DURATION: MEASUREMENTS, PHONOLOGICAL FUNCTIONS, THEORETICAL IMPLICATIONS

ILSE LEHISTE

Department of Linguistics
The Ohio State University
Columbus, OH 43210-1229, U.S.A.

ABSTRACT

The paper starts with a survey of the linguistic functions of duration. A specific case is then discussed in greater detail: the durational structure of Estonian disyllabic words with a three-way quantity contrast. Measurements show that the durations of the two syllables exhibit the following typical ratios for the three quantities: Quantity 1 - 2/3, Quantity 2 - 3/2, and Quantity 3 - 2/1. Listening tests, using white noise signals, were given to 28 English-speaking and 28 Estonian-speaking listeners. The results showed that both groups perceive duration ratios of 2/3 as distinct from duration ratios of 3/2 and 2/1, but that they do not use duration ratios to separate the latter two quantities.

DURATION: MEASUREMENTS, PHONOLOGICAL FUNCTIONS, THEORETICAL IMPLICATIONS

The general theme of this session is "The role of phonetics in linguistic theory". This is a vast topic; I doubt whether I could do justice to it in anything shorter than a book-length manuscript. There are two alternatives--to survey the field at a very general (and probably somewhat superficial) level--or to narrow the topic so it could be treated at reasonable depth. I have decided in favor of the second alternative, and after a very brief survey, I will treat one subtopic in somewhat greater detail. A considerable part of the past thirty years, I have studied the temporal organization of spoken language; but there is still very much left to discover, and it so happens that I also still have something new to say.

Let us start with a survey of the various functions of duration. I have treated this topic before in a fair number of publications, and I have summarized the results of my own work, and that of others, in a paper entitled "The many linguistic functions of duration" (Lehiste 1984). While new details have been added to our knowledge since that time, the general picture does not seem to have changed, at least not as far as I am aware, so I will summarize briefly the findings presented in that paper. References to work by other scholars, up to the year 1984, will be found in that publication; I have added some more recent references to the bibliography of the current paper.

As a start, I would classify the linguistic functions of duration as follows.

1. Duration serves to establish the identity of a segment.
2. Duration serves to specify the meaning of a word. According to structuralist terminology, this would be called phonemic duration or quantity.
3. In many languages, duration is a strong cue for stress and emphasis.
4. Duration serves to indicate the position of a linguistic unit within a higher-level linguistic unit:
   - the position of a segment within a syllable
   - the position of a syllable within a word
   - the position of a word within a phrase and a sentence
   - the position of a sentence within a unit of discourse
5. Duration functions in establishing boundaries between linguistic units.

When we talk about the role of duration in establishing the identity of a segment, we are dealing, first of all, with non-contrastive, intrinsic duration. It is well known that some sounds are longer than others, all other factors being kept constant. There are also language-specific durational phenomena at this level: subphonemic differences between sounds can serve to identify the sounds, even though such differences are not used independently for contrastive purposes. For example, in English duration serves as a strong perceptual cue distinguishing certain inherently long and short vowels, and duration of a syllable nucleus serves as a cue to the voicing or voicelessness of a postvocalic consonant.

There are also many languages in which duration can be independently contrastive at the segmental level; at least that is the traditional way of analyzing oppositions between long and short vowels and consonants in languages like Finnish. I believe, however, that in most cases contrastive segmental duration is further modified by durational patterns that apply at the next higher

levels--at least at the level of syllables and words. Contrastive segmental duration is integrated into durational patterns that apply at higher levels.

I will come back to that later in the current paper, when I present some new data. Let me mention just now that in Estonian polysyllabic words, the durational pattern is normally distributed over a disyllabic sequence. Thus the durations of the first and second vowel in minimal triples like sada - saada! - saada have an inverse relationship: lengthening of the first vowel is accompanied by shortening of the second vowel. The relationship between the durations of the two syllables appears to play a strong role in the listener's perception of the identity of the word; the durational pattern has the whole disyllabic sequence as its domain.

One result of the temporal relationships just referred to is to keep the duration of disyllabic words approximately constant--at least there is a clear tendency for maintaining something that might be called the "temporal integrity of the word". This means also that the duration of a syllable depends on the total number of syllables within the word: the tendency to keep the duration of the word close to some average level causes the syllables to become shorter when there is a larger number of them in the word.

Studies by Nooteboom (1972) and Lindblom and Rapp (1973) have shown for Dutch and Swedish respectively that duration of a stressed long vowel is longest in monosyllables and decreases systematically with the addition of further unstressed syllables. Tarnóczy showed that for Hungarian already in 1965 (Tarnóczy 1965). In a study which I published in 1975 (Lehiste 1975a) I showed that in English, a stressed syllable nucleus is longest in a monosyllabic word and shorter in polysyllabic words; I showed likewise that a stressed syllable is longer than an unstressed syllable in the same position, and that a syllable in final position is longer than the same syllable in non-final position.

It is thus clear that the position of a syllable within a word influences its relative duration. As part of the same study, I also found that the duration of a test word depends on the length of the frame in which it appears: test words were longest in the shortest frame, and shorter in two long frames used in the study. The way the duration of the test words interacted with the duration of the frames shows that the speakers integrate the test words into the utterance at the level at which the time program for the whole sentence is generated.

I got similar results in my first study of paragraph structure (Lehiste 1975b): sentences were longer when they were produced in isolation, and shorter when they were part of a paragraph--which indicates that the temporal planning extends to units larger than a single sentence. Furthermore, sentences in final position within a paragraph were longer than the same sentences when they occurred in initial or medial position within a paragraph.

This lengthening is part of a more general process which I have called pre-boundary lengthening. Pre-boundary lengthening is also used extensively to indicate the position of syntactic boundaries within a sentence. I have carried out several studies of pre-boundary lengthening, relating it to the rhythm of the spoken utterance. In this context, I would like to review briefly my first study dealing with the disambiguation of syntactic ambiguity (Lehiste 1973). In that study, listeners were able to identify correctly such sentences in which difference in meaning was correlated with a difference in surface syntactic bracketing. Successful disambiguation was achieved when the speakers had increased the interstress interval that contained the relevant boundary. The speakers had used several ways to achieve the same aim; the most straightforward one was the insertion of a pause, but equally successful were other means like the lengthening of one or more segmental sounds preceding the boundary, i.e. pre-boundary lengthening.

My further studies of the relationship between syntactic and rhythmic structure of English sentences lead to the postulation of a connection between rhythm and syntax that operates in the following way. Speech is a rhythmic activity, as are most motor activities performed by human beings. Stressed syllables carry the greatest amount of information; therefore, attention has to be focussed on the stressed syllables. This is facilitated by setting up an expectation as to when the next stressed syllable is likely to occur. Producing sentences in such a way that stressed syllables occur at regular intervals contributes to optimal perception by the listeners whose attention is cyclically directed to the points in time at which the stressed syllables can be expected to be found (Martin 1972, Cutler and Darwin 1981). Furthermore, a disruption of the expected pattern--namely, lengthening of an interstress interval--can be used to convey crucial information about syntactic structure: the placement of a syntactic boundary. At least in English, the syntactic structure of a sentence is thus to a considerable extent manifested in the timing pattern of that sentence when produced orally by a native speaker of the language. Timing appears to me to be primary; whatever other cues may be present, they play a less effective role.

I base this claim on a study in which it was shown that syntactic boundaries can be effectively recognized when the test sentences have been reduced to monotone, thus eliminating any possible contribution from fundamental frequency (Lehiste, Olive and Streeter 1976). In a later study (Lehiste 1983), I confirmed these results from the point of view of perception, and provided additional evidence from the point of view of production.

In this brief survey of the linguistic functions of duration, I have discussed the function of duration in establishing the identity of a segment; I have talked about contrastive duration at the word level; I have also discussed the ways in which duration functions to indicate the position of a linguistic unit within a higher-level linguistic unit, and the ways in which duration functions to establish boundaries between linguistic units. I have not talked much about duration as a cue to stress and emphasis-mainly because by now this function of duration appears to be generally known and accepted. (For a recent treatment of the topic, cf. Beckman 1986.) There may be other linguistic functions of duration that I have unintentionally overlooked. But I hope the general picture is clear: duration plays a part at a number of levels, and no linguistic description of a language is complete without reference to the function of duration within the system.

I would like to return now to a very specific case in which the role of duration appeared to me to need further study. This is the question of the three-way contrast between disyllabic word structures in Estonian--a topic that has considerable theoretical interest.

In my first extensive study of segmental and syllabic quantity in Estonian (Lehiste 1960), I made the observation that the factor that determined whether a disyllabic word was in quantity 1, 2 or 3 was word structure--more specifically, the ratio between the durations of the first and the second syllable. Listeners assigned the word to quantity 1, when speakers had produced the word in such a way that the ratio was approximately 2 to 3; the word was assigned to quantity 2, when the ratio was approximately 3 to 2, and to quantity 3, when the ratio approximated 2:1. The average durations of the syllables analyzed in the study were 106 and 151 msec for words in quantity 1, 295 and 187 for words in quantity 2, and 435 and 195 msec for words in quantity 3.

Duration of the first syllable is, of course, contrastive at the syllabic level; the average durations of the first syllable can be classified into the three categories of short, long, and overlong, as has been traditional in Estonian phonetics and phonology. And the durational differences are accompanied by different fundamental frequency patterns. All three factors are phonetically present; several linguists have been interested in determining the hierarchy of importance among these three factors, and much work has been done in the description of the three-way quantity opposition in Estonian. Let me mention here just the most recent papers by Arvo Eek and several other scholars associated with the Institute of Language and Literature of the Estonian Academy of Sciences and Tartu University (cf. Eek 1983, which contains these references).

In his very thorough study of 1983, Eek related the ratios between syllable durations to speech tempo and to fundamental frequency patterns within the disyllabic sequence. Basically, words were heard as being in quantity 1, when the ratio of the second vowel and the first vowel was equal to or larger than 1.2. The word was assigned to quantity 2, when the ratio V2:V1 was between 0.57 and 0.81, and to quantity 3, when the ratio was equal to or smaller than 0.43. Differences in tempo and in Fo played important roles. According to Eek's study, quantities 1 and 2 differ primarily in duration, since Q1 could be turned into Q2 and vice versa by manipulation of duration alone. Additional phonetic features are required for the perception of Q3.

I had a problem with the ratios described by Eek: they are presented as having a fairly large range of values, and these values appeared too precise and too complex. Already in 1960, I had described the ratios in terms of simple numbers: 2:3, 3:2, 2:1. It seemed intuitively obvious to me that contrastive structures would be based on simple notions; and it appears that there is some experimental support to this idea. I would like to summarize now a paper by Dirk-Jan Povel entitled "Internal representation of simple temporal patterns" (Povel 1981).

Povel started from a study by Fraisse (1946), who had discovered a remarkable phenomenon in the production and perception of durations. Fraisse found that subjects who were asked to produce temporal patterns by tapping basically used only two durations; the longer duration was typically approximately twice as long as the shorter duration, with a ratio of 2:1. Povel investigated the limitations present in the perception of temporal sequences by having subjects imitate sequences of 150-msec beeps whose onset intervals were varied in a systematic fashion. The duration ratios of the intervals between beeps were relationships numerically expressible as 1:4, 1:3, 2:5, 1:2, 3:5, 2:3, 3:4, and 4:5. (Note that Povel always presented the shorter duration first, resulting in ratios smaller than 1.) The results of two experiments yielded the finding that the only duration ratio that was correctly reproduced was 1:2 (i.e. .50). The errors in production were systematic: there was a tendency toward the 1:2 interval ratio, so that smaller ratios were increased and larger ratios were made smaller. For example, a ratio of .40 was reproduced as .45, and a ratio of .66 was reproduced as .49. Under certain special conditions set up for a third experiment, subjects were also able to imitate interval relations of 1:3 and 1:4 accurately in the contexts used in the experiment.

Povel carried through his experiments at Indiana University, and his subjects were presumably native speakers of American

English. It is legitimate to ask whether similar results would be obtained, if the subjects were speakers of a language in which duration plays a contrastive role. The experiments which I am about to report were carried out to test precisely this question: are there any differences in the perception of durational ratios that are correlated with the linguistic use of duration in the native language of the subjects being tested. The experiments were carried out in collaboration with Dr. Robert Fox at the Ohio State University,. and will be described in more detail in a joint publication (Fox and Lehiste, in preparation).

Let us recall here that in Estonian, there exist sets of three minimally contrastive disyllabic words, consisting of the same segmental sounds. One measured characteristic of such words is the durational ratio between the two syllables, which I had already in 1960 observed and formulated as ratios 2:3 for words in Q1, 3:2 for words in Q2, and 2:1 for words in Q3. A considerable literature has grown up in the meantime; the work of Eek is particularly significant in this context (cf. Eek 1983). There is no doubt that measurements do not yield very precise ratios, and that there is a certain amount of variation to be found under different speech conditions. Other phonetic factors are likewise present in spoken utterances, such as the duration of the first syllable nucleus itself (relative to some possible average internal standard) and the fundamental frequency contour applied to the disyllabic sequence. Which of these phonetic factors is contrastive needs to be established by means of listening tests--measurements alone are not enough.

The theoretical interest of the problem is at least two-fold. There have been linguistic schools that claim that all linguistic oppositions are binary; sounds can be short or long, there are no three-way durational contrasts. Fraisse's and Povel's findings seem to support this point of view. If speakers of a language with a three-way quantity opposition likewise can only identify durational ratios of 1:2, the three-way opposition must be manifested by other means. If, however, speakers of such a language can identify additional durational ratios, especially such that occur in their native language, then it is true that the native language of a subject influences his performance in psychoacoustic tests.

Our experimental procedures differed considerably from those used by Povel, since we did not just want to replicate his experiment, but wanted to use stimuli that could be directly related to Estonian disyllabic word patterns. We used pairs of noise bursts with controlled durations. The ratios that we employed were those found in Estonian disyllabic words: 2:1, 3:2, 2:3, and 1:2, numerically equal to 2.0, 1.5, .66, and .5. In Povel's study, the first temporal interval was always shorter than the second; he seems to have assumed that the ratios 2:1 and 1:2 are perceived in the same fashion, but since in Estonian, quantities 1 and 2 contrast, having the ratios 2:3 and 3:2 respectively, we felt that this assumption would not be justified. The ratio 1:2 was included for symmetry's sake, even though it is not regularly found in Estonian disyllabic words.

Each experimental trial consisted of presenting two such paired signals, separated by very short pause. The subjects were asked to state whether the duration ratio of the first noise sequence was the same as or different from the durational ratio of the second noise sequence. There was a 500-msec pause between each experimental trial.

We wanted to ensure that subjects were comparing duration ratios and not, for example, the durations of the first noise burst in each sequence. (This would be comparable to assigning an Estonian disyllabic word to a quantity category on the basis of the duration of the first syllable.) With this concern in mind, a second factor was introduced in constructing the experimental tokens: overall duration of the noise sequences. In particular, in half of the noise sequences, the duration of noise bursts 1 and noise burst 2 summed to 350 msec. In the other half of the sequences, the noise burst summed to 450 msec. Sequences with the same overall duration (LONG-LONG or SHORT-SHORT) alternated, in random order, with sequences of different overall durations (SHORT-LONG or LONG-SHORT). It was hoped that in this way duration ratio differences among the experimental trials would not be confounded with noise burst duration differences. These factors were explicitly discussed in the instructions, and examples of duration ratio differences vs. overall duration differences were included at the start of the stimulus tape. Equal numbers of "sames" and "differents" were included in the experimental tape, and subjects were also informed of this fact so that their responses would not be skewed in one direction or another. Guessing was encouraged. In the actual administration of the test, subjects were asked to encircle the appropriate letter (standing for "same" or "different") on the test sheet.

The test was first administered to 28 subjects at The Ohio State University in Columbus. These subjects were native speakers of English, with minimal exposure to languages in which duration plays a contrastive role. The same test, using identical tapes, but appropriately translated instruction and test sheets, was administered to 28 subjects in Tallinn. (The help of colleagues Arvo Eek, Mati Hint, and Kullo Vende in carrying out the tests is gratefully acknowledged.) These subjects had Estonian as their native language, and they were tested in Estonian. The analysis of the responses was carried out in Columbus. Detailed results will be presented in a separate publication (Fox and Lehiste, in preparation); below are some preliminary results.

The results are presented in the form of five tables. The first four tables have the same structure. The ratio of the first sequence is indicated on the vertical axis, the ratio of the second sequence on the horizontal axis; the numbers in the cells of each matrix represent the percentage of "SAME" responses. Table 1 gives the responses of English listeners to stimuli in which both sequences had equal durations.

Table 1

Percentage of "SAME" responses given by English-speaking listeners to pairs of stimuli consisting of a sequence of two noise bursts. Ratio of the noise burst durations of the first sequence is indicated on the vertical axis, ratio of the stimuli of the second sequence on the horizontal axis. Both sequences had equal total duration.

		Ratio of second sequence 1:2	2:3	3:2	2:1
Ratio	1:2	93.9	77.7	18.3	11.6
of	2:3	78.6	92.3	33.0	17.9
first	3:2	18.3	14.7	93.6	91.1
sequence	2:1	14.2	15.2	89.2	93.3

Table 2 presents the same information for English subjects reacting to stimuli in which the two sequences had different durations.

Table 2

Percentage of "SAME" responses given by English-speaking listeners to pairs of stimuli consisting of a sequence of two noise bursts. Ratio of the noise burst durations of the first sequence is indicated on the vertical axis, ratio of the stimuli of the second sequence on the horizontal axis. The sequences differed in duration.

		Ratio of second sequence 1:2	2:3	3:2	2:1
Ratio	1:2	63.8	50.9	15.2	17.4
of	2:3	51.3	60.7	22.3	13.8
first	3:2	12.1	16.5	69.0	62.1
sequence	2:1	11.6	10.7	52.2	75.4

Table 3 shows the responses of Estonian listeners to stimuli in which both sequences had equal durations; Table 3 thus corresponds to Table 1.

Table 3

Percentage of "SAME" responses given by Estonian-speaking listeners to pairs of stimuli consisting of a sequence of two noise bursts. Ratio of the noise burst durations of the first sequence is indicated on the vertical axis, ratio of the stimuli of the second sequence on the horizontal axis. Both sequences had equal total duration.

		Ratio of second sequence 1:2	2:3	3:2	2:1
Ratio	1:2	96.7	89.7	10.7	9.8
of	2:3	85.7	95.5	18.8	8.9
first	3:2	6.7	12.9	97.6	93.3
sequence	2:1	5.8	6.3	93.3	98.7

Table 4 presents the responses of Estonian listeners in cases in which the sequences differed in duration; Table 4 thus corresponds to Table 2.

Table 4

Percentage of "SAME" responses given by Estonian-speaking listeners to pairs of stimuli consisting of a sequence of two noise bursts. Ratio of the noise burst durations of the first sequence is indicated on the vertical axis, ratio of the stimuli of the second sequence on the horizontal axis. The sequences differed in duration.

		Ratio of second sequence 1:2	2:3	3:2	2:1
Ratio	1:2	61.6	52.2	9.8	10.3
of	2:3	45.5	53.7	12.9	8.9
first	3:2	8.9	14.2	68.1	53.5
sequence	2:1	7.6	6.3	52.2	73.8

Let us compare first Tables 1 and 3 with Tables 2 and 4. The cells starting at the top on the left and descending diagonally show the identification as "SAME" of signals in which the ratios were in fact identical (e.g. cases in which the first sequence and the second sequence both had ratios of 1:2). Correct recognition was evidently more difficult in cases when the sequences differed in duration: the percentages in the cells constituting the diagonal are considerably lower in Tables 2 and 4, also indicating, among other things, that the two groups of listeners reacted to the differences in overall sequence duration in the same general fashion.

The results presented in Tables 2 and 4 reflect listeners' reactions to ratios in cases in which the overall duration of the stimuli provided a conflicting cue: they were identifying sequences as "same" in spite of the fact that overall durations were clearly different. On the basis of durations alone, the listeners should have identified all the stimuli serving as basis for Tables 2 and 4 as "different"; and if they had been simply guessing, the scores would have been close to 50%. It is obvious that in many cases, listeners were able to identify ratios correctly even when the signals differed in duration; the statistical significance of these results will be discussed in detail in the forthcoming publication referred to earlier (Fox and Lehiste, in preparation).

Let us look now at the four tables from the point of view of successful discrimination between the four ratios. Here the results are likewise quite clear: the listeners, both English-speaking and Estonian-speaking, recognized only two contrastive patterns--sequences that had a first element that was longer than the second element, and sequences that had a first element that was shorter than the second element. This result emerges from the fact that ratios 1:2 and 2:3 are not distinguished from each other, the same being true for ratios 3:2 and 2:1. The percentage of "correct positive" decisions is somewhat higher than the percentage of "incorrect positive" decisions, but the difference appears not to be statistically significant.

What about the difference between the linguistic backgrounds of the two groups of listeners? Table 5 provides some information that is relevant in the present context.

Table 5

Average percentages of "SAME" responses given by English-speaking and Estonian-speaking listeners to pairs of stimuli consisting of a sequence of two noise bursts. "Correct positive" refers to cases in which duration ratios that were actually identical were identified as "SAME". "Incorrect positive" refers to cases in which duration ratios of 1:2 and 2:3 on the one hand, and 3:2 and 2:1 on the other hand, were identified as "SAME". "Wrong" refers to cases in which ratios of 2:1 and 2:3, or 1:2 and 3:2, were identified as "SAME".

	Average "correct" positive"	Average "incorrect positive"	Average "wrong"
English-speaking	80.3	69.1	16.4
Estonian-speaking	80.6	70.7	9.9

This table presents average percentages, calculated on the basis of the data presented in Tables 1-4. Average "correct positive" decision refers to cases in which, for example, the ratios of 2:1 and 2:1 were identified as "SAME". "Incorrect positive" refers to cases in which, e.g., the ratios 2:1 and 3:2 were identified as "SAME". Average "wrong" decision gives the percentage of "SAME" decisions involving pairs of opposite durational ratios (e.g. 2:1 and 2:3). And it is here that a difference between English-speaking and Estonian-speaking listeners emerges: the Estonian-speaking listeners appear less likely to call such ratios "same". The difference of 6.5 percentage points is in fact significant--and it is the only significant difference between the two groups of listeners.

Let us return now to the theoretical questions that were raised at the beginning of the paper. The listeners seem in fact to have been capable of distinguishing between shorter and longer signals, and to have been able to decide whether the first or the second member of a sequence was longer. Under the conditions of this experiment, the listeners did not distinguish between the ratios 1:2 and 2:3 on the one hand, and the ratios 3:2 and 2:1 on the other hand. The linguistic background of the listeners did not have any effect on this aspect of the outcome; but Estonian listeners were much less likely to confuse the ordering of longer or shorter elements within a sequence than were English-speaking listeners.

From the point of view of Estonian prosody, the following conclusions may be drawn. The results clearly show that words in Q 1, with a duration ratio of 2:3, are perceived as distinct from words in quantities 2 and 3, with duration ratios 3:2 and 2:1. These two long quantities, however, are not distinguished on the basis of duration ratio. Since under normal conditions listeners do indeed recognize the difference between words in quantities 2 and 3, other phonetic factors must provide the decisive information. Fundamental frequency contours are the most likely candidate, but further research may bring new information and new ideas. The present experiment suggests that Estonian should rightfully be considered an accent language, in which other phonetic factors besides durational ones play a significant role. The experiment also demonstrates that phonetics does indeed provide crucial information that must be taken into account when questions of linguistic theory are to receive satisfactory solution.

BIBLIOGRAPHY

Mary E. Beckman (1986), Stress and Non-Stress Accent. Foris Publications: Dordrecht, Holland/Riverton-U.S.A.

R. Carlson and B. Granström (1986), "A search for durational rules in a real-speech data base". Phonetica 43:140-154.

Elizabeth Couper-Kuhlen (1986), An Introduction to English Prosody. Max Niemeyer Verlag: Tübingen.

Anne Cutler and Christopher J. Darwin (1981), "Phoneme-monitoring reaction time and preceding prosody: effects of stop closure duration and fundamental frequency". Perception and Psychophysics 29:217-24.

Arvo Eek (1983), "Kvantiteet ja rohk eesti keeles (1)". Keel ja Kirjandus 26, 9:481-489, 10:549-559.

Robert A. Fox and Ilse Lehiste (in preparation),"Perception of duration ratios".

P. Fraisse (1946), "Contribution a l'étude du rythme en tant que forme temporelle". Journal de psychologie normale et pathologique 39:283-304.

Ilse Lehiste (1960), "Segmental and syllabic quantity in Estonian". American Studies in Uralic Linguistics, Bloomington, pp. 21-82.

Ilse Lehiste (1973), "Phonetic disambiguation of syntactic ambiguity". Glossa 7, 2:107-122.

Ilse Lehiste (1975a), "Some factors affecting the duration of syllable nuclei in English". Salzburger Beiträge zur Linguistik 1:81-104.

Ilse Lehiste (1975b), "The phonetic structure of paragraphs". A. Cohen and S. G. Nooteboom (eds.), Structure and Process in Speech Perception. Springer-Verlag: Berlin- Heidelberg-New York. Pp. 195-206.

Ilse Lehiste (1983), "Signalling of syntactic structure in whispered speech". Folia Linguistica 17, 1-2: 239-245.

Ilse Lehiste (1984), "The many linguistic functions of duration". James E. Copeland (ed.), New Directions in Linguistics and Semiotics. Rice University Studies, Houston, Texas. Pp. 96-122.

Ilse Lehiste, Joseph P. Olive, and Lynn A. Streeter (1976), "Role of duration in disambiguating syntactically ambiguous sentences". Journal of the Acoustical Society of America 60:1199-1202.

Björn Lindblom and Karin Rapp (1973), "Some temporal regularities of spoken Swedish". Papers from the Institute of Linguistics 21, University of Stockholm.

J. Martin (1972), "Rhythmic (hierarchical) vs. serial structure in speech and other behavior". Psychological Review 79:487-509.

S. G. Nooteboom (1972), Production and Perception of Vowel Duration. Doctoral Dissertation, University of Utrecht.

Dirk-Jan Povel (1981), "Internal representation of simple temporal patterns". Journal of Experimental Psychology 7, 1:3-18.

T. Tarnóczy (1965), "Can the problem of automatic speech recognition be solved by analysis alone?" Rapports du 5e Congrès International d'Acoustique, Vol. II, Conférences générales. Liége: D. E. Commins. Pp. 371-387.