TEMPORAL COMPENSATION IN A QUANTITY LANGUAGE

ILSE LEHISTE

This paper* explores the temporal relationships between the segments in three sets of Estonian words. While a considerable amount of interest has been shown previously in the duration of sounds with contrastive function, the temporal structure of the whole word of which the contrastive sounds constitute a part has received relatively little attention.[1] It is the thesis of this paper that the word is programmed as a whole, and that significant relationships exist among all segments that constitute a word, although not all segments participate in segmental quantity oppositions.[2]

The material analyzed for the study consists of three minimal triples: *vaga* 'pious' (nom. sing.), *vaka* 'bushel' (gen. sing.), *vakka* 'bushel' (part. sing.); *sada* '100' (nom. sing.), *saada* 'send' (2. sing. imperative), *saada* 'send' (-da-inf.); *saag* 'saw' (nom. sing.), *saak* 'prey' (nom. sing.), *sakk* 'sawtooth' (nom. sing.). According to traditional analyses, the intervocalic consonant /k/ is in short, long and overlong quantities in the first set (short, long and overlong will be referred to as quantities 1, 2 and 3); in the second set, the contrastive sound is the vowel of the first syllable, which appears in quantities 1, 2, and 3; and in the third set, /a/ is in quantity 3 in the first two words and in quantity 1 in the third, while final /k/ is in quantity 1 in the first word and in quantity 3 in the second and third.[3] These words were recorded by two speakers, each of whom repeated each test word between 100 and 110 times in sequence. The recordings were made in Tallinn in the autumn of 1970.[4] The tapes were processed through a Frøkjær-Jensen transpitchmeter and intensity meter; the curves were displayed by means of an Elema-Schönander Mingograf (at a speed of 10 cm/second). Measurements of duration were made from mingograph traces;

* This research was supported, in part, by PHS Research Grant No. 1 RO3 MH18122-01 from the National Institute of Mental Health, and in part by Grant No. 534.1 from the National Science Foundation to the Computer and Information Science Research Center, The Ohio State University.

[1] The problem is surveyed, and literature cited, in Lehiste, 1970.

[2] For a discussion of the problem, cf. Lehiste, 1971.

[3] The words are given in standard spelling. The letter *g* stands for a voiceless lenis plosive, which is the realization of /k/ in quantity 1. Traditional spelling does not distinguish between long and overlong vowels, both of which are written with two vowel letters.

[4] I would like to thank my informants for the generous contribution of their time, and the researchers at the Institute for Language and Literature and the Laboratory of Experimental Phonetics of

the results were analyzed statistically by means of an IBM 360 computer.[5] To normalize for variations in tempo, the average durations of all words were computed, and a subset of 50 utterances whose durations were closest to the mean duration was extracted for each word. Further computations were performed on these subsets. This procedure of tempo normalization is essentially the same as that employed by Ohala and by Kozhevnikov and Chistovich in previous temporal studies.[6]

One form of temporal compensation within Estonian words has been frequently referred to in previous descriptions. This is the compensation in the duration of the vowel of the second syllable, which adjusts itself inversely to the duration of the first syllable, so that a first syllable in quantity 1 is followed by a so-called half-long vowel in the second syllable, and first syllables in quantity 2 and 3 are followed by successively shorter second syllable vowels. Evidence for this type of compensation, which is part of the phonological structure of Estonian words, is given in Table 1 and in Figures 1, 2, and 3.

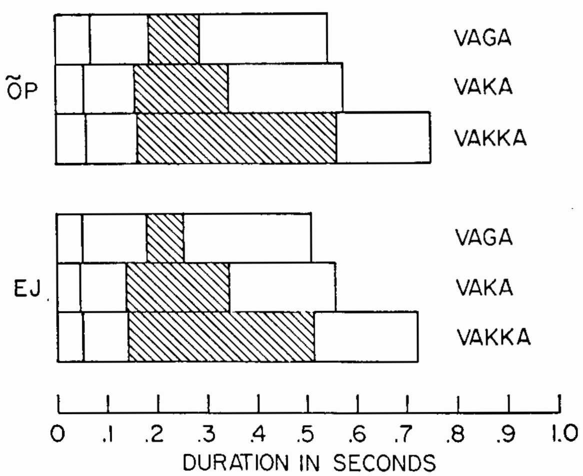

Fig. 1. Average durations of segments in the three words *vaga*, *vaka* and *vakka*, produced by two informants.

Figure 1 shows the average durations of segments in productions of *vaga-vaka-vakka* by the two speakers. As may be seen, there is some adjustment in the duration of the vowel of the second syllable, which partly compensates for the increasing duration of the intervocalic consonant. A similar observation may be made with respect to the set *sada-saada* (2) — *saada* (3), displayed on Figure 2. Here the duration of the vowel of the second syllable is inversely correlated with the vowel of the first syllable, whose duration is contrastive.[7]

the Academy of Sciences of the Estonian S.S.R. for their cooperation and assistance in making the recordings.

5 The analysis techniques are described in detail in Shockey, Gregorski, and Lehiste 1971.

6 Ohala 1970, Kozhevnikov and Chistovich 1965.

7 There is some controversy over the question whether a first syllable in quantity 2 is followed

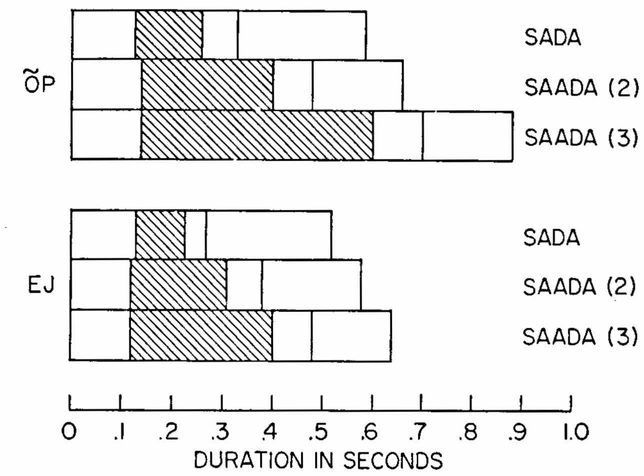

Fig. 2. Average durations of segments in the three words *sada*, *saada* (2) and *saada* (3), produced by two informants.

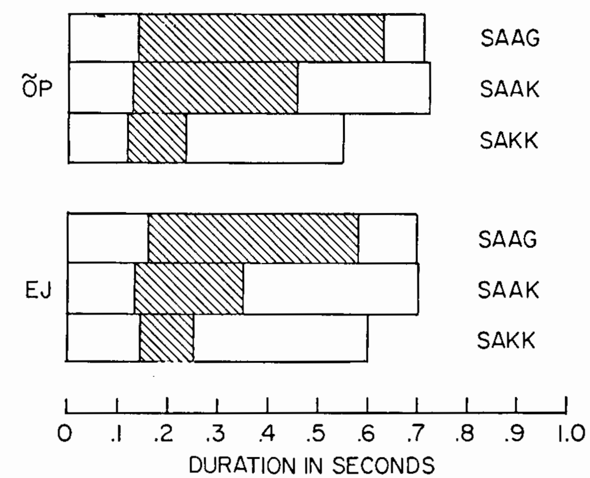

Fig. 3. Average durations of segments in the three words *saag*, *saak* and *sakk*, produced by two informants.

Figure 3 shows the set *saag-saak-sakk*. Here the compensation is between the vowel and the final consonant, both of which are contrastive on the segmental level. It is interesting that the total durations of the words *saag* and *saak* are practically identical: the compensation is complete, and the two monosyllabic words really differ only in the distribution of duration among the two contrastive segments. The third member of the set, *sakk*, contains what is commonly analyzed as a quantity 1 vowel and a quantity 3 final consonant. In terms of measurable durations, this adds up to some-

by a half-long vowel or not. In the present set of data, one of the speakers had successively shorter second-syllable vowels in *vaga-vaka-vakka*, the other in *sada-saada-saada*.

what less than the 3 + 1 sequence in *saag*, where a vowel in quantity 3 is followed by a consonant in quantity 1. As far as the word *saak* is concerned, the assignment of the vowel and the final consonant to phonemic quantities remains ambiguous on phonetic grounds. For both speakers, the duration of /a/ in *saak* is longer than that of /a/ in *saada* (2), but shorter than /a/ in *saada* (3); the duration of /k/ in *saak* is likewise between the durations of /k/ in *vaka* and *vakka*. The pertinent data are given in Table 1.

TABLE 1

Mean durations (in milliseconds) of segments in nine test words produced by two speakers. N = 50.

Word and speaker	C_1	V_1	C_2	V_2
ÕP *vaga*	71.34	120.74	98.68	257.26
ÕP *vaka*	57.24	103.54	188.82	223.84
ÕP *vakka*	57.86	105.54	397.26	187.68
EJ *vaga*	51.40	128.22	71.94	257.24
EJ *vaka*	45.06	94.54	204.20	210.28
EJ *vakka*	50.18	92.42	376.32	203.02
ÕP *sada*	131.30	128.38	73.56	251.98
ÕP *saada* (2)	136.72	255.88	76.96	190.68
ÕP *saada* (3)	139.30	454.56	100.80	185.40
EJ *sada*	130.96	101.62	53.84	232.16
EJ *saada* (2)	117.88	191.46	72.36	196.44
EJ *saada* (3)	122.94	275.96	78.10	165.76
ÕP *saag*	141.14	486.20	85.96	
ÕP *saak*	136.78	316.46	271.00	
ÕP *sakk*	119.86	115.98	316.70	
EJ *saag*	158.16	419.82	118.40	
EJ *saak*	131.72	222.64	351.14	
EJ *sakk*	143.68	104.82	350.80	

The temporal compensation with which we are primarily concerned in the current paper is of a different kind. It is manifested not in the pattern itself, but in its realization. We hypothesize that there exists a temporal program for the production of an utterance. At a certain level in the process of the production of the utterance, the sequence of articulatory gestures is programmed, and the utterance is assigned an overall basic duration. If this is true, then repeated productions of the same utterance will aim at a duration close to the average for a series of productions. In order that this may be accomplished, temporal adjustment will take place between successive segments during a single production: if one of the segments is produced with a duration that is longer than its own average, another segment within the same utterance will be relatively shorter than its respective average, so that the duration of the word as a whole will remain more or less constant, i.e., vary as little as possible from the average duration programmed for the word. Each segment will, of course, have some variability, which may be statistically expressed in terms of variance. If the segments were independent of each other, their variances would be additive and the variance of the whole word would be the sum of the variances of the segments. If, however, there is temporal compensation among the segments constituting the word, the variance of the word should be less than the sum of the variances of the segments.

TABLE 2

Mean durations (in milliseconds) and variances of nine test words produced by two speakers. N = 50.

Word and speaker	Mean duration	Sum of variances of segments	Variance of word
ÕP *vaga*	548.02	527.33	118.38
ÕP *vaka*	573.44	582.72	84.06
ÕP *vakka*	748.34	1345.85	487.63
EJ *vaga*	508.80	796.56	199.75
EJ *vaka*	554.08	598.35	132.31
EJ *vakka*	721.94	1138.20	193.13
ÕP *sada*	585.22	452.00	238.56
ÕP *saada* (2)	660.24	896.20	161.56
ÕP *saada* (3)	880.06	1257.32	305.63
EJ *sada*	518.58	610.43	203.75
EJ *saada* (2)	578.14	390.80	70.88
EJ *saada* (3)	642.76	751.75	213.50
ÕP *saag*	713.30	601.10	297.19
ÕP *saak*	724.24	592.22	173.06
ÕP *sakk*	552.54	621.51	202.13
EJ *saag*	696.38	1753.70	655.06
EJ *saak*	705.50	1196.46	264.25
EJ *sakk*	599.30	1005.59	330.63

Table 2 contains the mean durations of each test word, the sum of variances of the segments, and the variance of the word taken as a whole. Figure 4 presents the same data graphically for the *vaga-vaka-vakka* set. As is obvious from the table and the figure, temporal compensation is indeed present in all test words, and in general the hypothesis appears to be validated. The study was continued to establish the statistical significance of correlations between all subsets of segments in each test word. A summary of the results is presented in Table 3 and in Figures 5-7.

Figure 5 shows the correlation cœfficients (Pearson correlations) for all segments contained within the words *vaga*, *vaka* and *vakka* produced by the two speakers. Specifically, these correlations show the relationship of the first three segments to the fourth. (Correlations between various other combinations of segments are given in

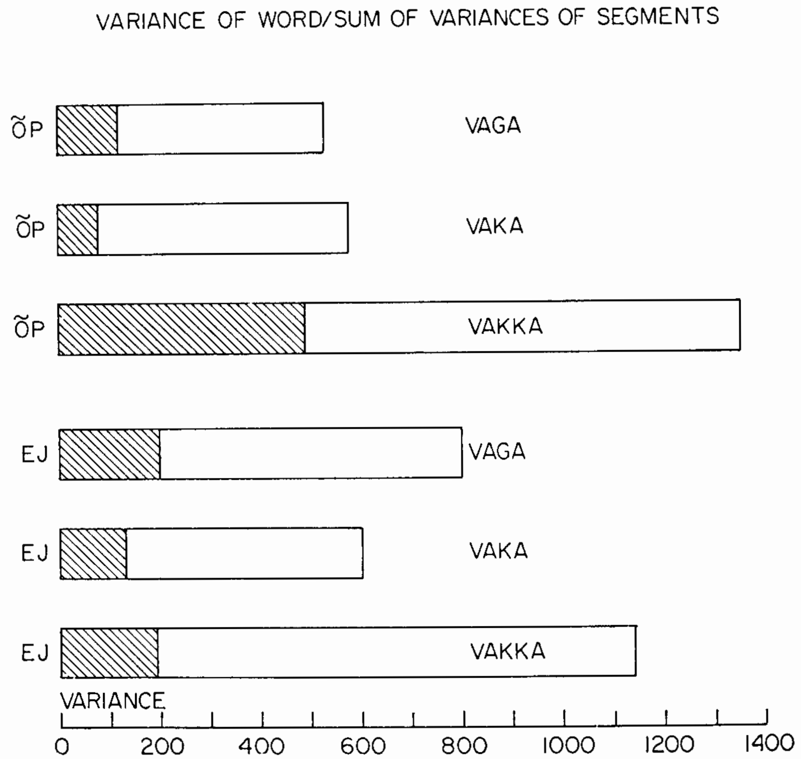

Fig. 4. Variance of the word (diagonally hatched) and sum of variances of segments in productions of the words *vaga*, *vaka* and *vakka* by two speakers. Variance of word is superimposed on the sum of variances of segments.

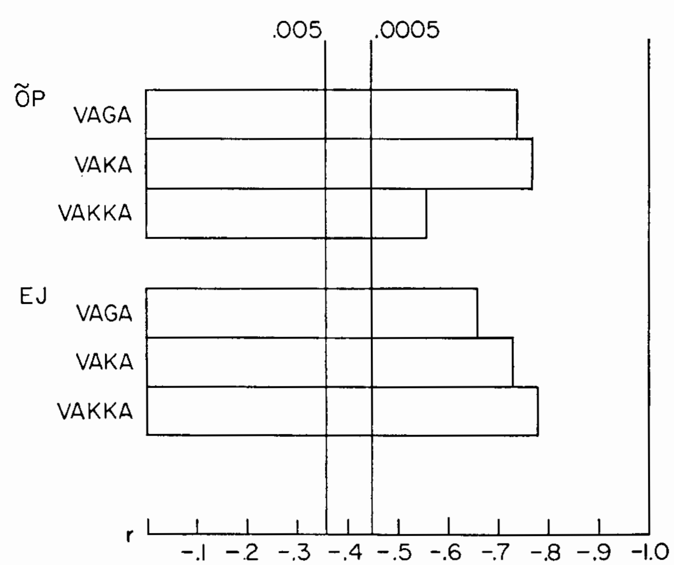

Fig. 5. Correlation coefficients (r) between the first three segments and the fourth segment contained in the words *vaga*, *vaka* and *vakka* produced by two speakers.

TABLE 3

Correlation coefficients between various combinations of segments in productions of nine test words by two speakers.

$$N = 50;\ r = \frac{1}{N}\sum\left(\frac{X-\bar{X}}{\sigma X}\right)\left(\frac{Y-\bar{Y}}{\sigma Y}\right).$$

Significance of r at .235 — .95, at .279 — .99, at .361 — .995, and at .451 — .9995.

Word	Segments involved in the correlation	Correlation coefficient	
		Speaker ÕP	Speaker EJ
vaga	1, 2	–0.178	–0.468
	2, 3	–0.142	–0.367
	3, 4	–0.353	–0.386
	1, 2, 3, 4	–0.738	–0.660
vaka	1, 2	–0.310	–0.148
	2, 3	–0.046	–0.333
	3, 4	–0.652	–0.523
	1, 2, 3, 4	–0.770	–0.730
vakka	1, 2	0.058	–0.035
	2, 3	–0.186	–0.219
	3, 4	–0.509	–0.717
	1, 2, 3, 4	–0.556	–0.776
sada	1, 2	–0.236	–0.170
	2, 3	–0.404	–0.428
	3, 4	0.166	–0.391
	1, 2, 3, 4	–0.530	–0.543
saada (2)	1, 2	0.121	–0.381
	2, 3	–0.328	–0.314
	3, 4	0.104	–0.116
	1, 2, 3, 4	–0.834	–0.765
saada (3)	1, 2	–0.467	–0.288
	2, 3	–0.382	–0.609
	3, 4	–0.139	0.028
	1, 2, 3, 4	–0.667	–0.668
saag	1, 2	–0.155	–0.191
	2, 3	–0.604	–0.529
	1, 2, 3	–0.637	–0.590
saak	1, 2	–0.257	–0.312
	2, 3	–0.475	–0.205
	1, 2, 3	–0.668	–0.768
sakk	1, 2	–0.100	0.112
	2, 3	–0.406	–0.326
	1, 2, 3	–0.708	–0.685

Table 3.) As may be seen, the degree of negative correlation is extremely high. The two vertical lines on the figure represent r values that show significance at the .005 and .0005 level respectively; the actually obtained correlations are significant at an even higher level.

Figure 6 presents correlation coefficients for the words *sada*, *saada* (2) and *saada* (3). In this case, the displayed negative correlations were found to obtain between the two syllables — segments 1 and 2 on the one hand, and 3 and 4 on the other hand. As before, the correlations are highly significant.

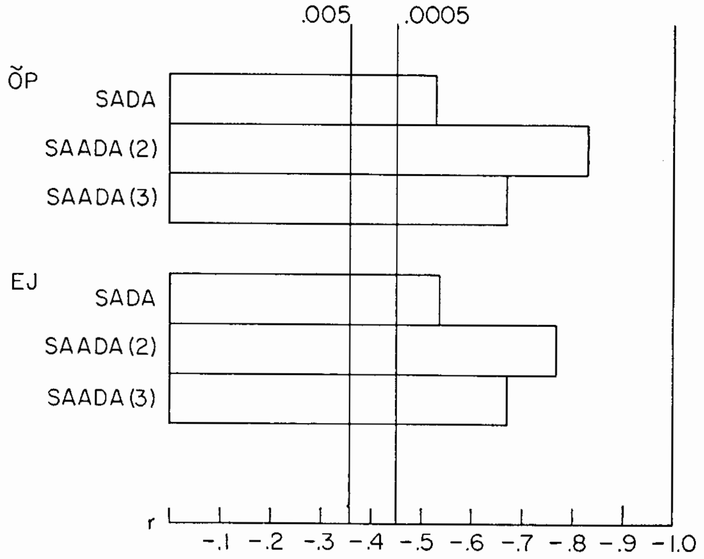

Fig. 6. Correlation coefficients (*r*) between the first two and last two segments contained within the words *sada*, *saada* (2) and *saada* (3) produced by two speakers.

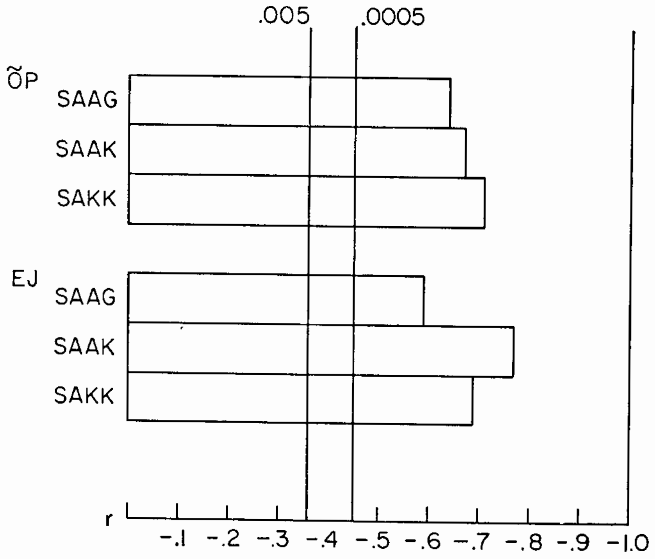

Fig. 7. Correlation coefficients (*r*) between the first two segments and the third segment contained within the words *saag*, *saak* and *sakk* produced by two speakers.

Figure 7 presents similar data for the monosyllabic words *saag*, *saak* and *sakk*. Here the first consonant and vowel have been correlated with the final consonant. Again, the degree of negative correlation is highly significant.

Not all combinations of segments yielded equally high negative correlations. In most cases, correlations involving the initial consonant and other parts of the word were either significant at a lower level or not significant at all. This may reflect the fact that the duration of the initial consonant is non-contrastive at the segmental level. However, combinations that involved all segments yielded significant negative correlations in all cases.

The hypothesis at the beginning of this paper was that words are programmed as units, and that significant relationships exist among all segments that constitute a word. The results of the study have clarified these relationships: the durations of segments constituting a word are negatively correlated, and the level of significance of these negative correlations is much too high to be attributed to chance. Since the timing patterns extend over the whole word, it may be concluded that words do indeed constitute units of programming. Further research is needed to establish to what an extent these patterns are modified when the word becomes part of a higher-level unit such as a phrase or sentence.

Department of Linguistics
Ohio State University
Columbus

REFERENCES

Kozhevnikov, V.A. and L.A. Chistovich
1965 *Speech: Articulation and Perception*, trans. J.P.R.S. (= *JPRS* 30, 543) (Washington, D.C.) (Moscow, Leningrad).

Lehiste, I.
1970 *Suprasegmentals* (Cambridge, Mass., M.I.T. Press).
1971 "Temporal Organization of Spoken Language", in *Form and Substance, Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen*, L.L. Hammerich, R. Jakobson, E. Zwirner, eds. (Copenhagen, Akademisk Forlag) pp. 159-169.

Ohala, J.
1970 *Aspects of the Control and Production of Speech* (= *UCLA Working Papers in Phonetics* 15) (Los Angeles).

Shockey, L., R. Gregorski, and I. Lehiste
1971 *Word Unit Temporal Compensation* (= *Ohio State University Working Papers in Linguistics* 9) (Columbus).

DISCUSSION

BREND (Ann. Arbor, Mich.)

The last two sentences of your paper answered my major question, regarding the relationships of your findings with the occurrence of words in higher-level phonological units. However, I would also be interested to know if you envision any studies regarding tempo of words with regard to their grammatical function.

LEHISTE
I have completed another study, dealing with the temporal effects of morpheme boundaries and syntactic boundaries. While a review of this study would require too much time, I can state the general results as follows: the temporal structure gives no clue as to any possible differential effect of the two kinds of boundaries.

OHALA (Berkeley, Calif.)
There may be another factor which would lead to a negative correlation between adjacent segments in a given word or syllable, namely, an error in the placement of the boundary between the two segments. This error may be due to two sources: (1) human error or sloppiness (I know Professor Lehiste is careful in the treatment of data, however, so perhaps this is not a significant factor in this case), (2) uncertainty in picking a boundary point that is identical to the boundary point used by the brain. If one suggests that the brain compensates for temporal deviations it follows that the brain must know when a given segment begins and ends. (Even if there is no temporal compensation accomplished by the brain, but instead there exists something approaching the 'chain' model, the brain must still know when a given segment ends in order to know when to execute the next segment.) If the boundary point used by the experimenter does not coincide with that used by the brain then the first segment may be too long and as a result the second too short, and vice-versa. Thus a negative correlation may appear that is in no way indicative of temporal compensation.

LEHISTE
I am fully aware of the difficulties involved in making duration measurements. The test material was selected with relative ease of segmentation in mind. If one uses the same criteria for establishing a boundary each time (e.g., the plosive release or the onset of voicing), errors of measurement will be reduced to human error in drawing a boundary line to coincide with this pre-determined segmentation point. Now each line was drawn on the basis of an independent decision; I did not first establish the beginning and end of the word and then determine the duration of each segment within the word boundaries. I am inclined to believe that errors in making each measurement independently will cancel each other out; it is not immediately obvious to me that they cause temporal compensation.

It is true that we do not know whether the segmentation points established on the basis of major changes in the acoustic signal correspond to the reference points used by the brain. This is a separate hypothesis; I am presently trying to find ways to test it. However, I believe that some fixed reference point must exist, since a large number of languages make linguistic use of length distinctions. If length oppositions

are part of a phonological system, and if observed differences in the duration of speech sounds are correlated with differences in linguistic function, it appears plausible that the measurements have a reasonable claim for perceptual reality.

GREGG (Vancouver)
Is it likely that this temporal compensation is valid not only for languages with clear-cut length oppositions but for others i.e., might it possibly be a language universal?

LEHISTE
I agree that higher-level temporal programming is likely to be a universal characteristic of spoken language. Some work has been done in our laboratory on temporal compensation in English, yielding similar results.