Segment Duration, Voicing and the Syllable

Leigh Lisker*

ABSTRACT

The syllable has been proposed as the basic unit of articulatory organization. One kind of evidence cited is the well-known relation between vowel duration and the duration of a following stop; the relatively long voiceless stop is preceded by a shorter vowel than is the shorter voiced stop. Thus the combined durations of vowel and stop "tend" to be equal. One problem with the syllable is that of deciding where to locate its boundary in utterances reported to consist of more than one syllable. The decision criteria most generally used are linguistic (phonotactic) rather than phonetic in nature. Durational data derived from nonsense words of $C_1aC_2i$ type were subjected to the variance test to determine whether durational relations pointed to a syllabification $C_1a\cdot C_2i$ or $C_1aC_2i$. Aside from the known defects of the variance test, it failed to point consistently to either solution for the different "words" measured. Moreover, the same test failed to yield consistent results when applied to monosyllabic utterances of $C_1aC_2$.

THE SYLLABLE AS AN INTUITIVE GIVEN

The development of writing has reached full flower in an alphabetic system exemplified by a phonetic transcription in which each so-called speech sound is represented by a unique letter shape. This kind of system has, it is felt, a closer relation in general to the facts of speech perception and production than do other kinds of writing. It is true that representing speech by a sequence of letters, and thus by implication representing it as a sequence of distinct sounds, quite obscures another property of speech signals, namely their quasicontinuous nature. This aspect of speech, however, only becomes apparent when we look at it with the help of laboratory instruments; the ear that listens to speech is still convinced of its discrete (or at least quasidiscrete) nature. (One may wonder whether the lip-reader sees speech as a sequence of discrete elements corresponding to the phonetic segments.) Of course, visible patterns that reflect more faithfully the quasicontinuous nature of speech signals, for example waveforms and spectrograms, seem to be harder to read than any writing system known, and it is possibly true that only a discrete representation can ensure the tolerably

*This paper was given at a Symposium on Segment Organization and the Syllable, Boulder, Colorado, 21–23 October 1977.

†Also University of Pennsylvania.

[HASKINS LABORATORIES: Status Report on Speech Research SR-54 (1978)]
efficient transmission of linguistic messages by graphical means. What is somewhat harder to understand is why alphabetic systems fail to incorporate, as a regular feature, any device for indicating the so-called syllabic organization of speech, although the listener is as strongly convinced (we think) that speech is composed of syllables as he is that it is decomposable into sounds. Linguists moreover, particularly phoneticians, show no reluctance about invoking the syllable, and yet the transcriptions they use provide no mark with which to indicate syllable boundaries. If, for example, we are inclined to feel that a word pair like plight-polite differ in syllable structure at least as much as in sound composition (and we might even believe that they differ not at all in the latter), nevertheless our inclination is to diagnose the difference as one of segment composition.

DEFINING THE SYLLABLE: COUNTING VS. DELIMITING

At the same time that the syllable is regarded as a pervasive feature of speech, it is notoriously difficult to make explicit the notion of the syllable or of syllable organization. It has something to do, it seems, with the near-dichotomous classification of the speech sounds into vowels and consonants, and the strong conviction that there exist special relations between a vowel and abutting consonants. These relations are of two kinds at least: in one the vowel and consonant or consonants are intimately connected, and in the other they are products of two distinct articulatory gestures. Speech is then decomposable into subsequences of phonetic elements, each consisting of a vowel and those adjacent consonants having the closer connection to it. Aside from the fact that the separation into vowels and consonants is not always made independently of an analysis into syllables, or at least the identification of syllable "nuclei," it has proved difficult to develop criteria of a purely phonetic nature by which to determine the boundaries between adjacent syllables. Proposals that the syllables be defined language-specifically as an element with respect to which regularities of sound-type distribution may be described have undoubted merit, but the unit so defined is not to be lightly identified with the syllable of our phonetic intuition. The utility of a phonologically defined syllable says nothing one way or the other about whether we may reasonably hope to define the syllable in purely phonetic terms. Of course, if our conviction that it is worthwhile to continue the effort to define the syllable as a phonetic element is based on our phonetic intuition, then we must be clear as to just what that intuition tells us. What it seems to tell us most reliably (how reliably?) is how many syllables comprise an utterance. It informs us much less precisely about where we should feel that one syllable ends and the next begins, if indeed it tells us much at all on this matter that is uncontaminated by knowledge of the phonotactics and grammar of the specific language. The fact that rules devised to decompose speech into syllables are generally not at all what we should call phonetic--they are formulated either on a phonotactic or strictly acoustic basis--would in itself suggest the possibility that syllables are not physically discrete entities at all. Thus, while we are able to locate the syllable "peaks," the "syllabic" elements of the sound sequence, in the case of the consonants we feel impelled to locate them to one side or the other of a syllable boundary; it is not enough for linguists to say simply that they constitute syllable "troughs." Perhaps the problem of locating a syllable boundary is serious only because there is a problem of
spelling encountered anytime we feel the need to hyphenate a word in a phonetically justifiable manner. In introducing a hyphen, we in effect treat syllable boundary as just another element in the linear sequence of letter-shapes. There are, to be sure, two situations in which we are very sure of where to place such a boundary: the onset and the termination of speech coincide, respectively, with the onset and termination of a syllable. It is not obvious, however, that the phonetic properties of speech onset and termination should serve as the basis for analyzing speech into syllables generally.

**COARTICULATION: EVIDENCE FOR THE SYLLABLE?**

The difficulties encountered in trying to define the syllable as a phonetic unit so that it corresponds exactly to our intuition do not mean that phoneticians refrain from speaking facilely on the subject. As already mentioned, there is the notion that the syllable is a unit of speech production, in that the sequence of sounds composing this unit is generated by a single unanalyzable gesture,¹ and not by a sequence of freely commutable ones. Complementary to this is the assertion that we apprehend speech in units of syllable-size, since there is ample evidence that the phonetic evaluation of an acoustic segment is not context-independent. These phenomena of "coarticulation" (and "copercception") provide some of the strongest motivation for viewing the syllable as a phonetic unit, provided we suppose that these effects are coextensive with the syllable. Coarticulatory effects are all those phenomena that cannot be accommodated by a strictly segmental hypothesis of speech production, according to which the speech sounds of a language are analogous to the letters of typescript, each with a shape unaffected by its neighbors. The evidence that speech is not produced in this way is not necessarily evidence for the syllable, for it is not clear that the scope of all the coarticulatory effects so far observed coincides with the syllable, insofar as its boundaries are well established on intuitive, distributional or other grounds. All that we do know is that when we closely examine the operation of the various structures making up the vocal tract, we find that they do not in general shift position in close synchrony and at the rate at which the phonetic segments are emitted. Thus lip rounding and nasalization, two favorite topics in the study of coarticulation, are segmental features (in English) that refuse to be confined within their "proper" segmental boundaries. It is not obviously true that they pay more attention to syllable boundaries, wherever these may be located. Nasalization, which we associate in English linguistically with consonants, may color a preceding vowel, and rounding, a feature of vowels, may color preceding consonants; we

¹The term "ballistic" has sometimes been applied, but it is not clear what the term conveys. If it can with some plausibility be used in referring to either an opening or a closing movement of the articulators, it seems to stretch the connotation of the word beyond recognition if it is used to describe an opening-closing sequence as a single movement. Moreover, if it is legitimate to call this sequence a single ballistic gesture, why not as well apply the term to a sequence of closing and opening movements?
do not know that such anticipatory coarticulation depends crucially on the absence of a syllable boundary.

Coarticulatory phenomena, which have been understood to tie in somehow with the syllable, are of course measured in time, and the temporal dimension is connected with the syllable in another way also, in that it is the syllable to which we appeal in treating the more narrowly temporal properties of speech, namely tempo and rhythm. Connected with tempo and rhythm, though the connection is a complex one, are the durations of intervals in the speech signal that can be said to correspond more or less to the phonetic segments into which the listener/linguist resolves the signal. An extensive literature (Lehiste, 1970) reporting measured durations of acoustically defined intervals corresponding to vowels and stop-consonant closures has provided evidence that segment duration is determined significantly by a number of factors—by overall speech rate, by articulatory properties of the segment (setting its "intrinsic" duration), by rhythm (stress pattern), and by phonetic properties of neighboring segments. Of these, it is the last that has the most immediate bearing on the question of the syllable; tempo and rhythm clearly affect intervals that comprise more than one syllable, while intrinsic duration is by definition a segmental attribute. The existence of coarticulatory relations between adjacent segments does not constitute evidence for the syllable as a phonetic unit; for that we must find that coarticulatory linkages are markedly weaker between segments said to belong to different syllables.²

**Vowel and Stop Durations in Monosyllables**

One of the best-known cases of temporal coarticulation is the relation between vowel duration and the voicing of a following consonant, according to which vowels preceding voiceless stops and fricatives are shorter than before their voiced counterparts. This relationship is most clearly seen when we compare isolated monosyllables that differ only in the voicing of their final stops. In a word pair like cop-cob the durations of the vowels, however, defined acoustically, show a ratio of about 2/3. Moreover, if the final stops

---

²That differences in linkage between acoustic segments can affect listeners' place of a syllable boundary was demonstrated long ago by Malmberg (1955). Synthetic speech patterns interpreted as vowel-stop-vowel sequences were divided into syllables differently, depending on whether formant transitions were supplied before or after a silent interval corresponding to the stop closure. The stop was grouped with the vowel having the transition. Unfortunately, in natural speech such sequences show transitions both before and after the closure interval, so that, for example, black-out (assuming we agree to put a syllable boundary after /k/) will not lack a transition from the /k/-closure to the following vowel, and the postclosure signal will be a convincing gout when heard alone. If a recording of the expression Greek odds is segmented so that only the postclosure part is heard, the listener will report the same word that he hears when the same operation is performed on the expression Greek gods. The pair can be disambiguated by a talker, but that process certainly involves a disruption of the articulatory "plan" of normal fluent performance.
are released so as to produce a noise that allows us to measure the duration of stop closure from the acoustic record, we find that the duration of /p/ is longer. Since the combined durations of vowel and following closure tend thus to be more nearly equal for the two words than either vowels or stops taken separately, and since vowel and stop must here be tautosyllabic, it seems reasonable to suppose that this temporal interdependence is a feature of syllable organization, especially if it turns out that the entire CVC sequence tends to be of equal duration for the two words. We should then perhaps be able to say that the voicing difference associated with the final stops is to be related very directly to the voicing difference, let us say, between coo and gob, where the onset of voicing relative to the articulatory program of the syllable is differently timed. For cop vs. cob the difference is in the timing of the devoicing gesture relative to the end of the syllable; with the earlier onset of devoicing it happens in English that the accompanying stop closure is also advanced in time—perhaps because articulatory closure is itself a devoicing maneuver! To the extent that temporal relations of this kind can be established for isolated monosyllables, we may then turn to polysyllabic utterances, for example the "words" /kābi/ and /kāpi/, whose intervocalic stops are not so readily assigned to one syllable or the other. If we find that the temporal relationship between stressed vowel and following stop is like the one in the monosyllables, then we may claim that the argument for placing a syllable boundary before the stop is thereby weakened. Of course, if no sequence of vowel + stop can be turned up that fails to show this relationship, no matter what its location with respect to the syllable boundary of our phonetic (?) intuition, then the argument for the syllable as a phonetic unit is undermined.

STRESSED VOWEL AND STOP DURATION IN CVCV

In order to have some specific numbers to talk about in this connection, I had a single native speaker of American English produce, in random sequence, a number (61) of tokens in the following sequences: /kāpi/ /kābi/ /kāmi/ /gāpi/ /gābi/ /gāmi/. (The forms with /m/ were included in order to have a bilabial closure associated with nondistinctive voicing.) These were recorded and analyzed spectrographically (Kay Sonograph Model 7029A). The following points were fixed on the spectrograms: onset of the release burst of the initial velar, the onset of voicing following the /k/-aspiration, time of establishment of the bilabial closure and release of the bilabial closure.

CVC as Temporal Unit

Figure 1 shows the frequency distribution of durations measured from onset of the initial burst to the termination of the medial closure, an interval here labeled "stressed syllable," though no prejudice is implied favoring the analysis of CVCV into CVC-V over CV-CV. For this particular speaker, and for the particular set of forms elicited (over six sessions in the course of several weeks), the measured interval averages about 300 msec, with a standard deviation (s) of just over 23 msec. We might want to regard this 300 msec as the "target" or "intrinsic" duration for CVC-sequences generally, or perhaps for sequences of velar stop + /a/ + bilabial stop. Consideration of the durations of these sequences in the six forms taken separately (Figure 2) suggests that they may differ significantly for the
DURATION OF "STRESSED SYLLABLES" IN $C_1V_1C_2V_2$

$C_1 = /k, g/$  
$V_1 = /i/$  
$C_2 = /p, b, m/$  
$V_2 = /i/$  

Mean = 299 msec  
$s = 23.2$  
$n = 366$

Figure 1: Frequency distribution of durations measured from release of initial velar stop to release of medial bilabial closure. Pooled data of six CVCV sequences.
Figure 2: Durations of intervals from initial to medial stop releases. Upper panel: mean durations of measured intervals for each CVCV sequence type. Lower panel: frequency distributions of CVC for the six types.
Figure 3: Durations of intervals from /k/-releases to onset of glottal pulsing. Left display represents summation of data shown on right.
Figure 4: Top row: distributions of durations from /g/-release to establishment of medial closure. Middle row: durations of intervals from /k/-release to beginning of medial closure. Bottom row: durations from onset of glottal pulsing following /k/-release to beginning of medial closure.
DURATIONS OF ORAL CLOSURE IN MEDIAL CONSONANTS

/m/
mean: 86.2 msec
s: 11.04
n: 122

/kami/
mean: 82 msec
s: 11.3
n: 61

/gami/
mean: 90 msec
s: 9.4
n: 61

/b/
mean: 94.8 msec
s: 10.42
n: 122

/kabi/
mean: 91 msec
s: 10.9
n: 61

/gabi/
mean: 98 msec
s: 8.8
n: 61

/p/
mean: 135.8 msec
s: 11.72
n: 122

/kapi/
mean: 131 msec
s: 11.1
n: 61

/gapi/
mean: 140 msec
s: 10.6
n: 61

Figure 5: Frequency distributions of medial bilabial closure for /p, b, m/.
different \( C_1 \) and \( C_2 \). While mean values for /kāmi/ and /gāmi/ are obviously not different, and /kāpi/ and /gāpi/ are not clearly different, /kābi/ differs from /gābi/ by an amount that is significant at the .005 level. And even if we entirely discount any differences between /k----/ and /g----/ forms, pairs differing with respect to their medial consonants show mean durational differences that in only one case, /gāmi/ vs. /gābi/, fail to meet a criterion of \( P < .01 \). From Figure 2 it appears that CVC- durations depend significantly on the second consonant: CÚp- > CÚb- > CÚm-. The role played by \( C_2 \) is quite evidently more important than any by \( C_1 \).

**Durations of CÚ and C_2 Taken Separately**

Turning now to the component segments to determine how much these vary and to what extent any variability can be ascribed to context, we see, first of all (Figure 3), that /k----/ forms show no significant differences in the intervals between onset of the release burst and onset of laryngeal pulsing. (This interval, the first measured segment of these forms, can be called, as it has been, either "aspiration" or "devoiced vowel," depending on whether one elects to say it is "part of" \( C_1 \) or V.) The 3-msec difference between the means of /kābi/ and /kāpi/, though there would be no difficulty in devising explanations for it, must be dismissed as statistically nonsignificant.

Instead of deciding what to define as the duration of the stressed vowel (a dubious enterprise of indeterminate meaning), I chose to adopt both practices that have been followed, measuring both from release of the initial stop and from the onset of voicing in the /k----/ forms. In these forms /a/ was taken to correspond both to the entire interval from release to closure (Figure 4, middle row) and to the voiced part of that interval (Figure 4, bottom row). Each of these intervals was compared with the first segments in /g----/, in which any lag in voice onset behind the release was neglected in my measurements. By both measures the /a/-durations differ significantly, depending on both \( C_1 \) and \( C_2 \); the one exception is in /gāmi/ vs. /gābi/. [These relations are very like those reported long ago in Peterson and Lehiste (1960), although their prenasal vowels were longer than those before voiced stops.] The ratios of /a/-durations before /p/ to those before /b/ have a mean value of .85, somewhat greater than the .67 more commonly reported for isolated monosyllables.

The mean values of closure duration are significantly different for the three bilabial consonants (Figure 5): \( p < b < m \). Moreover, there seems to be some kind of relation between the voicing of the initial stop and the duration of the medial closure, or perhaps it would be more reasonable to relate the duration of the closure to the duration from initial release to the beginning of the closure interval. The greater duration of closure in the /g----/ forms seems to represent an instance of "compensatory" lengthening, although in the case of /b/ this lengthening was not quite enough to equalize total CVC durations of /kābi/ and /gābi/ (they differ at the .01 level). On the other hand, the significant difference between /m/ and /b/ durations (\( P \ll .01 \)) is not matched by any appreciable difference in the intervals from /g/-release to onset of medial closure (\( P = .05 \), i.e., < .01). All this says no more than we have already seen, that total durations of CVC in the CVCV forms show generally significant (\( P < .01 \)) differences in their means, depending on the
second consonant particularly, and the differences are especially striking when one item of a comparison has the voiceless stop as \( C_2 \). (Note, however, that the difference between /kámi/ and /kábi/ is greater than between /kábi/ and /kápi/).

**The Variance Test—Tested and Failed**

There is another test that has been applied in similar situations, although its significance is questionable. It compares the sum of the variance of two continuous intervals with the variance of the summed intervals; if the latter quantity exceeds the former, then the two intervals have durations that are negatively correlated. Ohala (1973) has pointed out that variability in locating the boundary between segments will itself yield a negative correlation even if two intervals A and B are truly independent, that is, the relation between their variances \( (s^2) \) will be: \( s^2(A) + s^2(B) > s^2(A + B) \). On the other hand, a relation \( s^2(A) + s^2(B) < s^2(A + B) \) could well result from a variation in "overall speaking rate," a known variable whose effect on segment duration must somehow be cancelled if the "temporal patterning of speech" (a feature that is by definition independent of speech rate) is to be determined. If the first relation can reflect uncertainties in measurements as much as timing "errors" in articulation, and the second results from macroscopic rate variability with no clearly identifiable contribution of the "temporal patterning" factor we are looking for, then a comparison of variances provides data that are part of the problem rather than of the solution. Even if \( s^2(A) + s^2(B) = s^2(A + B) \), we are no better off, since there is no reason to exclude this result as a possible outcome of the two competing effects of measurement error and rate variation, given that we may encounter considerable difficulty in making precise determinations of the magnitude of those effects.

Despite these and other caveats that might be raised, the variance relations among the six forms measured are presented in Table 1. If we could ignore the factor of measurement error, we would infer a negative correlation between duration of aspiration and following voiced vocalic interval, while closure intervals are positively correlated with immediately preceding segments, except in the case of /gápi/. These findings are, at least intuitively, in conformity with our belief that the initial stop is tautosyllabic with the following vowel, while the syllable affiliation of the medial stop is uncertain. On the other hand, they are not in conformity with conclusions based on the relations among mean durations of the segments measured, since mean durations of /k/-aspirations show no differences for the different durations of the voiced intervals that follow; in contrast, the clear relationship between the latter interval and closure duration is not apparent in the variance relationship. Since we are not prepared to accept a view that /k/ is not tautosyllabic with /á/, we may by the same token be less willing to use the variance relation between vowel and following closure as the basis for deciding the place of a syllable boundary, that is, analyzing our forms as CV-CV.

On this last point, if the usefulness of the variance relationship has not already been sufficiently impugned by Ohala (1973), we ask what this measure reveals when applied to sequences that are considered monosyllables—
### TABLE 1: Variance relations.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(\frac{3}{1+2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s^2)</td>
<td>(vowel)</td>
<td>(closure)</td>
<td>(1 + 2)</td>
<td></td>
</tr>
<tr>
<td>/gápi/</td>
<td>228</td>
<td>112</td>
<td>303</td>
<td>.89</td>
</tr>
<tr>
<td>/gábi/</td>
<td>207</td>
<td>77</td>
<td>388</td>
<td>1.36</td>
</tr>
<tr>
<td>/gámi/</td>
<td>213</td>
<td>88</td>
<td>335</td>
<td>1.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(\frac{3}{1+2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s^2)</td>
<td>(aspiration)</td>
<td>(vowel)</td>
<td>(1 + 2)</td>
<td></td>
</tr>
<tr>
<td>/kápi/</td>
<td>90</td>
<td>222</td>
<td>262</td>
<td>.84</td>
</tr>
<tr>
<td>/kábi/</td>
<td>106</td>
<td>250</td>
<td>269</td>
<td>.76</td>
</tr>
<tr>
<td>/kámi/</td>
<td>104</td>
<td>190</td>
<td>231</td>
<td>.79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(\frac{3}{1+2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s^2)</td>
<td>(aspiration)</td>
<td>(closure)</td>
<td>(1 + 2)</td>
<td></td>
</tr>
<tr>
<td>+ closure)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/kápi/</td>
<td>262</td>
<td>123</td>
<td>543</td>
<td>1.41</td>
</tr>
<tr>
<td>/kábi/</td>
<td>269</td>
<td>119</td>
<td>506</td>
<td>1.30</td>
</tr>
<tr>
<td>/kámi/</td>
<td>231</td>
<td>128</td>
<td>388</td>
<td>1.08</td>
</tr>
</tbody>
</table>
the English words /kap/ and /kab/. Two sets of ten repetitions of these words were recorded by the same speaker who furnished the dissyllables just discussed. The variance ratios obtained were these:

<table>
<thead>
<tr>
<th>TABLE 2:</th>
<th>Trial 1</th>
<th>Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kap/:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh_a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh + a</td>
<td>.95</td>
<td>.76</td>
</tr>
<tr>
<td>ap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a + p</td>
<td>.65</td>
<td>.29</td>
</tr>
<tr>
<td>/kab/:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh_a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh + a</td>
<td>1.14</td>
<td>1.13</td>
</tr>
<tr>
<td>ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a + b</td>
<td>.73</td>
<td>1.10</td>
</tr>
</tbody>
</table>

If these figures mean anything (and, if anything, how much?), they indicate absence of any clearly negative correlation between aspiration duration and following voiced vocalic interval, possibly a significant difference in the degree of correlation between the duration of the closures of the final stops (all were produced with final releases) and the preceding segments. If we combined the voiceless and voiced intervals between the initial releases and final closures, we obtain the following ratios:

<table>
<thead>
<tr>
<th>TABLE 3:</th>
<th>Trial 1</th>
<th>Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kap/:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>khap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh_a + p</td>
<td>.47</td>
<td>.37</td>
</tr>
<tr>
<td>/kab/:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>khab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh_a + b</td>
<td>.96</td>
<td>1.21</td>
</tr>
</tbody>
</table>

These figures tempt us to suppose that a negative correlation exists between first and second intervals when /p/ is involved, but not when the final stop is /b/. If the variance ratio tells us something about "syllable organization," it presumably provides no guidance in the matter of syllable boundary placement, since I think we are unprepared to deny monosyllabic status to /kab/ or to believe that /kab/ and /kap/ differ in the number of syllables
composing them. If the differences in variance ratios have any meaning, they indicate only a difference in the temporal organization of the two syllables, nothing about how we should decompose phonetic sequences into syllables. If a ratio greater than unity may hold between segments belonging to the same syllable, in the case of a monosyllable, then the argument for using that ratio as a guide to syllable division becomes suspect. Since the ratios of Table 1 representing relations between medial closures and preceding segments are almost all greater than one, they do not point to any special relation between the two. If these ratios do not give the "right" answer in the case of the monosyllables, why should we use them to resolve the question of syllable division where phonetic intuition speaks with a forked tongue?

REFERENCES


---

3In this connection, Fujimura and Lovins (1977) propose that final voiced obstruents be considered "phonetic affixes" rather than "core" constituents of syllables. The data presented here are consistent with their proposal. However, an analysis of spoken English into sequences of syllabic "cores" and "affixes" would yield elements rather different from the syllables of our phonetic intuition.