WHAT HAPPENS TO VOWELS AND CONSONANTS WHEN WE SPEAK FASTER?

Sidney Wood

October 1973

SUMMARY

The subjects of this article are Kozhevnikov and Chistovitch's finding regarding the constancy of relative syllable durations and the inconstancy of relative speech sound durations for different rates of speech, and their proposal that vowels are elided because the necessary minimum duration of a consonant consumes all the time that happens to have been assigned to that syllable.

The relative consonantal durations of my seven informants (six languages) varied with speaking rate but mustly not as predicted by Kozhevnikov and Chistovitch. Instead of rising continuously at faster rates, the relative consonantal durations decreased again (Figs. 2 and 3). The variation was so small however, that a linear model, assuming a constant consonantal proportion, provides an excellent approximation to the results (Figs. 4, 5, and 6). A spot check on the syllables in two selected words in the German sample L revealed that the relative syllable durations were not less variable than the relative consonantal durations there (Fig. 9).

Several problems and difficulties related to speech reduction are discussed in general terms (§ 4.2). Are segments squeezed out when temporally constrained, or are they deliberately omitted? If segments or gestures are suppressed during production, does this occur peripherally or centrally? The view is expressed that most examples of segment synkope and syllable contraction in everyday speech are regular and habitual, and are not necessarily caused by increasing rate although their occurrence has the effect of accelerating the message.

1. INTRODUCTION

1.1. Aim

Kozhevnikov and Chistovitch (1965) found that "upon the increasing of the duration of a phrase, there is a clear lessening of the portion of the time occupied in it by the consonants" (p. 87).

The converse decrease of the relative vowel duration at faster rates was accompanied by the total elision of a vowel. In consequence of this, they went on to formulate a hypothesis of reduction, whereby a consonant must be executed in a necessary minimum duration, possibly leaving no time over for the vowel in that syllable. In their own words, the loss of a vowel occurs "when the interval between the syllables assigned in the programme is too small in order to accomplish both the closing and the opening of the organs which articulate the consonant in the case of a rapid rate of speech. As a result, there is simply not sufficient time for the vowel" (p. 89), and again "the changes of rate significantly change the relative durations of the consonant and vowel within a syllable. In the case of a rapid rate of speech the vowel can disappear completely (full reduction of the vowel) and in the case of a considerable slowing of the rate the duration of the consonant practically does not change and the prolongation of the syllable occurs at the expense of the vowel" (p. 89).

They had previously concluded that the variation of the relative duration of syllables was random. In view of the apparent inconstancy of phoneme relative durations, they decided in favour of the syllable as the basic programme unit of speech:

"If we examine the syntagma as a sequence of sounds of speech, we cannot find any constancy in its time figure. However, if we turn to the syllables and consider the syntagma as a sequence of syllables, its rhythmic figure is an invariant and completely independent of the rate of speech. From this conclusion naturally follows that in the programme of a syntagma the syllable commands are rhythmically controlled" (p. 89).

I wish to query the following two points:

Firstly, is it true that the relative durations of vowels and consonants vary with speech rate, and, if so, in the way described by Kozhevnikov and Chistovitch? This is important for their conclusion that syllables are the programme unit.

Secondly, are segments "squeezed out" for want of sufficient time for their execution? This is important for theories of reduction.

Nooteboom and Slis (1969) recognized that if Kozhevnikov and Chistovitch were right about the relative durations, it would be a "very important finding for the research on the control of timing and articulation". They therefore performed a similar experiment, but unlike Kozhevnikov and Chistovitch (who had aggregated the consonants and the vowels through the test sentence) they compared individual consonants and vowels in their respective positions in two nonsense words <u>mamamám</u> and <u>mɑmɑmɑ́m</u>. They were unable to confirm that segment relative durations were entirely correlated with speaking rate. At fast rates, they tended to remain constant (except in the internal weak syllable for which they accepted Kozhevnikov and Chistovitch's reduction hypothesis). At slow rates, the relative vowel durations did increase, more so in the long vowels <u>ɑ</u> and less so in the short vowels <u>a</u>. There was also a difference between the subjects. They conclude by suggesting that "a slow speaking rate distorts in some way or other the rhythmic figure of a word in the brain of the speaker".

Gaitenby (1965) found that the relative durations of words, syllables and phonemes remained constant at different rates. However, the rate differences were those between the normal speech of "fast", "medium" and "slow" speakers. This does not exclude the possibility that there are differences when each speaker varies his own rate of speech.

1.2. Some theoretical difficulties

1.2.1. In any investigation of tempo, it is necessary to distinguish between gross and net measures of rate (Clevenger and Clark [1963], Goldman-Eisler [1958, 1961], Kelly and Steer [1949] and Wood [1973:§1]).

The distinction between gross and net measures can be obscured by the varying degrees of abstraction of the speech units counted in the analysed sentence. It is possible to count segments actually represented in the speech wave, or segments deemed to have been present in some idealized underlying construction. The "phoneme" has been a notoriously vague concept in this respect, with much controversy between schools. But even "syllables" and "words" are not always as concrete as we might like to imagine. I have discussed this problem in greater detail in my 1973:§3. The difficulty of the degree of abstractness is especially relevant to the present problem. Kozhevnikov and Chistovitch had a 7 syllable test sentence Tonya topila banyu, where the second vowel disappeared at faster rates (i.e. at shorter sentence durations), yielding (I assume) a 6 syllable rendering similar to [toɲ topila banyu]. If this is so, then by taking the utterance duration to indicate tempo, they would have been using a gross type of measure for rate, disregarding the reduction - the sentence was uttered in a briefer duration hence the message was transmitted more quickly. On the other hand, if we were to take account of the fact that at shorter sentence durations there were fewer syllables generated, then we would be using a net type of measure that paradoxically might show a slower articulation rate after the reduction.

Suppose a sentence were uttered in an unbroken phrase of n syllables in t_1 seconds, and then repeated more quickly in the briefer duration t_2 seconds but now in (n - 1) syllables following reduction at the "faster" rate. It is given that $t_1 > t_2$ seconds, but any of the following relations between the net syllable rates are possible:

$$\frac{n}{t_1} > \frac{n-1}{t_2}, \quad \frac{n}{t_1} = \frac{n-1}{t_2}, \quad \frac{n}{t_1} < \frac{n-1}{t_2} \qquad \text{sylls/sec}$$

For example, in the case where n = 8 syllables and $t_1 = 2$ seconds, the following values of t_2 will make each of the three relations true: $t_2 > 1.75$ seconds, $t_2 = 1.75$ seconds and $t_2 < 1.75$ seconds respectively. If the 8 syllables are uttered in 2 seconds, we have a net rate of 4 syllables/second. If the message is transmitted "more quickly" in, say, 1.8 seconds but with contraction of 1 syllable, we find a net rate of 7/1.8 syllables/second (3.9) which is "slower". Clearly, it is necessary to choose between the two measures with care, when a "faster" rendering can be articulated "more slowly".

It might be argued that it is not a gross measure but a net measure that indicates how hard the articulators are working during an utterance and is therefore more relevant for discussions of reduction. In the paradoxical situation cited above, the load on the articulators seems to have been lightened by the syllable contraction, despite the briefer utterance duration. Consequently, any experiment designed to relate segment reductions to gesture durations and frequencies ought to be based on a concrete measure rather than an abstract measure. Fortunately, this difficulty can be circumvented by avoiding the type of sentence in which syllable contractions occur, so that there is always the same number of syllables generated at any rate of utterance throughout the entire experimental series. I have followed this policy for the experiments to be described below, and test sentences were carefully selected to avoid the possibility of syllable elision. In so doing, I am of course implying that elisions are not the result of increased speaking rate, and any attempt to prove that from the experiments would inevitably end in circularity. Instead, I shall put forward in the conclusion my belief that elisions are largely habitual and amount to rejection of redundancy, enabling the speaker to increase his message transmission rate (if he wishes) without having to speed up articulation to the same degree. The English sentence Mary wants to go to Swansea can hardly be uttered in anything but 8 syllables. In contrast, the sentence perhaps he will go to Brighton contains two possible syllable contractions, reducing to p'raps e'll go to Brighton, with the added complication that the final syllable might contain either a syllabic n or a vowel.

1.2.2. We all know intuitively what a vowel is but it has never been one of the more easily defined concepts, which can lead to segmentation difficulties. It is wiser to avoid test sentences containing other syllabic sonorants than vowels.

Similarly, while it is often profitable in phonology to describe long vowels and diphthongs as simple vowels followed by semivowels, there is no phonetic boundary within a long vowel or diphthong that could be unambiguously used for measuring purposes. All post-vocalic semivowels (if indeed they have concrete reality outside of phonology) have therefore been included in the vowel aggregates in my test sentences. I could have restricted the set of test sentences by excluding long vowels and diphthongs, but my possible test sentence structure is severely constrained as it is, and some of my subjects already had difficulty in helping me compose suitable sentences in their languages.

Prevocalic semivowels have been included in the consonantal aggregates, but they have been avoided as much as possible as the acoustic wave form does not always permit conclusive segmentation. Post-vocalic nasals, especially in a ..VNC.. combination, revealed a tendency to disappear in anything but slow careful speech, leaving only nasalization on the vowel. This is in itself an interesting observation of a possible phonetic universal, but it is nevertheless a disturbing factor in this type of investigation and several of my repetition series had to be discarded because it was impossible to resolve this difficulty.

The price of these precautions might be decreased generality of the results, but we hardly have any use for test sentences that we do not know how to measure.*

1.2.3. Kozhevnikov and Chistovitch's relative consonant duration refers to an average for the whole sentence, not for selected individual consonants. All consonant durations and all vowel durations were aggregated separately and the consonant durations expressed relative to the vowel durations for the entire utterance, unity being assigned to the vowels. In their own words:

"Since the fluctuations of the relative durations of the sounds of speech

* These problems suggest that the categories of "vowels" and "consonants" may be too comprehensive for this type of investigation. It might be more profitable to investigate the temporal structure of segments in greater detail by looking at the individual sub-classes, in specified syllable structures, as at Figure 9.

in individual syllables were considerable, use was made of an integral index consisting of the ratio of all the consonant intervals to the sum of all the vowel intervals" (p. 87).

Their Figure 3.9 (p. 88) showed that the value of the $\Sigma C/\Sigma V$ ratio increased as utterance duration shortened, indicating that the consonantal portion of the utterance increased at faster rates. With the help of their Figure 3.9, these ratios can easily be transformed into proportions of the utterance duration, which are more immediately comprehended (this was also the form in which Kozhevnikov and Chistovitch expressed their finding in their statement quoted in my opening paragraph). The consonantal proportions of the sentence fell roughly from 35 % to 29 % as the utterance duration lengthened from 0.8 to 3.0 seconds. The vowels are the complement of the consonants with respect to the sentence and we consequently find the vocalic proportions increasing from about 65 % to about 71 % as the sentence duration lengthened.

Incidentally, their Figure 3.3 (p. 81) gives an example of the oscillographic write-out of the articulator electrodes and the speech wave for one rendering of the test sentence. For this illustration, the $\Sigma C/\Sigma V$ ratio was about 0.58:1, the vowels occupying about 62 % of the utterance. These values are typical for short utterance durations according to their Figure 3.9. However, the duration of this rendering was about 2.6 seconds, which is near the longest of the series. This rendering indicates that considerable variation is to be expected in this type of experiment.

2. PROCEDURE

2.1. Methods

2.1.1. The experiment consists in having a suitable test sentence repeated several times at different rates and then making suitable numerical analysis of the measured consonant, vowel and sentence durations.

Kozhevnikov and Chistovitch's Russian sentence had consonants specially chosen to contain labial or coronal gestures that would close electrical circuits at electrodes suitably placed in the mouth, thereby providing automatic segmentation. Nooteboom and Slis also used lip electrodes to facilitate segmentation in their mVmVmVm test words. I have used natural sentences with consonants that are not too difficult to distinguish from vowels on an oscillogram of the wave form, but at the same time sufficiently varied to avoid monotony or articulatory difficulty.

2.1.2. Some experimenters investigating temporal phenomena in speech have found it desirable to dictate a rate to their subjects, especially to hold tempo constant. For the present type of experiment, the aim is the opposite, to elicit a wide variation of tempo. Kozhevnikov and Chistovitch nevertheless chose to dictate rates to their subjects, the closure of the first t initiating a time signal that interrupted the speaker after one of 7 set sentence durations from 0.8 to 3.0 seconds. The speaker had to learn to adjust his rendering to the given interval and as Kozhevnikov and Chistovitch report, "the duration of the first pronunciations, as a rule, differed significantly from the assigned durations as the subjects sought the required rate of speech". Unsuccessful attempts were then rejected. Nooteboom and Slis simply asked their subjects to speak slowly, normally and quickly. The three rates were freely chosen by the subjects and data was obtained on a vast range of rates of speech. I have followed a similar procedure.

2.1.3. Kozhevnikov and Chistovitch's two informants made 25 to 30 attempts at each of the assigned sentence durations and for three different stress positions, yielding a total of 1260 attempts. Each speaker's five closest attempts at each sentence duration for a given stress position were selected for further processing. The point at each of the 7 utterance durations on the graph at the authors' Figure 3.9 presumably represents the $\Sigma C/\Sigma V$ quotient for the average of five renderings. Nooteboom and Slis had at least 20 repetitions each of slow, normal and fast rates. The slow rate was subsequently defined as test word durations in excess of 700 ms (corresponding to net rates of less than 4.3 sylls/sec). Their results were then averaged within each of the three broad rate classes. I have sought to obtain a total of 20 to 30 repetitions of each sentence to cover all rates. Some informants provided up to 40 repetitions.

2.2. Measurement and treatment

The data resulting from the experiment are absolute consonant, vowel and sentence durations, which were measured on oscillograms of the wave form and intensity, the paper running at 125 mm/sec to line up with Voiceprint spectrograms (Fig. 1). The spectrograms were used to help solve segmentation difficulties. The aggregated consonant durations are then expressed as a proportion of the sentence duration.

2.3. Possible outcomes

We can expect any of several outcomes from the experiment when speech is accelerated, for example that the consonants occupy (a) a progressively larger proportion of the sentence (as described by Kozhevnikov and Chistovitch), or (b) a progressively smaller portion of the sentence (the opposite to their predicted tendency) or (c) the same portion of the sentence (relative durations remaining constant, contrary to the belief that they vary).

Such outcomes will be indicated by the degree and sign of the correlation between consonantal proportion and speaking rate.

Alternatively, these outcomes can also be revealed by the linearity of the regression of absolute consonant durations for different rates. If (a) is true, the consonant durations will fall more steeply as sentences become briefer. If (b) is true, the consonant durations will fall less steeply as sentences become briefer. If (c) is true, the consonant durations will fall uniformly as sentences become briefer. There will thus be non-linear regressions for (a) and (b) and a linear regression for (c). The steepness of the slope depends on the proportion of the sentence occupied by the consonants, a larger proportion giving a steeper slope. The changing slope of the non-linear regression will therefore reflect changing relative consonant duration. If one linear regression is steeper than another, then the constant relative duration will have been that much larger.

2.4. Informants and test-sentences

Seven informants took part - B (Southern British English), I (Chinese), J (Polish), K (Southern Swedish), L (German), M (Southern Swedish) and N (Egyptian).

The sentences were pronounced with normal sentence intonation (M also recorded a contrasting series with an alternative intonation):

B1 Mary wants to go to Swansea

B2 Spiker's kidney pies are bigger

I Shuo hua shi ren lei chuan you de xing wei

J Krakowski pociąg czeka tam

K Sven köpte det gamla huset

L Ich kaufte zwei gelbe Tische

M1 Sture fick ett tåg på teknis

M2 Sture fick ett tåg på teknis

N S̄id̄id katab qiṣṣagd̄ida

None of these sentences contained syllables that were likely to be contracted, which ensured a constant number of syllables throughout the experiment.

3. RESULTS

The consonantal proportions are plotted against sentence duration for each subject at Figures 2 and 3. Sentence duration can easily be transformed into rate (the number of syllables uttered is constant in each series, and syllable rate is the reciprocal of syllable duration). The net articulation rate has therefore been given alongside the sentence duration scale. The consonantal proportions have also been given in the table at Figure 4 for different rate classes, with a class interval of 1 syll/sec. The absolute aggregated consonant durations have been plotted against sentence duration (rate) at Figures 5, 6 and 7.

3.1. Relative consonant durations

Sample M2 (fig. 3) supports Kozhevnikov and Chistovitch's finding that the consonantal proportion continues to increase with speaking rate.

In samples B2 and N (Figs. 2 and 3), the consonantal proportion appears to vary randomly with speaking rate, suggesting it remains constant independently of rate variation.

For B1 (Fig. 2), the consonantal proportion also appears to vary randomly, although it is possibly a little smaller at slower rates. If so, this would agree with Nooteboom and Slis's finding that the relative consonant durations varied at slower rates.

The remaining samples I, J, K, L and M1 show a very different tendency. The consonantal proportion increases from slow to medium rates – as found by Kozhevnikov and Chistovitch and by Nooteboom and Slis – but then decreases again for fast rates (where Kozhevnicov and Chistovitch expect a continued rise). This type of outcome was not foreseen and could not have been expected from the previous investigations.

The table at Figure 4 highlights smaller differences than can be detected visually from the graphs at Figures 2 and 3.

The table not only confirms that M2 supports Kozhevnikov and Chistovitch's finding (the consonantal proportion rose continuously from 53 % to 58 %). It also shows that B1 follows the same tendency, althought with very small increments.

The table suggests that B2's may not vary as randomly as was thought and that this sample may follow the unforeseen tendency. But the variations between rate classes are extremely small.

The table shows that N's consonantal proportion may not vary randomly. It seems to fall slightly as rate increases (there are only examples of medium and fast rates for this speaker).

The tendency for the consonantal proportion to increase from slow to medium rates, and then decrease again to fast rates, was not foreseen as a possible outcome in § 2.3 above. This tendency was exhibited in most samples – I, J, K, L and M1. It is just possible that B2 and N also belong to this group.

3.2. Absolute consonantal durations

The aggregated consonantal durations are plotted against sentence duration (and rate) at Figures 5, 6 and 7. The tendency for relative consonantal durations to vary with rate will appear as a non-linear regression in this presentation (§ 2.3), but the relationship revealed in Figures 5, 6 and 7 seems very linear. The variations of relative consonantal duration indicated by the table at Figure 4 are so small (a few percent only) that they can hardly be detected on Figures 5, 6 and 7. The straight lines drawn on these graphs represent the constant consonantal proportions stated there and have been fitted by eye to pass through origo and the set of points.

The table at Figure 8 gives the product moment correlation coefficients for consonant and sentence durations and vowel and sentence durations. These are so close to 1 that they underline how close these regressions are to linearity (this correlation coefficient assumes a linear relationship between the variables).

4. DISCUSSION

4.1. Variations of relative consonantal durations

Practically all of the cases illustrated at Figures 2 and 3 and tabulated at Figure 4 show that there were some variations of relative consonant duration at different speaking rates, although a large majority do not follow the tendency described by Kozhevnikov and Chistovitch. The range of variation was small. Figures 5, 6 and 7 show that such small variations hardly deviate from a virtually linear relationship between absolute consonantal durations and sentence durations. If we look very closely at I, J, K, L or M1, we can just see that the absolute consonantal durations do rise a little more steeply at faster rates, and do flatten off a little at slower rates. But the changes of slope are very slight and I am sure they can only be spotted because we know the answer in advance from Figures 2 and 3. Had we been looking for a truly linear relationship in any other experiment, we would have been overjoyed to find a set of plots like those of Figures 5, 6 and 7. The proximity of the coefficients to 1 is an indication that a linear model is an excellent approximation to these results (Fig. 8). The vowels are equally intimately correlated with speaking rate.

The problem is, are the very small departures from the excellent linear model nevertheless sufficiently large to warrant the conclusion that relative consonantal durations are invariant? Kozhevnikov and Chistovitch believed so. The degree of constancy of their relative syllable durations (that they concluded were invariant) may provide a yardstick to judge this by.

Their Figures 3.6, 3.7 and 3.8 illustrate the syllable relative to its word. I am not convinced that the variation they show is as random as they suggest. A definite pattern can be discerned, especially in their word <u>topila</u>. These variations appear to be as large as the variations of their consonantal proportions in the sentence. For comparison, I have taken two words from my German subject <u>L</u>, <u>kaufte</u> and <u>gelbe</u>. Figure 9 shows the variations of the word durations relative to the sentence, the syllables to the words and the consonants to the syllables. The consonantal proportion of the sentence has been included for reference. Figure 9 shows that the relative duration of none of the units was invariant. The words show the least variation - they decrease slightly from slow to moderate rates, and then rise again at faster rates. The stronger syllables <u>kauf</u> and <u>gel</u> show greater variation than the words. The weaker syllables are the complement of the stronger syllables in their respective words and have not therefore been included on the graphs - <u>te</u> falls from 32 % to 27 % and then rises again to 33 %, while <u>be</u> rises from 31 % to 40 %. This quick spot check does not support the idea that the relative durations of syllables should be less dependent on speaking rate than are the relative durations of the consonants. It might be worth while to investigate other syllables and words from my other informants and languages.

Even if Kozhevnikov and Chistovitch had been mistaken about the relative durations of phonemes and syllables, it can still nevertheless be true that the syllable is the programme unit. Many intuitively feel that the syllable is a basic unit of speech production. But some other type of argument and evidence may be necessary to confirm it.

4.1. <u>Synkope and syllable contraction</u>

There are two basic problems that need to be solved before a final theory of reduction can be built. Firstly, if a particular speech unit disappears in some renderings, has it been squeezed out because there is insufficient time to accomplish it, or has it been deliberately omitted? Are the necessary gesture commands extinguished at some point during the production process, or are they never initiated? Secondly, are reductions peripheral or central from the neuromotor point of view? There are possibly occasions when any of these alternatives may be true. I doubt whether a complete answer can be provided yet, but some aspects can be discussed briefly.

It is not easy to define an increase in tempo that can be related to reductions of the synkope type. The observed speaking rate, computed by counting how many of a given speech unit are produced in the measured time, represents the sum of the various influences acting on the temporal characteristics of speech segments. It reflects the consequences of reductions rather than the drive that is postulated to have occasioned them.

There are two phenomena that contribute to a shorter utterance duration and hence an accelerated message - shorter segment durations and coalescence or loss of segments. The first is the area observed in investigations of articulatory undershooting. The second is the area of Kozhevnikov and Chistovitch's hypothesis of elision.

4.2.2. Undershooting of gesture targets has been observed and related to the time available for their execution (e.g. by Stetson et alia 1940, Lindblom 1963, Gay 1968, Kent and Moll 1972). This is doubtless peripheral in character - a body travelling at a given velocity or accelerating at a given rate will not move so far if the duration of the movement is shortened. So long as there is a gesture to measure, it is a simple matter to relate undershooting and speech rate. But what happens if the observed segment is elided?

The loss of a segment is perceived because the rendering can be compared with a well known complete form. It would be tempting to consider such an elision as a case of maximum undershoot and relate observed reductions to the rate of transmission of the complete forms. For example, I found in a sequence of General American speech (recorded from a radio interview) the five syllables of <u>the Americans</u> spoken in 0.38 seconds, a rate of 13.2 syllables/second. We cannot articulate syllables at that speed, the expected maximum being about 8, perhaps 9, syllables/second (Wood 1973: § 4). In fact, he uttered three syllables, [ð i m e r k n̩ z] at 7.9 syllables/second, a very plausible fast rate of articulation. But to relate the reductions to the transmission rate of the complete forms would take for granted that elisions are of the squeeze-out type, and ignore the possibility that some might instead be deliberate omissions from the message. The first type of elision can be attributed to temporal constraints on the articulatory programme (i.e. encoding constraints), including Kozhevnicov and Chistovitch's hypothesis, the second type concerns the composition of the underlying message that is to be encoded. These are two possible ways in which the brain can work in this situation and must be born in mind during any discussion or speculation about the planning or programming that precedes the articulation of speech. The two cannot be distinguished in the speech output.

4.2.3. Simple experiments like Lehiste's (1970:7) comparison of the frequency of free apical vibrations in a trilled <u>r</u> (28 per second) with the voluntary apico-dental gesture of <u>t</u> (7 per second) suggest that the limiting factor for voluntary gestures lies within the nervous system, the bottleneck being the rate at which some higher motor centre can transmit sequences of different commands to the articulators. In syllable repetition experiments (Kaiser 1934, Hudgins and Stetson 1937, Sigurd 1971) this limiting rate appears peripherally as the rate at which the speaker has switched between different gestures, for example t-a-t-a-t-a... where the maximum rate appears as 7 or 8 syllables/second. Lehiste's experiment indicates that the articulators can move more quickly than the motor centres ever require for a deliberate gesture. Where synkope is a case of gesture commands being extinguished, it seems likely this would be due to the inability of a motor centre to pass on or switch between the necessary coordinated commands at a sufficiently fast rate, rather than to the inability of the articulators to respond properly owing to mechanical constraints.

It is clear from my wording of the previous paragraph that I imagine individual gestures, rather than whole segments, suffering from that type of neuromotor constraint. In test sentence B1, examples of the following reductions occurred in __...wants to...__: [w ´˜ n t s t ŭ], [w ɔ̃ n s t u], [w ɔ̃ s t u]. Only a few gestures were lost at a time and not the entire packages of gestures needed for __n__ and __t__. Note particularly that when the dental occlusion vanished, the __s__ remained voiceless without assimilating to the voiced __n__; similarly the nasal cavity remains open during the vowel even when the __n__ has vanished. In contrast I would expect the sequence __wands to__ ("wands to do conjuring tricks with") to reduce as follows: [w ɔ̃ n d z t u], [w ɔ̃ n z t u], [w ɔ̃ z t u]. If it is assumed that these are cases of gesture commands being extinguished, then it must also be admitted that the occlusions were discarded subsequent to the assignment of voicing assimilation to the inflexions __s__. I usually hesitate to agree that the ordered rules of a transformational generative model for phonology necessarily always reflect cerebral processes that are part of speech production, but the parallel is undoubtedly striking in these examples.

4.2.4. Supposing the Kozhevnikov and Chistovitch hypothesis were true, that vowels are omitted when the consonant happens to need all the time available for the syllable. Segment elisions ought then to occur at random, whenever there was a momentary shortage of time. But we know that vowels do not disappear at random in speech. Elisions and syllable contractions are largely non-random and are instead habitual and predictable from the environment. This is a view that is supported by the regularity of such phenomena synchronically in everyday speech and diachronically in sound change. For example, it is typical of many languages that certain weak vowels are omitted between obstruents and liquids, as in French app'ler or English delib'rate. The conventional, rather than tempo-dependent, character of such reductions is underlined when hyper-correct forms appear with spurious vowels as in Latin saec(u)lum. If the reduced forms lose contact with their complete form, and ultimately become established as the normal form, evidence of a lost segment may only appear in morphological alternations as in English hist'ry-historical, vict'ry-victorious.

Habitual reductions become part of the common speech code shared by all members of a definable speech community, and comprehension is not endangered. I suspect that these reductions are not the consequence of a rapid speech rate but rather that they permit more rapid rates. The articulatory programme is simplified and shortened by dispensing with some of the redundancy in the speech signal, enabling the linguistic message to be transmitted in a briefer period of time, that is, more rapidly. It was demonstrated in § 1.2.1. above that syllable contraction leads to a slight reduction of the articulation rate while the utterance becomes briefer and the message is transmitted more quickly. This reorganization of the articulatory programme permits more rapid message transmission rates while a comfortable rate of articulation is retained.

BIBLIOGRAPHICAL REFERENCES

Clevenger T. and M.L. Clark. 1963. Coincidental variation as a source of confusion in the experimental study of rate. Language and Speech 6: 144-150

Gaitenby J.H. 1965. The elastic word. Status Report of Speech Research, Haskins Laboratories, SR-2:3.1-3.12

Gay T. 1968. The effect of speaking rate on diphthong formant movements. Journal of the Acoustical Society of America 44:1570-74

Goldman-Eisler F. 1958. Speech analysis and mental processes. Language and Speech 1:59-75

Goldman-Eisler F. 1961. The significance of changes in the rate of articulation. Language and Speech 4:171-174

Hudgins C.V. and Stetson R.H. 1937. The relative speed of articulatory movements. Archives Néerlandaises de Phonétique Expérimentale 13: 85-94

Kaiser L. 1934. Some properties of speech muscles and the influence thereof on language. Archives Néerlandaises de Phonétique Expérimentale 10: 121-133

Kelly J.C. and Steer M.D. 1949. Revised concept of rate. Journal of Speech and Hearing Disorders 14: 222-226

Kent R.D. and Moll K.L. 1972. Cinefluorgraphic analysis of selected lingual consonants. Journal of Speech and Hearing Research 15:453-473

Kozhevnikov V.A. and Chistovitch L.A. 1965. Speech, articulation and perception. Joint Publications Research Service, Washington

Lehiste I. 1970. Suprasegmentals. MIT Press

Lindblom B. 1963. Spectrographic study of vowel reduction. Journal of the Acoustical Society of America 35: 1773-1781

Nooteboom S.G. and Slis I.H. 1969. A note on rate of speech. IPO Annual Progress Report 4: 58-60. Eindhoven

Sigurd B. 1971. Maximal rate and minimal duration of repeated syllables. Papers from the Institute of Linguistics of the University of Stockholm 3

Stetson R.H., Hudgins C.V. and Moses E.R. 1940. Palatograms change with rates of articulation. Archives Néerlandaises de Phonétique Expérimentale 16: 52-61

Wood S. 1973. Speech tempo. Working Papers of the Phonetics Laboratory of Lund University 9.

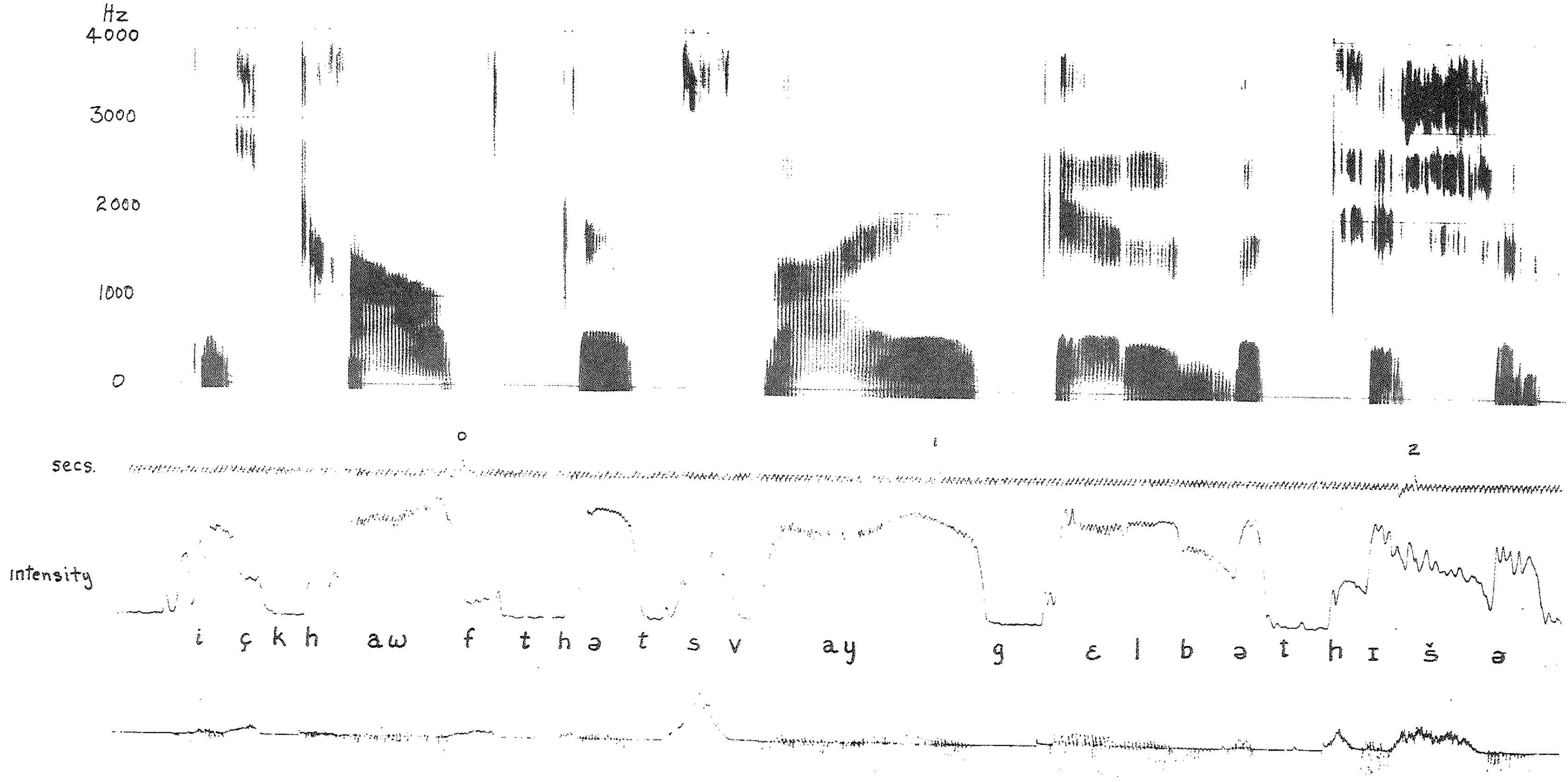

Fig.1a. A slow rendering (2.8 sylls/sec) of the sentence "Ich kaufte zwei gelbe Tische" (L).

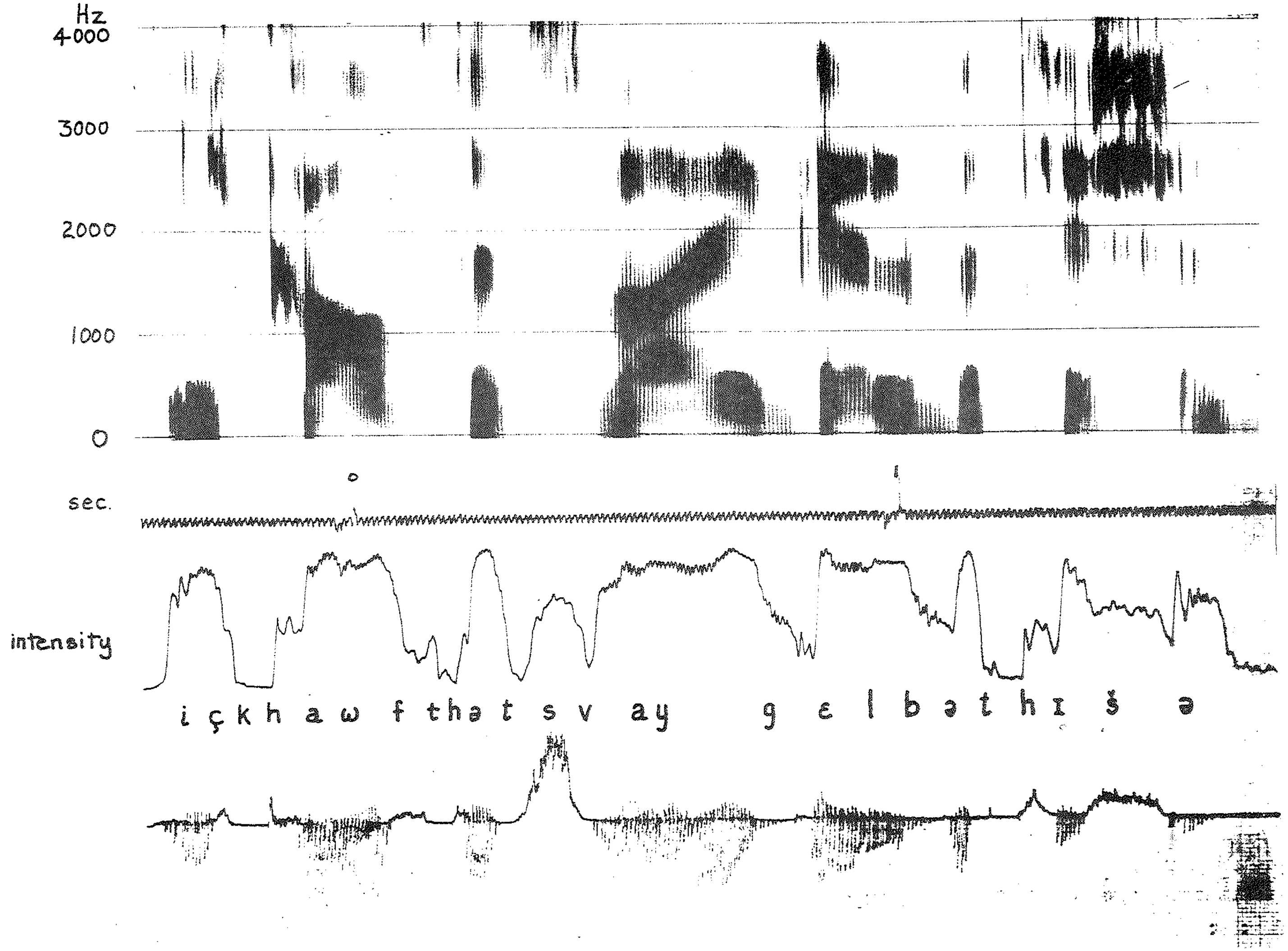

Fig.1b. A medium rendering (4.2 sylls/sec) of the sentence "Ich kaufte zwei gelbe Tische"(L).

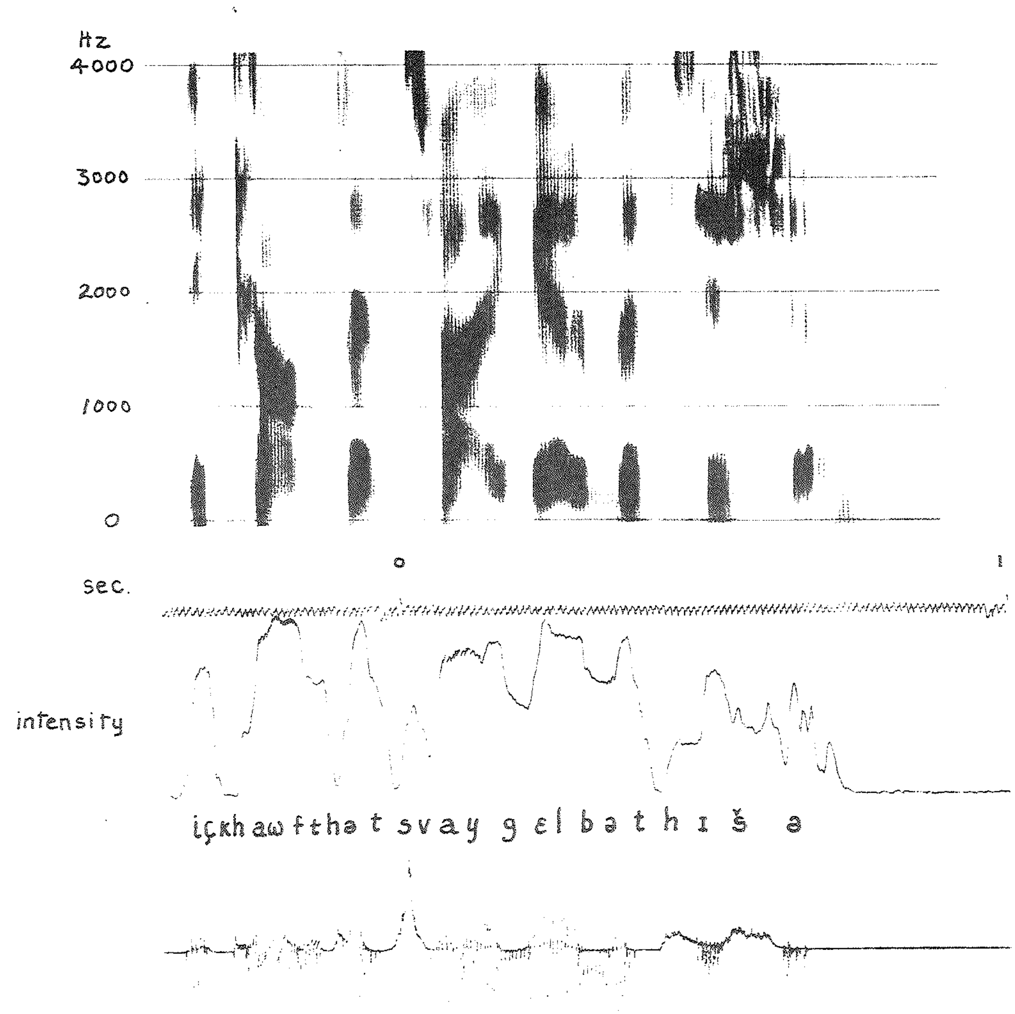

Fig.1c. A fast rendering (7.2 sylls/sec) of the sentence "Ich kaufte zwei gelbe Tische" (L).

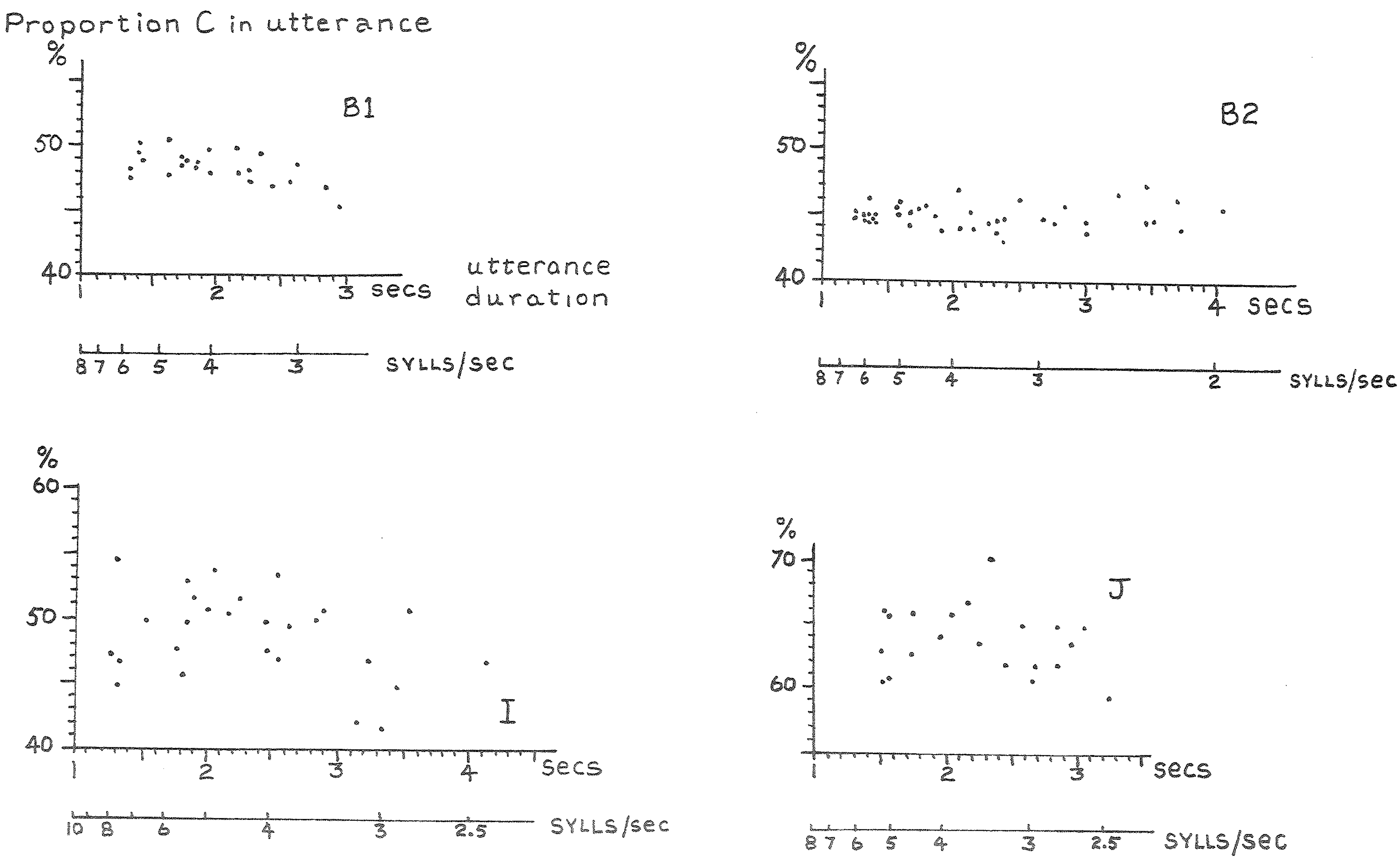

Fig. 2. The proportion of sentence duration consumed by the consonants in the sentence, for series B1, B2, I, J.

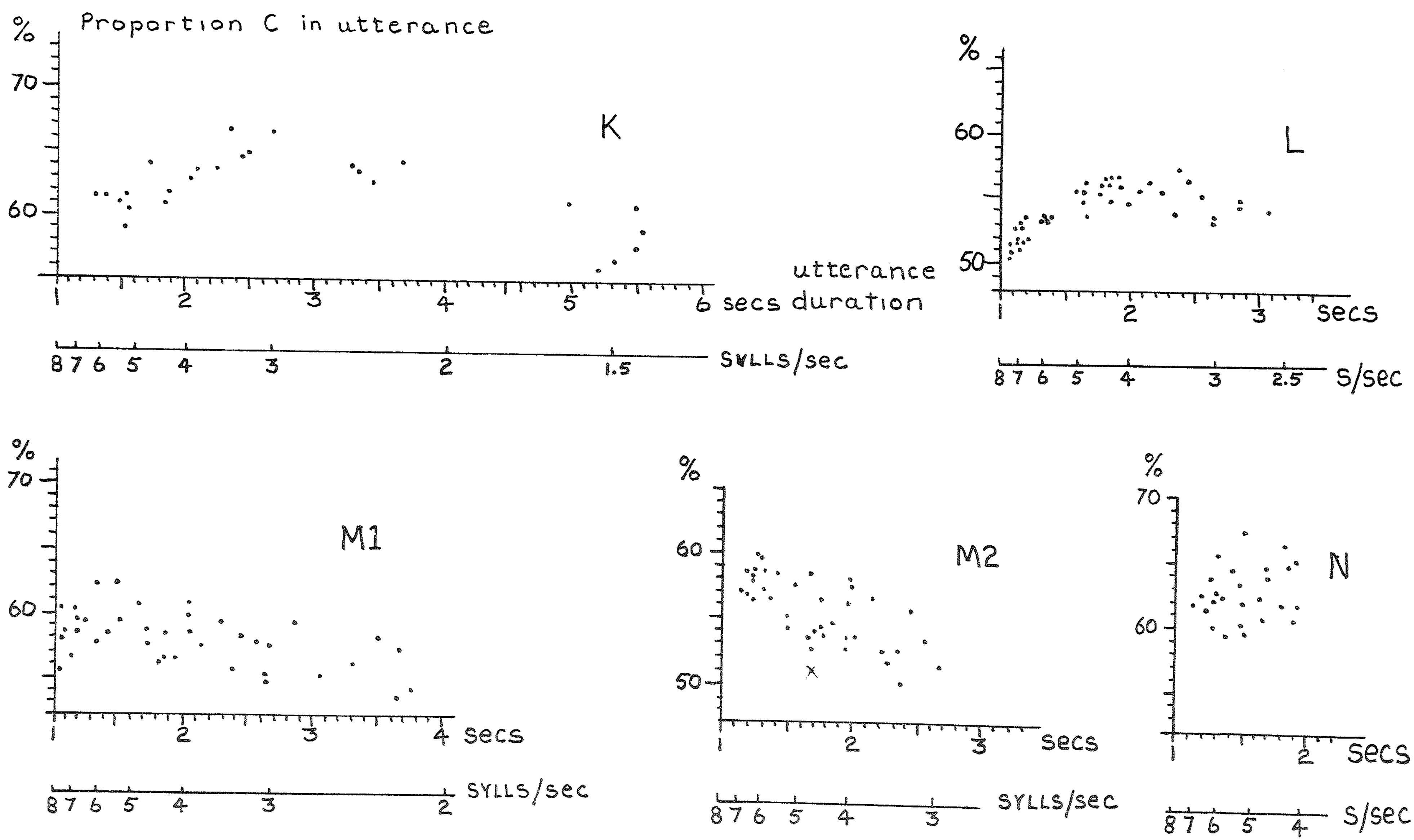

Fig. 3. The proportion of sentence duration consumed by the consonants in the sentence, for series K, L, M1, M2, N.

RATE CLASSES	CONSONANTAL PROPORTIONS (%) AT DIFFERENT RATES								
(Sylls/sec)	B1	B2	I	J	K	L	M1	M2	N
1.01–2.00					58.2				
2.01–3.00	45.5	44.7	45.8	58.7	61.9	54.8	56.3	53.0	
3.01–4.00	48.1	44.6	47.8	64.1	64.2	55.4	57.7	55.0	63.6
4.01–5.00	47.9	44.7	50.4	63.8	62.1	55.6	57.9	56.1	62.7
5.01–6.00	48.7	45.7	51.6	63.4	60.6	54.1	60.1	58.1	62.4
6.01–7.00		45.4				52.6	59.4		
7.01–8.00			48.1			51.6	57.8		

Fig. 4. The average consonantal proportions of the utterance duration for each speaker at different articulation rates.

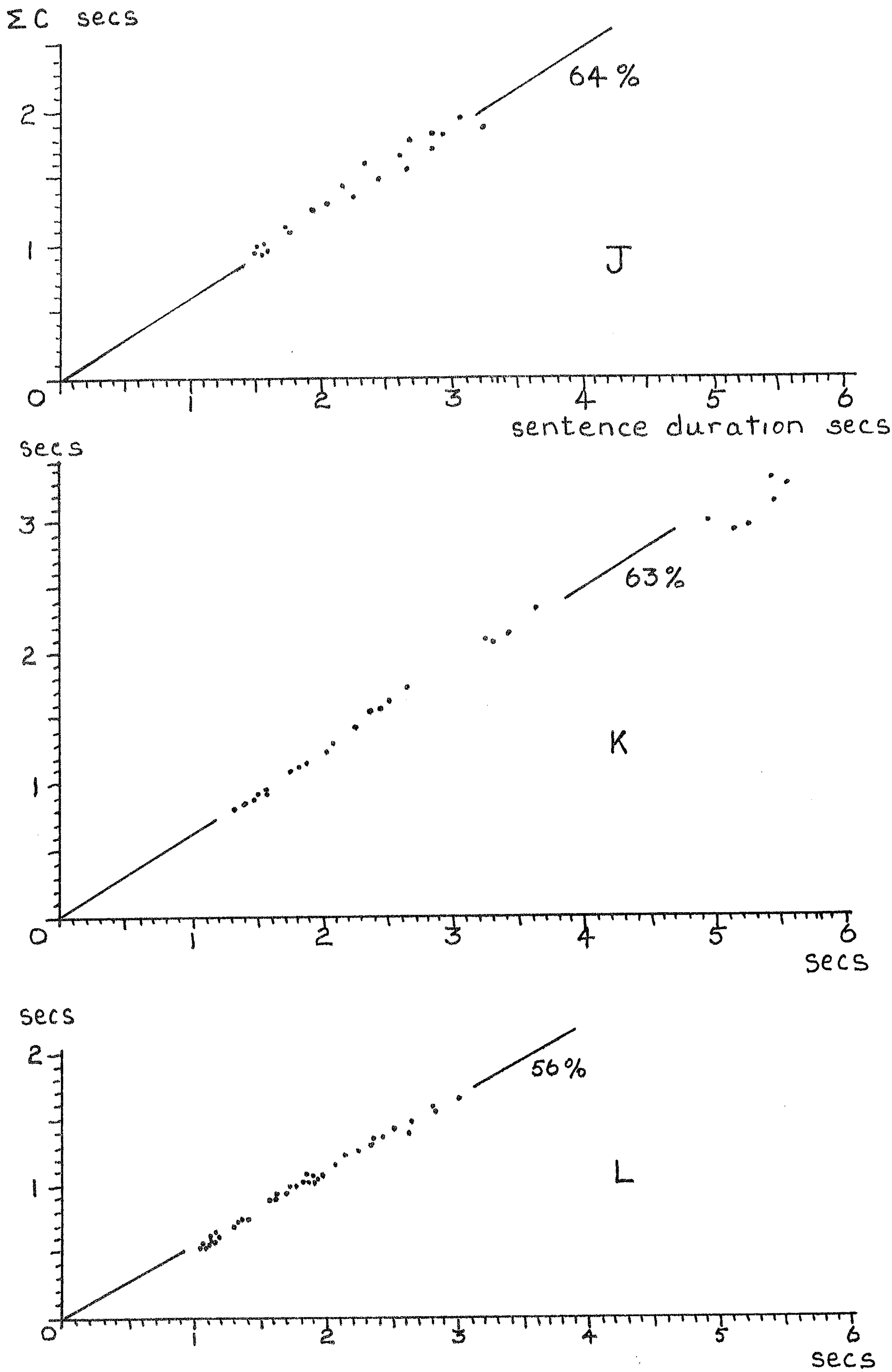

Fig. 6. The relationship between aggregated consonantal durations and sentence duratons in series J, K and L.

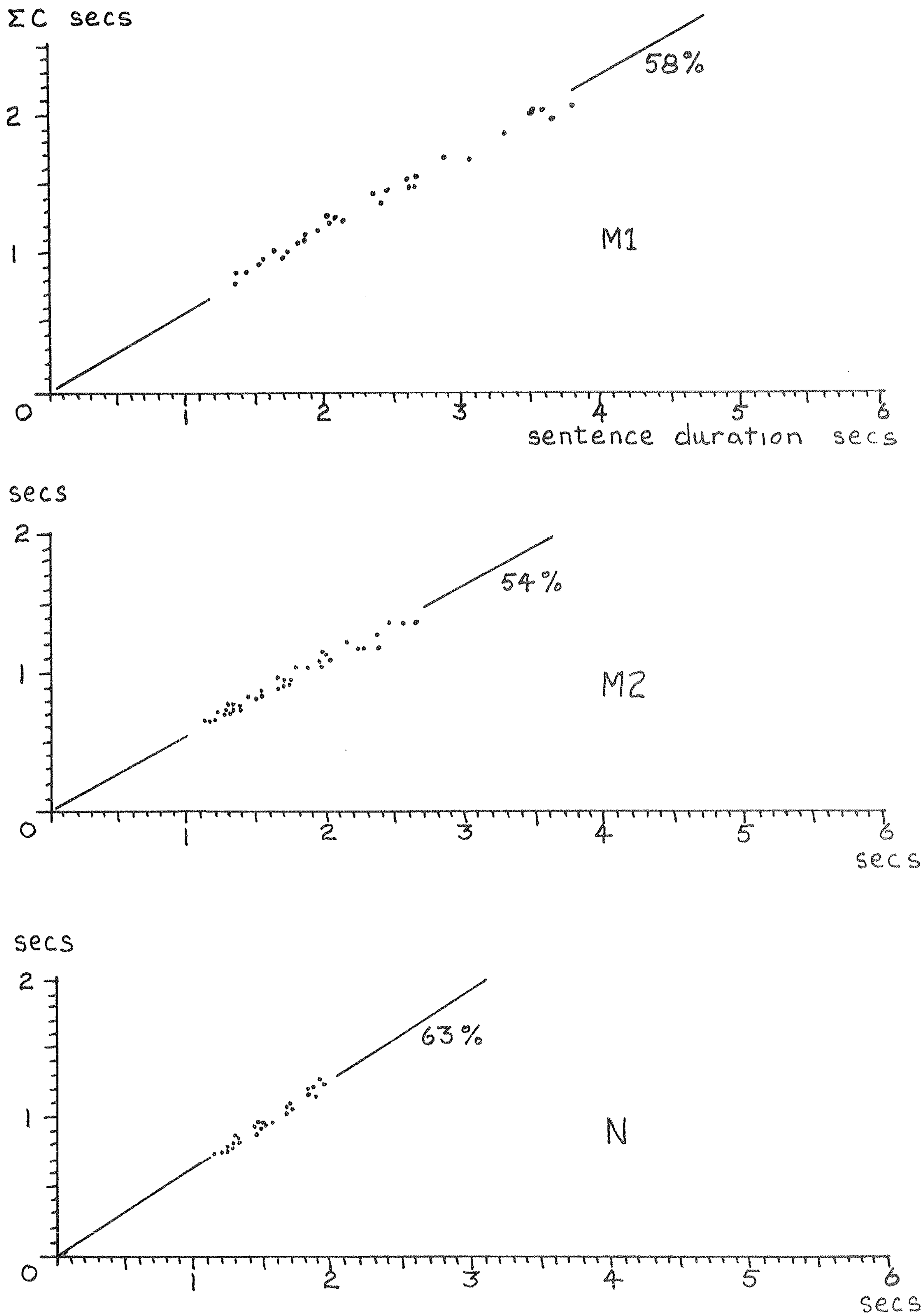

Fig. 7. The relationship between aggregated consonantal durations and sentence durations in series M1, M2 and N.

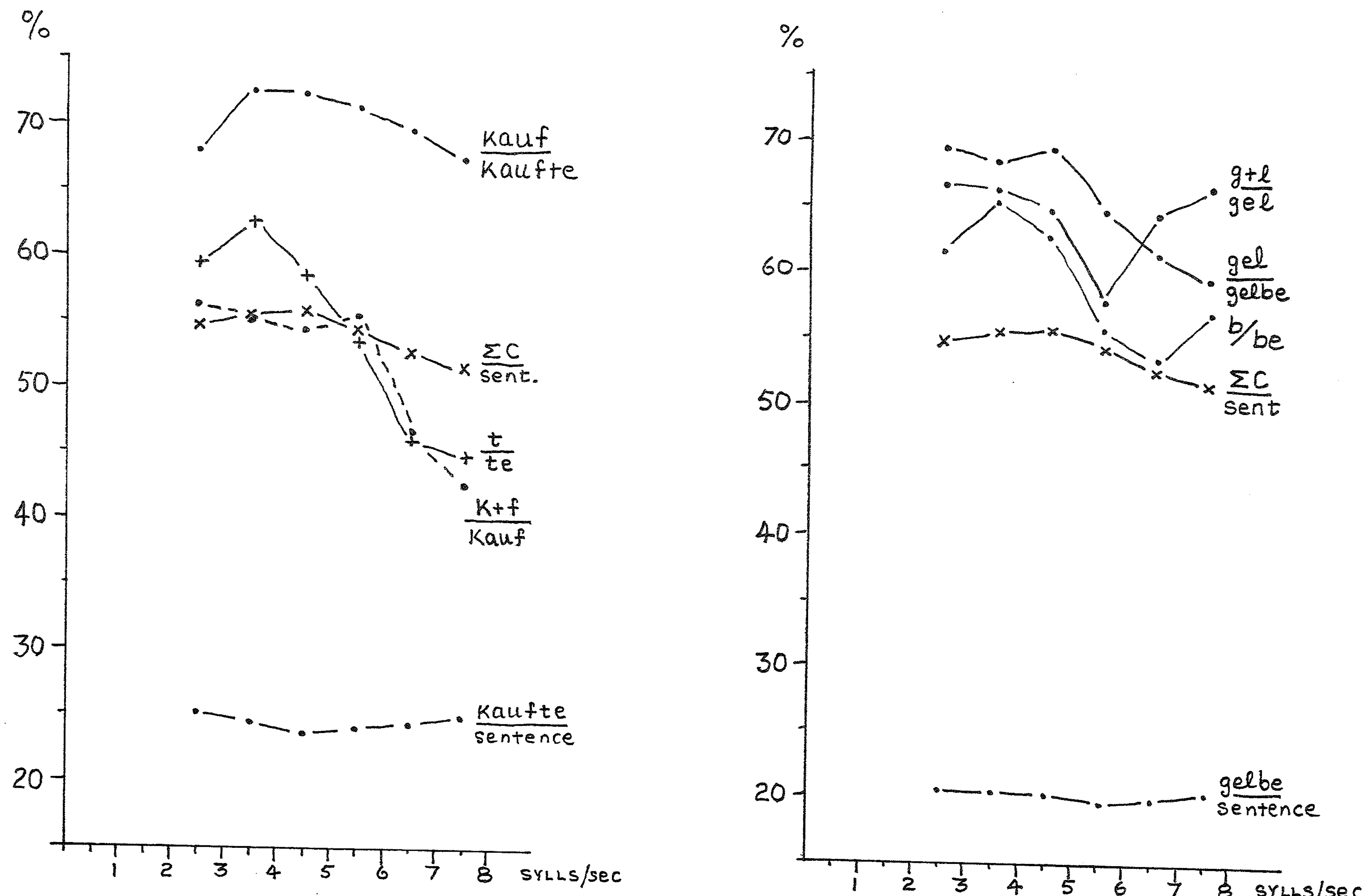

Fig. 9. The relative durations of selected words, syllables and consonants from L's series.

Speaker	Correlation coefficients for sentence duration and aggregated consonantal duration	aggregated vowel duration
B1	0.974	0.996
B2	0.998	0.998
I	0.967	0.982
J	0.984	0.965
K	0.954	0.960
L	0.982	0.994
M1	0.996	0.993
M2	0.985	0.986
N	0.981	0.942

Fig. 8. Coefficients for the correlations between sentence duration and the aggregated consonantal and vowel durations respectively in each rendering (cf. Figs. 5, 6 and 7).