Five subjects show greater absolute increments for unstressed vowels than
for stressed vowels. Two subjects show greater increments for stressed vowels. The average increment ranges from about 70 msec for two of the subjects to about 160 msec for one subject.

In Fig. 1(a), stressed vowel duration is plotted against number of syllables in words. The data are averaged over the seven subjects. The leftmost bar in each cluster of bars represents the average duration of the first vowel in words of the specified length, and the rightmost bar represents the average duration of the last vowel in the words of the specified length. The last vowel, the rightmost bar in each cluster, is invariably longer than any other bar in the cluster. The same final-syllable lengthening trend is evident in Fig. 1(b), where unstressed vowel durations are plotted. The same trend is evident for each of the seven subjects individually. In no case do the histograms for the individual subjects fail to show final-syllable lengthening.8

In addition to final-syllable vowel lengthening, three consonantal duration increments are found in the data. The two final-syllable consonants, i.e., the absolute final and the penultimate syllables, are both lengthened, and in addition the word-initial consonant is lengthened. Figure 1(d) plots duration of unstressed consonants against number of syllables in words. All consonants not immediately preceding stressed vowels are, in this study, considered to be unstressed consonants. The leftmost bar in each cluster represents the average duration of word-initial consonants, while the next-to-rightmost and rightmost bars in each cluster represent penultimate and absolute final consonant durations. Cross-hatched bars represent durations of medial consonants. Medial consonants are presumably unlengthened. As seen in Fig. 1(d), absolute final unstressed consonant durations are considerably greater than medial un

Consonant lengthening is of lesser absolute magnitude than final-syllable vowel lengthening. For unstressed consonants, the average increments are approximately 20 msec for absolute final consonants, 15 msec for penultimate consonants, and 30 msec for word-initial consonants. For stressed consonants the average increments are virtually zero for penultimate consonants and 20 msec for word-initial consonants.

Conclusions: The present data, in accordance with Lindblom’s, suggest that final-syllable lengthening consists mainly of increments in vowel duration, although consonantal duration increments also occur in final syllables.12 Lindblom does not, however, report word-initial increments.13 In the present data, the trend toward word-initial consonant lengthening is even stronger than the trends toward final-syllable consonant lengthening.

II. EXPERIMENT 2

In the first experiment, only utterances of the form “say a [bab],” “say a [babb],” “say a [babab],” etc., were tested. Such utterances have the intonation pattern of imperative sentences. To the author’s knowledge, no previous study has presented quantitative data specifically concerned with utterances having interrogative intonation. Furthermore, the first experiment did not present data concerning the most common sentence intonation, viz., the declarative intonation. In Experiment 2, the problem is to determine whether or not the effects found in the earlier study occur in utterances with interrogative and declarative intonation as well as in utterances with imperative intonation.

A. Methods

A native English speaker (hereafter referred to as S1) who, as all other subjects in the experiments, was unaware of the experimental hypotheses, read a randomized ordering of utterances of the form “was it a [bab]?” “was it a [babab]?” “was it a [babab]?” etc. The nonsense words are the same as those listed in Table I. All subsequent inventories in the present study are also of the same form, except that the carrier phrase or the structure of the nonsense syllables (though not their stress patterns nor their number of syllables) may vary.

As in all the experiments, S1 was asked to read each utterance as if it were a normal sentence. He was asked to pronounce each [b] with a clear plosive burst, and he was asked to release all final [b]’s. At least four samples were obtained of each utterance.

In the second part of Experiment 2, S1 and another native speaker of English (S2) read a randomized ordering of declarative utterances of the form “it was a [bab],” “it was a [babab],” “it was a [babab],” etc. The subjects were given the same instructions as in the case with interrogative utterances. Again, four samples of each utterance were obtained for each subject.
Wide-band spectrograms were made for each utterance, and segmentation was conducted in the same manner as in the earlier experiment, i.e., spectrographic evidence of lip closure and lip opening was used to determine the duration of consonants and vowels.

### B. Results

Both interrogative and declarative utterances show final-syllable vowel lengthening. Figure 2 presents six histograms in which vowel duration is plotted against number of syllables in words. For both stressed and unstressed data, the duration of final-syllable vowels, represented by the rightmost bar in each cluster, is invariably greater than that of any other vowel in words of the specified length. For the interrogative utterances ($S_1$), the increments (the average length of final-syllable vowels minus the average length of non-final-syllable vowels) are around 100 msec for stressed vowels and around 130 msec for unstressed vowels. For the declarative utterances the increments are not so great; for $S_1$, 90 msec for stressed vowels and 80 msec for unstressed; for $S_2$, 70 msec for stressed vowels and 50 msec for unstressed. It was thought that the difference between $S_1$'s increments on interrogative and declarative utterances might have been statistically significant. A $t$-test of unstressed data only, revealed significant differences between increments for interrogative and declarative utterances ($t=3.78$, df=3, $p<0.05$). But the $t$-test for stressed data only revealed no significant differences ($t<1$).

Figure 3 presents the data concerning consonant durations. For each of the six histograms, duration of consonants is plotted against number of syllables in words. Since the first experiment revealed lengthening trends for initial, penultimate, and final consonants, it seemed reasonable again to display the histograms in a way that differentiated between lengthened and unlengthened (medial) consonants. In Fig. 3, therefore, all bars representing durations of medial consonants are cross hatched.

The trend toward initial-consonant lengthening, which was reported in the earlier experiment, is clearly evident in data for both declarative ($t=7.2$, df=5, $p<0.001$) and interrogative ($t=7.1$, df=5, $p<0.001$) sentence intonations, and for both stressed and unstressed consonants within each type of intonation. Considering unstressed data, it can be seen that the trends toward penultimate-consonant lengthening (interrogative: $t=5$, df=2, $p<0.05$; declarative: $t=7.5$, df=2, $p<0.05$) are also in accord with the results of the earlier experiment. For $S_1$ the data on (unstressed) final consonants also reveals the lengthening trends (interrogative: $t=10.5$, df=2, $p<0.05$; declarative: $t=18.7$, df=2, $p<0.01$) found in the earlier study. $S_2$'s (unstressed) final consonants do not, however, show the lengthening trend. The earlier study revealed only a weak trend toward penultimate stressed-consonant lengthening, and the Fig. 3 data ($t<1$ for both interrogative and declarative data) are not at all inconsistent with those of the earlier study.
Table II displays the magnitudes of the average increments in consonant duration due to position-in-utterance. The data shown there are in basic accord with those of the earlier study except in that $S_2$ shows no absolute final consonant increments.

Conclusions: It is clear that changes in sentence intonational type do not eliminate final-syllable vowel lengthening. It is possible, in fact, that final-syllable vowels are lengthened even more in interrogative sentences than in declarative and imperative sentences. Initial-consonant lengthening appears to be essentially unaffected by changes in sentence type. Final-syllable consonant increments are probably not affected by changes in sentence type. The fact that one subject failed to show lengthening of absolute final consonants may be due to random variation of the data.

III. EXPERIMENT 3

Very little evidence has been presented to substantiate the claim that final-syllable duration increments occur in phrase-final as well as in utterance-final position. Furthermore, practically no quantitative data have been presented in phonetics literature to support Lindblom’s little-known claim that final-syllable duration increments occur in word-final but non-phrase-final position (i.e., that durational increments occur on words whether or not these words are in phrase-final position).

A. Methods

Three inventories were employed: (1) the declarative inventory described in Experiment 2: “It was a [bab],” etc.; (2) an inventory intended to test the claim concerning phrase-final lengthening: “The [bab] is on the table,” etc.; and (3) an inventory designed to test the claim concerning word-final lengthening: “The [bab] apple is on the table,” etc.

The subjects were $S_1$ and $S_2$. They read the 15 utterances of each inventory under the same conditions in all three cases. The subjects read the inventories separately and in order, i.e., Inventory 1 first, etc. For $S_1$ a second recording session was needed to obtain new samples of some of the utterances in Inventories 2 and 3. (During $S_1$’s first recording session, he had produced a considerable number of fricativized [b]’s, and, consequently, some of the recordings were invalid given the segmentation criterion employed in the study.)

The subjects were instructed to read the utterances of Inventories 2 and 3 without pause between the phrases and words of each utterance. It was particularly important that there be no pause between the nonsense word and the words following it. Had a pause occurred after the crucial nonsense word, it might have been argued that the test was still dealing with utterance-though not sentence-) final position.

B. Results

Figure 4 displays data on duration of vowels for Inventories 2 and 3. Figure 2, of course, includes the data from Inventory 1. Stressed and unstressed final-syllable vowels show considerable increments. For unstressed vowels, $S_1$’s average durational increment in phrase-final position (Inventory 2) is approximately 92 msec, and in word-final position the comparable value is 96 msec. For $S_2$ unstressed vowels are lengthened by about 40 and 48 msec for Inventories 2 and 3, respectively. $S_1$’s average increments for stressed vowels are about 52 and 48 msec; for Inventories 2 and 3, and $S_2$’s average increments for stressed vowels are about 40 and
The bab is on the table.

Fro. 5. Consonant durations for inventories with the nonsense word in phrase-final and non-phrase-final positions.

64 msec for the same two inventories. The corpus of data is probably too small to attempt to answer the question of whether or not final-syllables are lengthened more in word-final, phrase-final, or utterance-final position. In addition, the order of reading of the inventories constitutes an important variable that was not controlled for. Inventory 1 (utterance-final inventory) was read first, and since subjects apparently produced the slowest rate-of-utterance at the beginning of the session, they probably spoke slower on Inventory 1 than on the other two. The apparent difference between lengthening in word- and phrase-final positions as opposed to utterance-final position may then be artifactual. A t-test of difference between means revealed a trend approaching statistical significance ($t=1.92$, $df=15$, $p<0.10$) for the difference between vowel lengthening in utterance-final and word-final positions.

Figure 5 displays the data concerning consonant duration increments due to position-in-phrase and position-in-word. A trend toward initial-consonant lengthening is clearly present ($t=9.35$, $df=5$, $p<0.001$). It should be noted that initial consonants in Experiment 3 are not the initial consonants of phrases, although they are the initial consonants of words. The same was true of initial consonant data from Experiment 1. In all cases, initial consonants have been word- (but not phrase-) initial. It is not, therefore, surprising that the initial-consonant-lengthening trend continues to appear in Experiment 3.

Concerning the data for penultimate-consonant lengthening, $S_2$'s data ($t=10.9$, $df=5$, $p<0.001$) are in conformity with those of subjects from the first experiment, as well as being in conformity with the data presented in Fig. 6. $S_2$'s data are not drastically different, but his unstressed penultimate consonants do not show lengthening as conclusively ($t=1.1$, $df=2$, $p>0.10$) as was typical of the data of subjects from Experiment 1.

The data on final-consonant lengthening does appear to differ from data in Fig. 3 and from data from Experiment 1. Although the corpus of data is small, it can be seen that $S_1$'s data for phrase-final and word-final positions shows smaller increments ($t=1.75$, $df=2$, $p<0.10$) than his data for utterance-final position as exemplified in Fig. 3. $S_2$ fails to show final-consonant lengthening in Fig. 3 (Inventory 1), and in that respect he differs from subjects in Experiment 1. In any case, $S_2$ also shows small (by comparison with increments found in Experiment 1) final-consonant increments in phrase-final and word-final positions.

The average magnitudes of the consonantal increments are displayed in Table III. Consonantal increments for Inventory 1 are represented in the "declarative" rows in Table II. The average final-consonant increment for the seven subjects of the original experiment was 23 msec and this figure does not appear to differ significantly from $S_1$'s average for final-consonant lengthening in utterance-final position (Table II). However, no subject from the earlier experiment showed final-consonant increments as small as 4 msec, which appears to be the typical increment for consonants in phrase-final and word-final positions (Table III).

Conclusions: Clearly final-syllable vowel increments do occur in both phrase-final and word-final positions. The corpus of data is too small to answer the question of whether or not increments are of lesser magnitude for Inventories 2 and 3 than for Inventory 1. Initial and

Fig. 5. Consonant durations for inventories with the nonsense word in phrase-final and non-phrase-final positions.
penultimate consonant data from Inventories 2 and 3 reveal lengthening trends not unlike those from Inventory 1. It may be, however, that final consonants are not lengthened in phrase-final or word-final positions, or at least it may be that final consonants are not lengthened as much in phrase-final and word-final positions as they are in utterance-final position.

IV. EXPERIMENTS 4-8

Thus far, the data on position-in-utterance lengthening have been carefully quantified only for inventories in which syllables have been composed of [b]’s and [a]’s. All final syllables have been closed (CVC as opposed to CV). No consonant clusters and no complex vowels (diphthongs) have been employed. The problem in Experiment 4 is to examine position-in-utterance effects in inventories with varied syllable structures.

A. Methods

Experiment 4: A new subject (S1) read an inventory of the form “it was a [baib],” “it was a [baibiaib],” “it was a [baibibib],” etc. He was given the same instructions as in the [bab] experiments. Admittedly, the inventory is quite artificial in the [baib] case; even so the subject had little difficulty in performing the task even with the longest words.

Experiment 5: S2 read an inventory of the form “it was a [sas],” etc. The important difference between the instructions for this inventory and the instructions for other inventories is that in this case there was no release (of final consonant) requirement. The segmentation procedure for this inventory was based on spectrographic evidence of contact between the tongue tip and the alveolar ridge as well as evidence of frication. For absolute final [s]’s, segmentation was a problem, since there was no straightforward point at which frication terminated. A conservative segmentation criterion was adopted for final [s], i.e., if the experimenter was unsure which of two points should be considered the termination of consonant frication, the point which resulted in the smaller value for the duration of the consonant was chosen. It should be clear that values obtained for durations of final [s]’s are not strictly comparable with values obtained for final plosives where the final cutoff is determined by spectrographic evidence of plosive release. It is not clear that release ever occurs for final fricatives.

Experiment 6: S3 read an inventory of the form “it was a [stast],” etc. No release instruction was necessary in this case since final fricatives plus plosive clusters are normally released in English. The segmentation criterion was essentially the same as in Experiment 5 although some special problems were presented by the [st] clusters. The silent interval after frication of [s] and before vowel phonation was considered to be the [t].

Experiment 7: A fourth subject read an inventory of the form “say a [ba],” “say a [bab],” “say a [baba],” etc. The subject was instructed to pronounce all [b]’s as plosives rather than as fricatives. Here more than in any of the other experiments, the segmentation problem was great. Spectrograms showed final vowels which consisted of a period of phonation followed by a period of h-like friction. It was necessary to make a somewhat arbitrary decision about the point of termination of the vowel. It was decided that the most conservative segmentation (i.e., the segmentation resulting in the shortest duration for final vowels) should be used; therefore, the point of termination of phonation was considered to be the termination of the vowel. It should be clear that results obtained for the [ba] inventory are not strictly comparable with those obtained for inventories with final consonants. In the former case, the presumed mechanism of termination of the vowel is simply cessation of phonation, while in the latter case, the presumed mechanism is lip closure.

Experiment 8: S4 read an inventory of the form, “it was a [pap],” etc. No release requirement was needed since final voiceless stops are normally released in American English. Two segmentation procedures were employed for this experiment: one in which the period of aspiration following the release of [p] was considered to belong to the consonant, and one in which it was considered to belong to the vowel. Only one reading of the inventory was considered for this particular experiment.

B. Results

Results of Experiment 4 demonstrate that final-syllable vowel increments occur when the vowel is a diphthong. Figure 6(a) displays durational data for stressed [ai], and Fig. 6(b) displays unstressed [ai] data. The increment in final-syllable position can be seen to be entirely consistent. The average stressed vowel increment is 112 msec, but the average unstressed vowel increment is only 76 msec.

<table>
<thead>
<tr>
<th>Table III. Average magnitude of consonantal increments.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial (msec)</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>S1</td>
</tr>
<tr>
<td>stressed</td>
</tr>
<tr>
<td>unstressed</td>
</tr>
<tr>
<td>Word-final</td>
</tr>
<tr>
<td>unstressed</td>
</tr>
<tr>
<td>S2</td>
</tr>
<tr>
<td>stressed</td>
</tr>
<tr>
<td>unstressed</td>
</tr>
<tr>
<td>Word-final</td>
</tr>
<tr>
<td>unstressed</td>
</tr>
</tbody>
</table>
The consonantal data for Experiment 4 are displayed in Figs. 6(c) (stressed data) and 6(d) (unstressed data). Initial ($t=15.5$, $df=5$, $p<0.001$) and final ($t=5.9$, $df=2$, $p<0.01$) consonant trends agree well with values found earlier in experiments with the [bab] inventory. However, no lengthening trend is apparent for penultimate consonants in either stressed or unstressed conditions. In fact, a small, statistically insignificant trend toward penultimate consonant shortening ($t=1.17$, $df=5$, $p<0.20$) can be seen. Table IV(a) presents the magnitudes of the consonantal increments.

Results of Experiment 5 demonstrate that vowel increments occur in final syllables when the consonant is [s]. Figures 7(a) (stressed vowels) and 7(b) (unstressed vowels) show the data. Stressed vowel increments average 104 msec, while unstressed vowel increments average 80 msec.

Consonants also show lengthening trends in Experiment 5. Figures 7(c) and 7(d) show that both initial ($t=4.4$, $df=5$, $p<0.001$) and penultimate ($t=8$, $df=5$, $p<0.001$) stressed and unstressed [s]'s are lengthened. Figure 7(d) reveals the largest consonantal increment that has been found in any of the experiments by the present author. Utterance-final [s] appears to be lengthened ($t=64$, $df=2$, $p<0.001$) by even greater amounts than the vowels of final syllables. Table IV(b) displays the data on magnitude of increments.

Results of Experiment 6 demonstrate that final-syllable vowels are lengthened even when the syllable structure involves the consonant cluster [st]. The data are presented in Figs. 8(a) and 8(b). The averaged stressed vowel increment is 96 msec. The average unstressed vowel increment is 88 msec.

Consonantal durations are displayed in two ways for the [stast] experiment: (1) Figures 8(c) and 8(d) present the durational values for the [st] clusters, i.e., the values for a s-duration plus t-duration; (2) Figs. 8(e)-8(h) present the average durations for [s]'s and [t]'s separately. Durational values for the [s]'s only are presented in Fig. 8(e) (stressed [s]'s) and in Fig. 8(f) (unstressed [s]'s). Average durations for [t]'s only are presented in Figs. 8(g) (stressed) and 8(h) (unstressed).

![Fig. 7. Segment duration for the [sas] inventory.](image)

**Table IV.** Average magnitude of consonantal increments in various kinds of syllables.

<table>
<thead>
<tr>
<th>Type</th>
<th>Initial (msec)</th>
<th>Penultimate (msec)</th>
<th>Final (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) [bsb]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>16</td>
<td>-4</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>24</td>
<td>-4</td>
<td>16</td>
</tr>
<tr>
<td>(b) [sas]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>8</td>
<td>16</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>20</td>
<td>16</td>
<td>136</td>
</tr>
<tr>
<td>(c) [stast]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>24</td>
<td>8</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>32</td>
<td>16</td>
<td>48</td>
</tr>
<tr>
<td>(d) [bas]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>32</td>
<td>16</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>32</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>(e) [pap] (Segmentation 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>20</td>
<td>12</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>32</td>
<td>8</td>
<td>...</td>
</tr>
<tr>
<td>(Segmentation 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stressed</td>
<td>24</td>
<td>8</td>
<td>...</td>
</tr>
<tr>
<td>Unstressed</td>
<td>40</td>
<td>24</td>
<td>0</td>
</tr>
</tbody>
</table>

![Fig. 8. Segment duration for the [stast] inventory.](image)
Figures 8(c) and 8(d) show that both stressed and unstressed \text{[st]} clusters are lengthened in initial, penultimate, and final position. Furthermore, Figs. 8(e) and 8(f) show that for both stressed and unstressed clusters, \text{[s]}’s are lengthened in initial, penultimate, and final positions. In addition, from Figs. 8(g) and 8(h), it can be seen that \text{[t]}’s are lengthened in initial and final positions, although no such trend is apparent in the case of penultimate \text{[t]}’s.

Results of Experiment 7 suggest that, even given the conservative segmentation criterion outlined above, final-syllable vowels are lengthened in open syllables. Figures 9(a) and 9(b) display the data. The final-syllable vowel increments here are of a somewhat lesser magnitude than those found in earlier experiments. For stressed vowels, 48 msec is the average increment, and for unstressed vowels the average increment is 32 msec. Vowel lengthening here is significant past the 0.01 level for both stressed (t= 7, df= 3, p<0.01) and unstressed (t= 10.7, df= 3, p<0.002) vowels.

Consonantal duration data for Experiment 7 are displayed in Figs. 9(c) and 9(d). Both initial and penultimate (in this case the “penultimate” consonant is actually the last consonant in the word) consonant lengthening are apparent. The magnitudes of the increments are shown in Table IV (d). The consonantal increments are significant past the 0.01 level for both initial (t= 17.2, df= 5, p<0.001) and penultimate (t= 4.5, df= 5, p<0.01) consonants.

The results of Experiment 8 show that even with an inventory of nonsense words composed of vowels and unvoiced stops, final-syllable and initial-consonant lengthening continue to occur. Under both Segmentation 1 (aspiration included with vowel duration) and Segmentation 2 (aspiration not included with vowel duration), stressed vowels in final syllables are lengthened by 40 msec and unstressed vowels in final syllables are also lengthened by 40 msec. These data are presented in Figs. 10(a) and 10(b).

In Figs. 10(c) and 10(d), it can be seen that initial (Segmentation 1, t= 7.75, df= 5, p<0.001; Seg. 2, t= 4.8, df= 5, p<0.01) and penultimate (Seg. 1, t= 1.74, p>0.10; Seg. 2, t= 2.02, df= 5, p<0.10) consonant lengthening for voiceless stops differs little from comparable results for the other inventories. The fact that there is a very weak trend toward final-consonant shortening under Segmentation 2 (t= 1.8, df= 2, p>0.10) is not really surprising since the aspirated period following release of occlusion for absolute final \text{[p]}’s was not included in the duration of the segments, while the period of aspiration was included for nonfinal consonants. A fairer comparison of final to medial durations is that recorded for Segmentation 1 (though the lengthening trend is nonsignificant, t<1), where neither final nor nonfinal \text{[p]} durations included aspiration. The magnitudes of lengthening are included in Table IV (d).

**Conclusions:** It seems clear that final-syllable and initial-consonant durational increments are not limited to syllables in which the vowels are monophthongs nor to syllables in which the consonants are stops. In fact, absolute final \text{[s]} shows far greater increments than final \text{[b]}. In addition, all the previously found increments appear when the syllables include the consonant cluster \text{[st]}. Furthermore, it appears that the durational increments that have been found in earlier experiments are
not limited to closed (CVC) final syllables. There is, however, a weak indication (which should be tempered by knowledge of segmentation problems) that open syllables are not lengthened as much as are closed syllables. Finally, it appears that final syllables and initial consonants are lengthened in much the same way for syllables composed of voiceless stops and vowels as they are for syllables composed of voiced stops and vowels.

V. DISCUSSION

The consistency of the final-syllable lengthening effect both in this study and in others is impressive. The present data would suggest that the bulk of final-syllable increments is generally carried by vowels, although an altered inventory (including short vowels, for instance) might reveal other trends. In the present study, as in Lindblom’s, final-syllable consonant increments are consistently found. However, Lindblom did not report word-initial consonant lengthening (see Footnote 13). In the present data the initial consonant effect is even more conclusively demonstrated than most final-syllable consonant effects.

The results of the experiments also indicate that final-syllable vowels are lengthened under various intonational patterns. There is a weak indication that when the intonation on the final syllable is rising (as in the interrogative utterances), the final-syllable increment is of greater magnitude than in the case where there is falling intonation on the final syllable. Such a trend agrees with the results of Spang-Thomsen’s study of French, but apparently disagrees with Benguerel’s results in a study of French.

One of the most important results of the present study is the demonstration that final-syllable lengthening can occur in word-final position. The trend toward lesser magnitude of lengthening in word-final and phrase-final positions is consonant with the results of Joos’ work as well as Mattingly’s. But the amount of difference found in the present experiments between lengthening in absolute final and sentence internal positions does not correspond well with the predictions of either Joos or Mattingly. Joos’ results would suggest that there should be four times as much lengthening for absolute final syllables in declarative utterances as for final syllables of words in sentence internal “plus juncture” positions. Mattingly suggests that syllables just before word boundaries are lengthened by 25% and that syllables just before pauses are lengthened by 50%. The present results would suggest much higher lengthening percentages for syllables preceding word boundaries, and less difference between final-syllable lengthening in sentence-internal positions and absolute-final position.

Barnwell suggests that Mattingly’s synthesis rules provide no lengthening allowances for syllables in final position if the syllables end in voiceless stops. The results of Experiment 8 suggest strongly that syllables terminating in voiceless stops are lengthened in final position. The amount of lengthening appears to be only slightly less than the amount for syllables ending in voiced stops.

With respect to the consonant[s], utterance-final lengthening appears to be a clearcut result of the present study. The increments in final position appear to be of even greater magnitude than final-syllable vowel increments. MacNeilage has found similar results for absolute final [f], and Malcot has found similar results for other voiceless fricatives. Perhaps all fricatives show large increments in absolute final position.

It should be pointed out that fricatives in absolute final position differ from stops in final position in their manner of termination—this is so, at least in the context of the present experiments. Subjects in the experiments have been instructed to release final [b]s, and consequently it has been possible to specify a discrete point of termination of the final consonant; the point of termination corresponds to the last lip opening in the utterance. Nonfinal [b]s are also terminated by lip opening in the [bab] experiments. There was no release instruction in the [sas] experiment, since normal English speech does not include release of final [s]. Final [s] appears to be terminated by a shutting off of the airstream through the vocal cavity. But nonfinal [s]s in Experiment 5 were terminated in a quite different way—viz., by lowering of the tongue and onset of voicing for the following vowel. It would probably be an error to attempt to explain the great magnitude of absolute final fricative lengthening in terms of the theories of lengthening proposed below. A quite distinct explanation appears necessary in this case, and that explanation should be formulated in terms of the differential mechanisms of termination for absolute final fricatives as opposed to nonfinal fricatives.

A similar line of argument might be appropriate with respect to absolute final vowels. The termination of such vowels is accomplished in a different way from that employed in the case of final closed syllables. In the latter case, the articulation of the final consonant effects the termination of the vowel, while in the former case cessation of voicing terminates it.

Even so, it is important to keep in mind that there are important experimental problems in comparing absolute final vowels and absolute final fricative consonants with segments which are not in absolute final position. The different positions-in-utterance require different manners of articulation (i.e., termination). It will be easier to draw conclusions as to whether or not there are special lengthening mechanisms (not implied by the theories proposed below) for final vowels and fricatives after it has been determined whether or not the same kinds of increments occur for such segments in word-final (or phrase-final) position (not preceding a pause) as in utterance-final position.
The significance of durational variation of segments in different positions in utterances may be considerable. In the first place, most studies of segment duration, even if they are not directly concerned with position-in-utterance effects, should take such effects into consideration. For example, any study concerned with the question of whether or not speaking rate changes with word length should take final-syllable lengthening into account. Suppose that the duration of one-syllable words is compared with that of two-syllable words. It would appear that speaking rate increases in pronouncing two-syllable words, i.e., it would appear that the average duration of a syllable decreases for two-syllable words, but in fact it may only be the case that one-syllable words have the durational increments due to final-syllable lengthening and primary stress throughout all their segments, while only half the syllables have such effects in two-syllable words. Any account of average duration of syllables that fails to consider final-syllable lengthening will of necessity be incomplete.

In addition, studies of the physical correlates of stress should take position-in-utterance effects into consideration. Data from the present study as well as from others indicate that final unstressed syllables may be as long or longer than nonfinal stressed syllables. Comparisons between final unstressed syllables and nonfinal stressed syllables would suggest that duration is not a correlate of stress at all. Obviously, studies of the physical correlates of stress should compare the duration of final syllables only with that of other final syllables.

What explanations can be given for position-in-utterance effects? One suggestion is that lengthening as a function of position-in-utterance results in an acoustic cue which assists the listener in perception. Another suggestion is that some aspect of the articulatory control process causes durational increments in certain positions in utterances.

Haden suggests a perception-based explanation when he describes final-syllable lengthening as a "juncture cue." Listeners can be thought to locate the boundaries of linguistic units on the basis of durational differences. The demonstration in the present study that final-syllable increments occur at the end of phrases and words as well as sentences would appear to lend credence to the "boundary cue theory" of final-syllable lengthening.

One of the advantages of the boundary cue theory of final-syllable lengthening is that it can account for word-initial consonant lengthening. Since word-initial consonants and final syllables are contiguous, word-initial consonant lengthening can simply be thought of as providing an additional cue to the listener.

One of the possible production-based explanations for final-syllable lengthening is suggested by Lindblom. He argues that the effect is due to final-syllable intensity drops which are characteristic of Swedish (and English) declarative sentences. He refers to Öhman's model of word and sentence intonation in which there is a "general relaxation of speech gestures toward the end of utterances." Making the assumption that the same amount of physiological energy is expended for each syllable (excluding stress variations), Lindblom concludes that the lower intensity per unit time of final syllables is compensated for by increased duration of final syllables.

There are several reasons to doubt the Lindblom proposal. First of all, the present author is unaware of compelling evidence indicating that total speech energy remains constant on syllables of the same degree of stress. In fact, the problem of giving even a minimal characterization of the physical correlates of stress has been notoriously difficult.

A second reason for doubt is that Lindblom's proposal is inconsistent with data on amplitude of syllables occurring toward the end of utterances. Preliminary investigation of spectrograms (with amplitude displays) from the present study does indicate a drop in amplitude toward the end of utterances. But this amplitude drop may occur in two, three, or four syllables at the end of the utterance, while the durational increment applies only to the final syllable.

The most convincing reason to doubt the Lindblom theory comes from data on position-in-utterance durational effects in questions. Elsewhere I have presented evidence that although there is no amplitude drop (rather an increment) on the final-syllable vowels of many interrogative utterances, still final-syllable lengthening occurs.

Another possible speech production-based explanation of position-in-utterance effects is based on hypothesized characteristics of the motor command system. Some part of the process of pre-articulatory planning might be concerned with units stretching across words, phrases, and sentences. This planning might, for instance, determine the nature of prosodic patterns. The "planning" theory explanation of final-syllable lengthening suggests that the prearticulatory planning mechanism deals with segmental inputs in two stages. During stage one, linguistic units which may stretch across several syllables (for instance, prosodic patterns for words, phrases, and sentences) are planned. During stage two, each syllable is planned individually. It is assumed that before processing of a particular syllable (stage-two planning) can begin, all stage-one planning relevant to that syllable must have been completed. It is assumed further that processing for the two stages cannot be conducted simultaneously. Stage-two processing must shut off temporarily while stage-one processing is performed. Each time stage-two planning of a word-final syllable is performed, a durational increment is applied to that syllable in order to provide time for the stage-one planning relevant to the syllables of the next word. Planning theory assumes that the duration of word-final syllables is increased whether or not a "next word" actually exists. Thus, word-final syllables are lengthened even in utterance-final position.
As formulated, planning theory implies that final-syllable lengthening must be a language universal. It is not yet clear whether or not this implication is valid. Whether or not either of the proposed speech production theories of lengthening is valid, final-syllable and word-initial consonant lengthening may serve as powerful cues to the listener for the location of the boundaries of words, phrases, or sentences. It may be the case that one of the speech production explanations is correct, and that only incidentally do listeners make use of the effects in the process of speech comprehension.

ACKNOWLEDGMENTS

This study was conducted in the Linguistics Laboratory, Department of Linguistics of the University of Texas at Austin and was supported in part by NSF-USDP Grant GU-1598. The author would like to express his gratitude to a number of professors and students at the University of Texas who commented on early versions of this paper. A special thank you goes to Dr. Peter MacNeilage for his many helpful suggestions and for his encouragement. I would also like to thank the Child Development and Mental Retardation Center at the University of Washington for providing the final typing of this paper.

*Present address: Child Development and Mental Retardation Center and Department of Speech, University of Washington, Seattle, Washington 98195.

The term "lengthening" is used here in a descriptive sense only. N. Sobin and F. Brandon ["A No-Nonsense Approach to Position-in-Utterance Vowel Length," unpublished paper, Department of Linguistics Univ. of Texas (1971)] quite correctly point out that the problem of selecting "an optimal vowel length" (p. 4) has not yet been solved. Thus, we might just as correctly speak of non-final-vowel "shortening." Throughout the present study, the term "lengthening" is not meant to imply that short vowels are the basic forms nor that long final vowels are somehow derived forms.


Ref. 4.

The spectrographic data indicated that when subjects produced clear plosive bursts, there was an abrupt closure of the lips from 30–100 msec before the abrupt opening of the lips.


In the present paper as well as in Lindblom's, only primary-stressed syllables are called "stressed syllables." All syllables of other degrees of stress are referred to as "unstressed.


This does not refer to the penultimate syllable. The penultimate consonant is the next-to-the-last consonant in any utterance of the inventory. Thus it is the first consonant in every final syllable of the utterances.

If a different segmentation criterion had been employed (as for instance the criterion of Delattre, Ref. 3) or if a different segmental inventory had been employed (for instance, an inventory with short vowels), the results indicating that the bulk of increments is carried by vowels might have been different.

Word-final consonant lengthening is reported in A. Malécot, "The Force of Articulation of American Stops and Fricatives as a Function of Position," Phonetica 18, 95 (1968). But Malécot compares final consonant durations with penultimate (instead of medial) consonant durations. As a result, his data do not bear on "final-consonant lengthening" as the term is used in the present study.

Word-initial consonant lengthening is reported in I. Lehiste, "A Phonetic Study of Internal Open Juncture," Supplement to Phonetica 5 (1960); and in J. E. Hoard, "Juncture and Syllable Structure in English," Phonetica 15 (2), 96 (1966). However, the phenomenon reported in Lehiste's paper and Hoard's paper is of a different sort from that reported in the present study. Lehiste and Hoard compared initial consonant durations with final consonant durations, while I have compared initial consonant durations with medial consonant durations.

In the "Results" sections for Experiments 2–8, tests of statistical significance are generally reported only for consonant lengthening. In cases where no statistical tests are reported, the results are clearly significant.

Descriptive (structural) linguists generally use the term "phrase" in such a way that it is roughly interchangeable with "breath group." The term as used here refers to a more abstract structure which may or may not in any particular instance correspond with "breath group." Since in the present experiments all data where subjects paused following a "phrase boundary" were thrown out, we can safely assert that the term "breath group" does not apply to the sentence-internal syntactic boundaries of the present study's inventories.

Lindblom, Ref. 4. The only quantitative data I know of are presented in such a way that final-syllable durations include the duration of pauses following them.

A significance level cannot be reported here since, by chance, the differences for all the comparisons for the paired data t-test were equal. As a result, t = 0 even though the differences are very small.


For a review of articles which have failed to extract position-in-utterance effects from word-length results, see Chap. X of my dissertation (Ref. 9).

See Lindblom, Ref. 4, and Delattre, Ref. 3.
SPEECH SEGMENT DURATION IN ENGLISH

27 Lindblom, Ref. 4, p. 4.
29 The notion of prearticulatory planning is not uncommon. It is particularly important in Peter MacNeilage's, "The Motor Control of the Serial Ordering of Speech," Psychol. Ref. 77 (3), 182 (1970).
30 For a review of literature pertinent to this question, see my dissertation (Ref. 9, Chap. I).