https://doi.org/10.3176/lu.1990.4.05

ARVO EEK (Tallinn)

UNITS OF TEMPORAL ORGANISATION AND WORD ACCENTS IN ESTONIAN *

1. Units of temporal organisation

Although a complex hierarchical framework for prosody has been recognised, there is actually no generally accepted treatment of prosodic structure. Discussing the differences between a grammatical and a phonological hierarchy Anthony Fox (1986 : 104—105) concludes: «the phonological hierarchy is not just a set of formal and arbitrary relationships of an abstract kind, but is ultimately determined by physiological constraints on the articulatory process itself. Units such as the 'syllable' or the 'foot' are not merely categories of phonological organisation but have their roots in the neural or physiological aspects of speaking. This may be the reason for the absence of such disturbances as 'level-skipping' or 'back-looping': the phonetic status of the hierarchy precludes the kind of 'liberties' that are taken in grammatical structure».

Regardless of the multifunctionality of durational phenomena (about the linguistic functions of duration see, e. g., Lehiste 1984; 1987) in the present paper we shall use duration measurements to decide whether there is some kind of relevant units acting in a hierarchical manner. As far as meaning is transferred from the speaker to the hearer by means of a linear sequence of language units, the production and reception of speech are inevitably related to the time parameter. However, there is no psycholinguistically unambiguous answer to the question of the repertory of mental language units used by the speaker. It has remained unclear whether in the synthesis from meaning to text, i. e. in forming the predicative structure of sentences, speakers add affixes to selected lexemes, or whether they only order the memorised «ready» wordforms, or whether they do both (see, e. g., Help 1985; Viitso 1988). We have some evidence supporting both the generative character of morphological synthesis as well as the view that speakers use «ready» wordforms (see Table 1).

But it is not known to what degree, if at all, the temporal organisation of the speech flow is dependent on mental language units. Presumably there must be a one-to-one correspondence between at least some kind of units of an abstract language and a concrete speech production level. A full non-correspondence between the units of different levels would hardly facilitate the process of speech perception.

* This paper was presented on a seminar at the Royal Institute of Technology, Department of Speech Communication and Music Acoustics (Stockholm, March 27, 1990).

*Table 1**

What are the minimal morphological units represented in one's mental lexicon (in the speaker's mind)?

Deep structure **Surface structure**

(a) KALA + Sg + Nom; Sg → Ø, Nom → Ø: → *kala*
KALA + Sg + Gen; Sg → Ø, Gen → Ø: → *kala*
KALA + Sg + Part; Sg → Ø, Part → Ø: → *kala*
KALA + Sg + Illat; Sg → Ø, Illat → *sse*: → *kalasse*
KALA + Sg + Iness; Sg → Ø, Iness → *s*: → *kalas*
KALA + Sg + Elat; Sg → Ø, Elat → *st*: → *kalast*
KALA + Sg + Allat; Sg → Ø, Allat → *le*: → *kalale*
KALA + Sg + Adess; Sg → Ø, Adess → *l*: → *kalal*
KALA + Sg + Ablat; Sg → Ø, Ablat → *lt*: → *kalalt*
KALA + Sg + Translat; Sg → Ø, Translat → *ks*: → *kalaks*
KALA + Sg + Essive; Sg → Ø, Essive → *na*: → *kalana*
KALA + Sg + Terminat; Sg → Ø, Terminat → *ni*: → *kalani*
KALA + Sg + Abessive; Sg → Ø, Abessive → *ta*: → *kalata*
KALA + Sg + Komitat; Sg → Ø, Komitat → *ga*: → *kalaga*
(b) SUPP, SUPI + Sg + Nom; Sg → Ø, Nom → Ø: → *supp*
+ Sg + Illat; Sg → Ø, Illat → *sse*: → *supisse*
etc.
TUPP, TUPE + Sg + Illat; Sg → Ø, Illat → sse: → *tupesse*
etc.
KUPP, KUPU + Sg + Illat; Sg → Ø, Illat → *sse*: → *kupusse*
(c) KÄSI, KÄE, KÄT- + Sg + Nom; Sg → Ø, Nom → Ø: → *käsi*
+ Sg + Part; Sg → Ø, Part → *t*: → *kätt*
+ Sg + Illat; Sg → Ø, Illat → *sse*: → *käesse*

Historical sound changes: *käsi* < **käti*
käe < *käen* < **käťen* < **käten*
kätt < **kättä*

Articulated morphemes or words do not consist of a simple concatenating string of phonemes, and sentences are not a simple sequence of wordforms. It is easy to demonstrate an indispensable hierarchical organisation of units in speech flow since it is impossible to synthesise understandable wordforms from an adequately ordered sequence of sustained speech sounds, and to get prosodically natural and unambiguously comprehensible sentences from the sequence of wordforms uttered in isolation.

To a certain extent, articulatory, phonatory and respiratory gestures are independent of each other and as such can be autonomously controlled. Various inherent characteristics of speech production mechanisms as well as their subjection to separate control permits one to surmise that

---

* Comments to Table 1:
(a) Presumably speakers have learned to use certain rules of morphological operation for the generation of necessary wordforms anew in every concrete case of synthesis, selecting suitable root and affix morphemes from a memorised lexicon. We have lots of evidence about the so-called active morphology from the language acquisition period in early childhood (e. g. *joos* as *käis* pro *jõi*).
(b) The basic meaning of the root morphemes of nouns may be represented in the memory by two variants enabling one to identify the stem vowel. The quality of the stem vowel cannot always be derived from the constituent phonemes of a root morpheme.
(c) It is highly probable that irregular (inflexional) variants of roots are memorised bunchwise. It is hard to believe that a chain of regular sound changes, taken place in language history, would behave as an instruction for the generation of wordforms also in the present day.
Of course, indirect data being scanty, various morphological conceptions developed by linguists need not coincide with the real processes going on in the speaker's mind.

viewed from these different aspects speech flow can be divided into different coherent segments which may be called phonetic frames (cf. the phrasal frame of Fujimura 1981) within which, among other things, durational phenomena could also be regulated. It is assumed that the duration of the elements of a phonetic frame is regulated according to the cohesive power (isochrony) of the frame, but so that the temporal compression, or the restriction of lengthening, need not be the same for every constituent part of the frame.

The units of a language system are traditionally expressed by means of phoneme sequences whose articulation is realised by continuous alternations of the narrowing (i. e. consonantal) and the widening (i. e. vocalic) gestures in the vocal tract. The minimal (elementary) articulatory gesture seems to be a demisyllable (Fujimura 1981; 1987; 1988; Eek, Help 1987) where the movement from a constriction or closed phase to an open phase or vice versa makes up a sonority cycle (see, e. g., Fujimura, Lovins 1978; Clements 1988). The demisyllabic unit is coherent in the sense that a demisyllabic command scans all muscle channels exploited in the generation of the respective minimal articulatory sequence and switches on simultaneously those channels whose activity is not contradicting the basic state of the movement. This multidimensionality causes difficulties of segmentation on the articulation level as articulators are subjected to separate motor control and their different structure and mass cause differences in their moving speed. The inherent duration of a segment is nothing but the result of different movements. The segmentation of an acoustic signal is a little simpler, as the interruptions of the continuous speech flow (e. g., implosion/ /explosion) serving as fixed points for the measuring of temporal structure are discriminated more easily. However, in order to understand durational phenomena one has to turn to articulation as the essential basis of speech. The demisyllabic CV- or -VC elementary (transitional) gesture is a relatively non-compressible segment of speech flow and is not subject to considerable durational changes (see, e. g., an iceberg model: Fujimura 1982). Only the quasi-stationary part of vocalic and consonantal phases of CV- and -VC demisyllables respectively is subject to durational variation that may be caused by general or local change of speech rate or by stressedness. In many languages, including Estonian, the variability of the quasi-stationary part is used to mark the phonological short/long opposition. The allophonic variants generally arise from demisyllabic affiliation. The demisyllable is the minimal segmental unit of articulation.

What I have been saying is wholly connected with the coordination of the movements of articulators. As any motor activity speaking must also be rhythmical as muscles cannot be kept working continuously at a constant tension level. The articulation of speech flow can be regarded as a succession of energy impulses whose minimal (integral) unit is the foot and whose general shape is physiologically determined by an alternation of the tension (more energy) and the relaxation (less energy) phases (for more detail see Eek, Help 1987; see Fig. 1). The total amount of physiological energy spent on a foot by expiratory, phonatory and articulatory activities may be called stress. Stress is not considered as some special reinforcing amount of energy added to a certain independently defined unit of speech flow. Stress itself is intrinsically segmented into units rather than appears something like an energetical increment of stressed syllables relative to unstressed syllables. The foot is the domain of accentuation, i. e. the purposeful variation in the amount of stress.

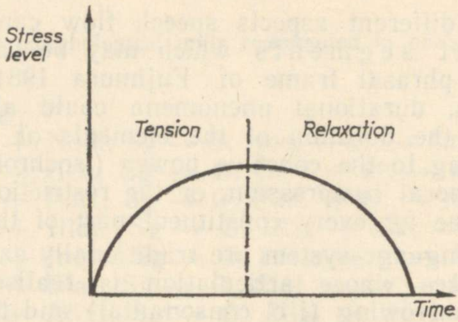

Fig. 1. A model of a foot.

What is the sequence of articulation units characteristic of the tension and relaxation phases of a foot? The simplest case is where the phase is formed by consequent demisyllables (e. g. *ka-na*, see Fig. 2). Here the demisyllable coincides with a syllable. The situation gets more complicated, however, if there are intersyllable consonant clusters (e. g. *kan-da; kant-ti*). Here the tension phase of a foot consists of two elementary units of articulation making up a coherent CV+VC sonority cycle. In acoustics and perception, there are clearly identifiable intra-foot units characterised by separate temporal intensity envelopes which have been defined as syllables. The s y l l a b l e is a relatively homogeneous sequence of demisyllabic unit(s) delimited by opening gestures of the vocal tract. There are many examples, however, where certain segmental units are left out of the syllable as just defined, e. g. in *kant-s s* is an appendage or affix of the syllable (cf. Fujimura, Lovins 1978; Eek 1987 : 208). Consequently the minimal articulatory units can be divided into demisyllables and affixes. A syllable cannot be called an independent unit outside an energetic minimal unit that is called the foot. A syllable is always a coherent element of a foot phase. A stress unit (foot), not a syllable, is the structural unit within which to seek phonotactic generalisations (e. g. the restricted occurrence of vowels in a relaxation phase, traditionally in an unstressed syllable). A foot may consist of one, two or three syllables.

The well-known division of speech rhythm into two types — stress-timing and syllable-timing — cannot be employed in the classification of all languages. We have proposed the model of speech rhythm in which 'stress' and 'foot' are considered as basic notions rather than 'stress' and 'syllable'. As for the basic non-timing nature of the distinction, the actual temporal rhythm pattern of a language may be viewed as deriving from two essentially nontemporal appearance of feet. Whether a language is stress-timed or not, depends on the interrelationship between feet and stress beats, i. e. on the frequency in which feet appear in continuous speech flow in the accentuated (reinforced) form displaying such stress beats. Whether a language is syllable-timed or not, depends on the interrelationship between feet and syllables, i. e. on the manner in which syllables associate with the internal structure of feet. As regards the fuzzy boundary between the rhythm types in cross-linguistic research, the foot perspective entails a much more complicated picture of factor interplay than a strict two-fold opposition between stress-timing and syllable-timing languages (Eek, Help 1987).

Some durational as well as qualitative phenomena accompanying these types probably indicate different exactness of the articulation commands in the foot phases. Detailed modifications of a stress impulse may be controlled by two kinds of language-specific hypothetical energetic

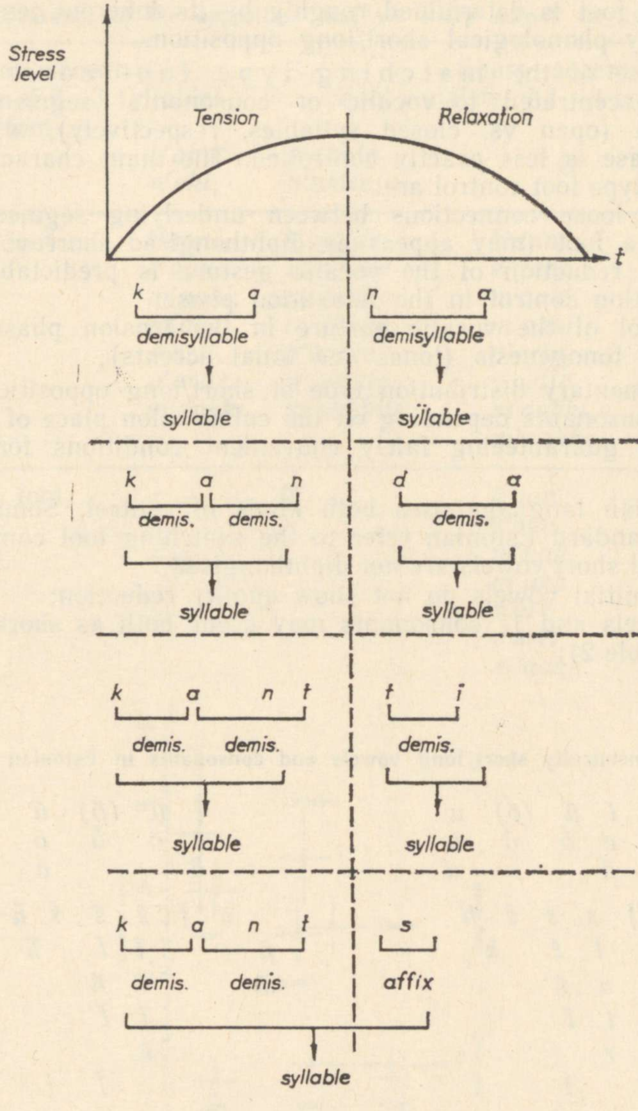

Fig. 2. Foot and syllable.

commands — switching and matching (cf. Eek, Help 1987 : 220—221). In the case of the switching-type foot control the quality of the tract-widening segments is (relatively) exactly controlled in both foot phases, but it does not define the temporal characteristics of vowels themselves, the duration of latter follows mainly from the inherent temporal properties of the corresponding demisyllables. The main characteristics of the s w i t c h i n g - t y p e f o o t c o n t r o l are:
(1) the target of the following vowel is clearly defined;
(2) there is no perceivable diphthongisation of short vowels;
(3) in foot-final syllables — an open set of vowels to choose from;
(4) in foot-final position — no reduced vowels;
(5) opposition between short and long phonemes in both phases of a foot;
(6) relatively weak foot-level and syllable-level isochrony, because the

duration of a foot is determined roughly by its inherent temporal properties and by phonological short/long opposition.

In the case of the matching-type foot control speech energy is concentrated to vocalic or consonantal segments of the tension phase (open vs. closed syllables, respectively), whereas the relaxation phase is less exactly controlled. The main characteristics of the matching-type foot control are:
(1) relatively loose connections between underlying segments at the beginning of a foot (may appear as diphthongised short vowels);
(2) a quality reduction of the vocalic gesture is predictable by less exact articulation control in the relaxation phase;
(3) the control of the vocalic gesture in the tension phase develops conditions for tonogenesis (tones and tonal accents);
(4) a complementary distribution type of short/long opposition between vowels and consonants depending on the culmination place of an energy impulse, thus guaranteeing fairly convenient conditions for foot isochrony.

The Estonian language uses both kinds of control. Some essential features of Standard Estonian refer to the switching foot control:
(1) foot-initial short vowels are not diphthongised;
(2) non-foot-initial vowels do not show quality reduction;
(3) all 9 vowels and 17 consonants may occur both as short and long phonemes (Table 2);

Table 2

Phonemically short/long vowels and consonants in Estonian

			i	*ü*	*(õ)*	*u*					*ī*	*ǖ*	*(ȭ)*	*ū*
			e	*ö*	*õ*	*o*					*ē*	*ȫ*	*ȭ*	*ō*
			ä			*a*					*ǟ*			*ā*
	v	*f*	*s*	*ś*	*š*	*h*			*v̄*	*f̄*	*s̄*	*ś̄*	*š̄*	*h̄*
p			*t*	*ť*		*k*		*p̄*			*t̄*	*ť̄*		*k̄*
m			*n*	*ń*				*m̄*			*n̄*	*ń̄*		
			l	*ľ*							*l̄*	*ľ̄*		
			r								*r̄*			
				j								*j̄*		

(4) a short/long opposition does not display any rules of complementary distribution (Table 3).

There are also some substantial characteristics of the matching foot control in Estonian:
(1) the occurrence of vowels in non-foot-initial positions is restricted (*a*, *e*, *i*, *u*, *(o)*);
(2) all feet occur in either a flat or in a sharp accent (see Table 3, Figure 3*);
(3) a strong temporal compression: mono- and disyllabic feet, irrespective of the accent type and segmental duration, differ durationally less than intra-feet segments (Fig. 4);
(4) there exists a significant temporal relationship among all segments that constitute a foot (Lehiste 1972; Eek 1974);
(5) the foot tends to act as an isochronic unit of a temporal program (Eek, Remmel 1974).

Table 3

The appearance of the opposition 'flat vs. sharp accent' in Estonian

	Flat accent /'/		Sharp accent /`/	
	Foot consists of 2—3 syllables		Foot consists of 1—3 syllables	
Polysyllabic foot with a short first syllable	*k'ana,* *n'aki,*	*k'anale* *n'akile*	Ø	
Polysyllabic foot with a long first syllable	*k'aani,* *k'auna,* *m'aaki,* *m'akki,* *p'arti,* *p'artte,* *m'aakki,*	*k'aanile* *k'aunale* *m'aakile* *m'akkile* *p'artile* *p'arttele* *m'aakkile*	*k`aani,* *k`auna,* *m`aaki,* *m`akki,* *p`arta,* *p`artte,* *m`aakki,* *m`aakke*	*k`aanima* *k`aunaki* *m`aakiki* *m`akkiki* *p`artaki* *p`artteki* *m`aakkiki*
Monosyllabic foot, obligatorily long syllable		Ø	*k`aan* *k`aun* *m`aak* *m`akk* *p`art* *p`artt* *m`aakk*	

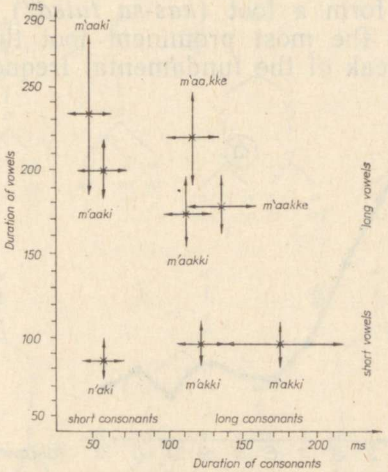

Fig. 3.* Durations of Estonian vowels and consonants in some word structures.

---

* Comments to Figure 3:

(1) The largest standard deviation is observed in the duration of sounds which carry the energy peak of the sharp accent.

(2) Foot is an integral energy unit:

a) each syllable-initial vowel is longer than the vowel following a syllable-initial consonant (cf. *isa — kisa*);

b) a phonological long vowel is longer when it is syllable-final, i. e. when the tension phase of a foot consists of one demisyllable, and a phonological long vowel is shorter, when the tension phase of a foot consists of two demisyllables.

(3) In vowel-dominant sharp accent the F_0 peak and intensity maximum are located in the initial part of the vowel (at about one-third of the duration), whereas in the flat accent they are located at about twothirds of the duration.

(4) The flat and sharp accents are differentiated by the even vs. localised distribution of the pronunciation energy.

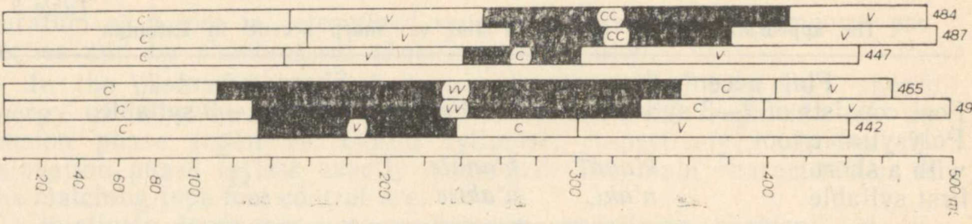

Fig. 4. Average durations of sounds in the following types of words: C'VCV, C'VVCV, C`VVCV; C'VCV, C'VCCV, C`VCCV.

Returning to grammatical units, we can see that there is no direct correspondence between grammatical units and phonetic ones. Morpheme boundaries generally do not coincide with articulatory boundaries of foot phases. But there are also cases where the morpheme boundary is marked phonetically: *võita* 'da-inf.' — *või/ta* 'abess. sg.'; *kaust* 'nom. sg.' — *au/st* 'elat. sg.'. If consonant matter belongs to the word stem then pronunciation energy culminates on the consonant and as a result the preceding vowel shortens; if consonant matter does not belong to the word stem (affixes) then pronunciation energy culminates on the preceding vowel and as a result this vowel lengthens.

As far as words are concerned, the grammatical wordform and the phonetic word are in one-to-one correspondence in the case of content words. If we have an auxiliary word, however, it can join its neighbouring word to form a foot (*kas-sa tuled?*). A word becomes a coherent whole due to the most prominent foot that carries the main stress marked by the peak of the fundamental frequency. Intra-word feet

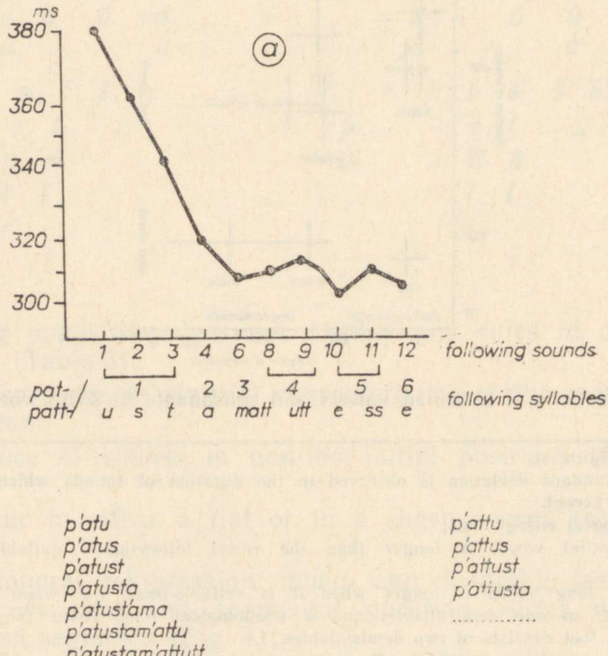

Fig. 5a. Average durations of part of the first foot in the words listed.

are somewhat shorter than the same feet constituting separate words. It seems that the universal phenomenon according to which the duration of the sound/syllable is inversely proportional to the word length is restricted by the rhythmic context. In agglutinative languages with very long words the shortening of the main-stressed vowel/syllable operates within a foot, creating the impression that one long word is made up of two or three separate words (see Fig. 5a, b).

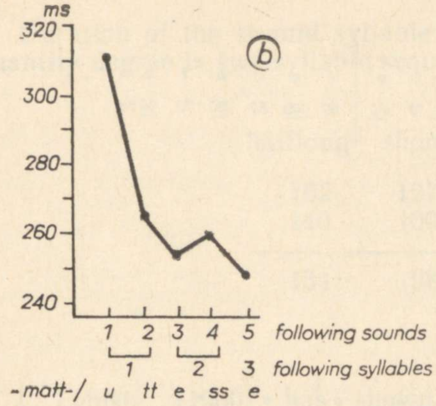

Fig. 5b. Average durations of the fourth, secondary-stressed syllable -matt- (a part of the second, tonic foot) in the words listed in Fig. 5a.

Level pronunciation is avoided within the foot (see above) as well as in a word-internal foot sequence. The intra-foot rhythm is noticeable in a two or three-syllable foot (*kava* 12; *kavala* 132, see Fig. 6). The same tendency can be observed in a word of several feet (*kava/lama* 12; *kava/lama/legi* 132). The general scheme is violated only by those suffixes that carry morphemic stress (*töö/lise/legi* 123). However, there is a strong tendency to avoid two neighbouring stressed syllables (*tööli/-selegi* 12; Fig. 6). Duration varies accordingly (see Fig. 7).

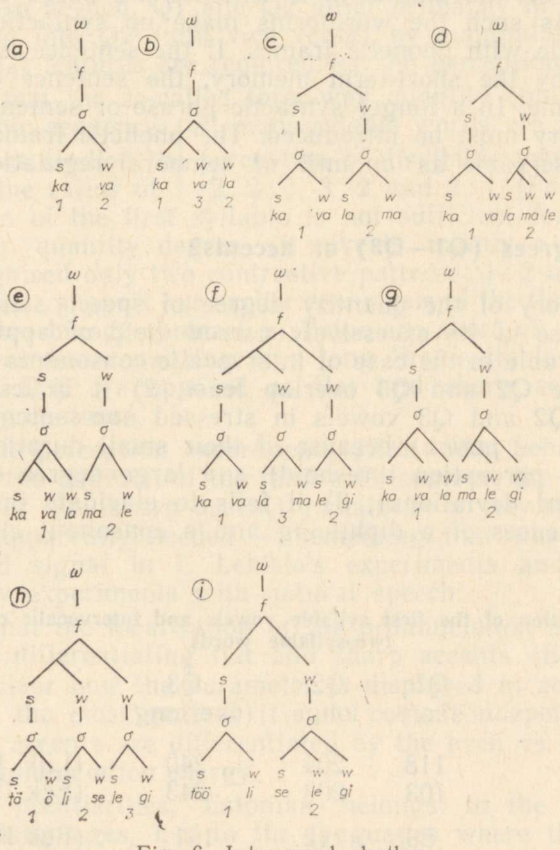

Fig. 6. Intra-word rhythm.

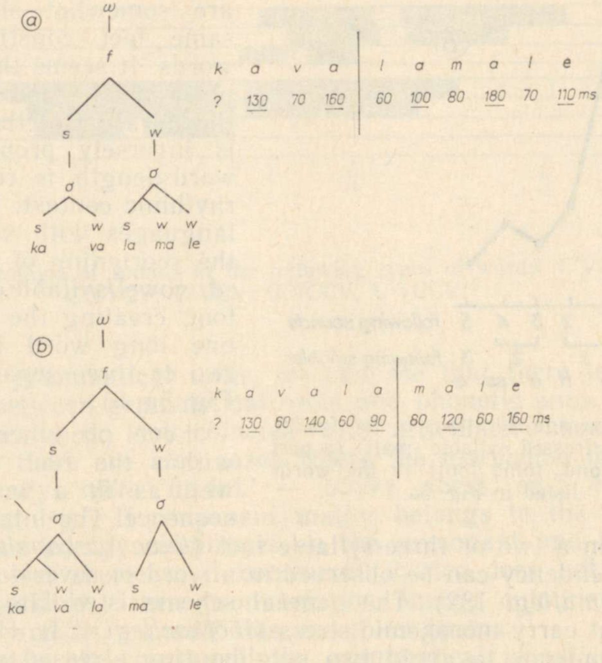

Fig. 7. Rhythmic structure of a wordform and duration of sounds.

Relations of the wordforms in a sentence are based on hierarchical subordination. As such the wordforms make up syntactic phrases that generally coincide with phonetic frames. If the sentence is short enough to be handled by the short-term memory, the sentence coincides with the phonetic frame. In a longer syntactic phrase or sentence, however, a phonetic boundary must be introduced. The phonetic frame, not a whole sentence, is observable as an unit of temporal regulation (Eek, Rannut 1988).

2. Quantity degrees (Q1—Q3) or accents?

The ternary theory of the quantity degree of speech sounds should be discarded because of its excessively narrow field of application: 1) at best it is employable in the case of intervocalic consonants (*lina* — *linna* — *linna*) where Q2 and Q3 overlap least; 2) it is less suitable for describing the Q2 and Q3 vowels in stressed non-sentence-final words (*jama* — *jaama* — *jaama*) because of their small durational difference (verging on the perception threshold) and large degree of overlapping (Fig. 3: standard deviations); 3) it fails to elucidate the cases of Q2 and Q3 in sequences of a diphthong and a consonant cluster.

Table 4

Segmental duration of the first syllable vowels and intervocalic consonants in two-syllable words

	Q1 short	Q2 long	Q3 (overlong)	
CV(V)CV	118	204	240	(Liiv 1961)
CV_1V_2CV	103	203	243	(Eek 1974; 1975)
CV(C)CV / CVC_1C_2V	58	112	161	(Eek 1975)

Duration of the second syllable vowels is inversely proportional to a quantity degree in two-syllable sequences:

halflong	short	undershort	
162	127	93	(Liiv 1961)
140	100	62	(Eek 1974; 1975)
134	98	74	(Eek 1975)

I. Lehiste (1960) has shown that speakers produce words in Q1 with a ratio of the duration of the first syllable and the second syllable of approximately 2 : 3, in Q2 3 : 2 and in Q3 2 : 1.

In Q1 and Q2 words the F_0 peak and intensity maximum are located at about two-thirds of the first syllable duration (mainly rising F_0). In words of Q3 — in the initial part of the vowel (mainly falling F_0), whereas in all cases F_0 continues the fall on the second syllable vowel.

Eek (1980; 1983) found that words were heard as being in Q1 when the ratio of the second vowel to the first vowel was equal to or larger than 1.2. The word was perceived in Q2 when the V_2/V_1 ratio was between 0.57 and 0.81, and in Q3 when the ratio was equal to or smaller than 0.43 but in the latter case the score remains low and the falling F_0 contour was also needed. Q1 and Q2 differ primarily in duration, as Q1 could be turned into Q2 (and vice versa) by manipulation of duration alone. Additional features are required for the perception of Q3.

It is clear that the so-called three-way quantity contrasts require at least a disyllabic sequence (all monosyllabic words are in Q3).

I. Lehiste has studied the perception of disyllabic tokens whose durations were in the ratios of 1 : 2, 2 : 3, 3 : 2 and 2 : 1 (Fox, Lehiste 1987; 1989). Duration of the first syllable is not sufficient for the perception of a particular quantity degree (it reveals mainly a speech tempo). Listeners recognised only two contrastive patterns: 1 : 2 and 2 : 3 vs. 3 : 2 and 2 : 1. Her data support our earlier reanalysis of the three-way quantity contrast into two binary contrast in Estonian: short vs. long segmental quantity whereas feet consisting of a long first syllable (i. e. in earlier terms Q2 and Q3) are opposed as flat and sharp accents based on differences in F_0 contours.

However, the perception experiments (Eek 1980; Lehiste 1988) show that even with the addition of the falling F_0 contour, the highest identification scores for Q3 (sharp accent) remain at a very low level. Something else is apparently needed — something that was not present in the synthesised signal in I. Lehiste's experiments and that was not modified in my experiments with natural speech.

I suppose that the localisation of the pronunciation energy may also play a role in differentiating flat and sharp accents (Eek 1986). However it is not clear how the parameter is displayed in acoustics. I hypothesise that in the most general (i. e. in context-independent) level the flat and sharp accents are differentiated by the even vs. localised distribution of the pronunciation energy.

Despite its peculiarities, Estonian belongs to the *Sprachbund* of Baltic accent languages, i. e. to the languages where the long tension phase of a foot is differentiated phonologically.

References

Clements, G. N. 1988, The sonority cycle and syllable geometry. — Sixth International Phonology Meeting, Third International Morphology Meeting. Abstracts, Krems, 21.

Eek, A. 1974, Observations on the duration of some word structures I. — Estonian Papers in Phonetics 1974, 16—30.

—— 1975, Observations on the duration of some word structures II. — Estonian Papers in Phonetics, 7—55.

—— 1980, Further information on the perception of Estonian quantity. — Estonian Papers in Phonetics 1979, 31—57.

—— 1983, Kvantiteet ja rõhk eesti keeles (I). — KK, 481—489, 549—559.

—— 1986, Problems of Estonian word prosody. — Estonian Papers in Phonetics 1984—1985, 13—66.

—— 1987, Kvantiteet ja rõhk eesti keeles (II). Seisukohavõtt. — KK, 153—160, 205—213, 277—284, 344—352.

Eek, A., Help, T. 1987, The interrelationship between phonological and phonetic sound changes: a Great Rhythm Shift of Old Estonian. — Proceedings of the XI International Congress of Phonetic Sciences 6, Tallinn, 218—233.

Eek, A., Rannut, M. 1988, Temporaalisen kompensaation morfologiset ja syntaktiset tekijät. — Symposium 1988. Läänemeresoome keelte sümpoosion. Teesid, Turku, 11—12.

Eek, A., Remmel, M. 1974, Context, contacts and duration: two results concerning temporal organization. — Preprints of the Speech Communication Seminar, Stockholm, aug. 1—3, 1974. Vol. 2. Speech Production and Synthesis by Rules, Stockholm, 187—192.

Fox, A. 1986, Dimensions of prosodic structure. — The University of Leeds, Working Papers in Linguistics & Phonetics 1986 4, 78—127.

Fox, R. A., Lehiste, I. 1987, Discrimination of duration ratios by native English and Estonian listeners. — Journal of Phonetics 15, 349—363.

—— 1989, Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners. — Journal of Phonetics 17, 167—174.

Fujimura, O. 1981, Temporal organization of articulatory movements as a multidimensional phrasal structure. — Phonetica 38, 66—83.

—— 1982, Relative invariance of articulatory movements. An Iceberg Model. — Paper presented at XIIIth International Congress of Linguists. Tokyo, August 1982, Working Group on Speech Production.

—— 1987, A linear model of speech timing. — In Honor of Ilse Lehiste. R. Channon, L. Shockey (eds), Dordrecht, 109—123.

—— 1988, Some remarks on consonant clusters. — Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo, Annual Bulletin 22, 57—64.

Fujimura, O., Lovins, J. B. 1978, Syllables as concatenative phonetic units. — Syllables and Segments. A. Bell, J. B. Hooper (eds), North-Holland Publishing Company, 107—120.

Help, T. 1985, Üks võimalus eesti verbimorfoloogia käsitlemiseks. — KK, 212—226.

Lehiste, I. 1960, Segmental and syllabic quantity in Estonian. — American Studies in Uralic Linguistics. Indiana University, Bloomington, 21—82.

—— 1972, Temporal compensation in a quantity language. — Proceedings of the VII International Congress of Phonetic Sciences, Mouton, The Hague — Paris, 929—939.

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

—— 1987, Duration: measurements, phonological functions, theoretical implications. — Proceedings of the XIth International Congress of Phonetic Sciences 4, Tallinn, 9—15.

—— 1988, Current debates concerning Estonian quantity. — FUSAC '88. Proceedings of the Sixth Annual Meeting of the Fenno-Ugric Studies Association of Canada, Lanham — New York — London, 77—86.

Liiv, G. 1961, Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid. — KK, 412—424, 480—490.

Viitso, T.-R. 1988, Eesti keele kujunemine flekteerivaks keeleks. — Symposium 1988. Läänemeresoome keelte sümpoosion. Teesid, Turku, 114—115.

АРВО ЭЭК (Таллинн)

ЕДИНИЦЫ ВРЕМЕННОЙ ОРГАНИЗАЦИИ И СЛОВЕСНЫЕ АКЦЕНТЫ В ЭСТОНСКОМ ЯЗЫКЕ

Данная статья представляет собой доклад, прочитанный автором 27 марта 1990 г. в Стокгольме на семинаре, который проводился Отделом речевой коммуникации и музыкальной акустики Королевского технологического института.

В первой части статьи автор, опираясь на измерения длительности, пытается обнаружить в речи сегменты, в рамках которых связанные с длительностью явления регулируемы.

Во второй части статьи на базе проведенных А. Ээком и И. Лехисте тестов восприятия дается акцентологическое толкование эстонских т. н. трех степеней долготы.