



A Metrical Theory for Estonian Quantity

Author(s): Alan S. Prince

Source: *Linguistic Inquiry*, Summer, 1980, Vol. 11, No. 3 (Summer, 1980), pp. 511-562

Published by: The MIT Press

Stable URL: <https://www.jstor.org/stable/4178178>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



The MIT Press is collaborating with JSTOR to digitize, preserve and extend access to *Linguistic Inquiry*

JSTOR

Alan S. Prince

A Metrical Theory for Estonian Quantity*

Da stieg ein Baum. O reine Übersteigung!
O Orpheus singt! O hoher Baum in Ohr!
Und alles schwieg. Doch selbst in der Verschweigung
ging neuer Anfang, Wink and Wandlung vor.

— Rilke

Two things of opposite nature seem to depend
On one another. . . .

— Stevens

0. Introduction

It is remarkable that Estonian distinguishes three degrees of length in both consonants and vowels. But if this bare, exotic fact were the end of the matter, one would simply adjust one's conception of possible quantity contrasts, wonder at the unimagined fecundity of natural things, and leave it at that. What compels attention is not so much the fact itself, but that the three-way contrast, and in particular the distribution of overlength, is richly and curiously connected with patterns of morphology, syllable structure, and stress. The system of dependencies calls for some depth of explanation.

To give the flavor of the phenomenon, let us list here seven of the major systematic properties of overlength. (Following standard usage, we refer to overlong segments as being in “Q3”—the third degree of quantity; long segments as being in “Q2”; short segments as in “Q1”.)

(I) *Association with Stressed Syllable.* Q3 segments occur only in connection with stressed syllables.¹ There is no such constraint on Q1 and Q2.

(II) *Monosyllables.* All major-category monosyllables *must* contain at least one

* This research was supported by NSF grant BNS 7705682 and by the Sloan Center for Cognitive Science at MIT.

I would like to thank Jane Grimshaw, Morris Halle, Paul Kiparsky, Mark Liberman, and John McCarthy for valuable comments on earlier drafts of this article. I am particularly grateful to Kiparsky for discussion of the material relating to Finnish.

I would also like to thank Rev. Ewald Mänd for various phonetic judgments and performances, and R. Jochanan b. Qarathi for help with glosses in the ecclesiastical tongue.

¹ Tauli (1968, 525) attributes this observation to Harms (1962).

segment in Q3. More generally, any word-final syllable that bears main stress will contain a Q3 segment.²

(III) *Adjacent Stresses*. Stressed syllables within a word are usually separated by one or two unstressed syllables. But a stressed syllable associated with a Q3 segment may be followed immediately by another stressed syllable.

(IV) *Sequencing of Q2 and Q3*. According to Lehiste (1965a), a vowel in Q2 may not be succeeded by a consonant in Q3, nor may a vowel in Q3 be succeeded by a consonant in Q2; all other combinations are possible. (Other authors, for example Tauli (1954) and Eek (1975), offer a somewhat different version of the facts; the significance of this will be discussed below in section 1.)

(V) *Intonation*. Syllables associated with Q3 segments are known to have special intonational properties. For example, Lehiste (1960) studies the "falling intonation", typical of statements, as it is realized on two-syllable words (first syllable stressed). She reports (pp. 60–61) that an "abrupt fall" takes place on the second syllable, in the absence of Q3. However, with Q3 in the first syllable, "the falling intonation pattern was characterized by a steep drop on the *first* syllable, followed by level pitch on the second syllable . . ." (p. 61). Thus, Q3 attracts the major pitch movement of the intonation pattern to its own syllable.

(VI) *Grade Alternation*. In the morphophonology of gradation (*Stufenwechsel*), Weak Grade Q1 and Q2 consonants have Q3 counterparts in the Strong Grade; similarly, Weak Grade Q2 vowels (and diphthongs) appear as Q3 in the Strong Grade. But Q1 vowels, unlike Q1 consonants, cannot show up as Q3 in the Strong Grade. Furthermore, with the exception of one loan word, Q1 never alternates with Q2.

(VII) *Syllable Conflation*. Q3 usually appears in the Strong Grade. But when consonant deletion in the Weak Grade reduces a bisyllabic sequence CVCV to a single syllable CVV, the resulting vowel or diphthong VV is always overlong. Conflation of two syllables produces a single syllable with a Q3 nucleus.

Notice that most of these properties have nothing to do with the question of whether the three-way surface contrast arises by (morpho-)phonological rule from a tractable two-way underlying distinction. And indeed, although we shall have something to say about this question, it will not be our focus of attention. If every occurrence of Q3 had to be spelled out in the lexicon, word by word, we could still profitably ask the central question: what is Q3 that it should behave as it does?

Our hypothesis will be that Q3 has a status in phonological representation generically different from the status of segmental length. We will be assuming that the phonological string is organized hierarchically into syllables, groups of syllables, and even higher-order structures, in accord with the basic slant of the metrical theory of stress. (See, for example, Liberman and Prince (1977), Liberman (1975).) We will aim to show that Q2 and Q3 elements are indistinguishable at the *segmental* level in such

² The only polysyllables with final main stress are relatively unassimilated loanwords, e.g., *avenüü*: and so this generalization is particularly impressive.

representations. Only above that level, where aspects of metrical structure come into consideration, does Q3 show itself as a distinct phonological entity. Our essential proposal will be that a syllable phonetically associated with Q3 is in itself a minimal metrical unit [s w]—a *foot*. This prosodic status will provide the environment for the rules assigning duration to segments, and will prove to be the basis for the entire range of properties that we have just noted.

In the second part of the article, we will see how the metrical analysis reveals a deep but previously unsuspected relationship between Q3 and certain other durational phenomena studied by impressionistic and acoustic phoneticians. The hierarchical category “foot”, generalizing across intra- and intersyllabic structure, will show itself as the basic determinant of the relationship. With this result, metrical theory can claim to be opening up new territory lying well beyond the purview of rival segmental and syllabic theories of phonetic representation.

The nonspecialist, or *am ha-aretz*, is fortunate in that there exists a significant body of descriptive and theoretical work on the relevant issues, work done by native speakers, readily available in a familiar scholarly language. Linguists such as Mati Hint, Ilse Lehiste, and Valter Tauli have, besides giving detailed accounts of the factual situation, put forth ideas that lead directly to solution of many of the basic analytical problems. I will draw heavily on their work.

0.1. Phonetic Preliminaries

A somewhat bastard orthography will be used to transcribe Estonian words. Long segments will be written double, VV or CC, as in *linna* ‘town (gen. sg.)’, *tuuma* ‘kernel (gen. sg.)’. Overlong segments will be marked through use of the colon: thus, in *tuu:ma* and *lin:na* (both partitive sg.), the reader is to understand that the *u* (= *uu:*) of the first word and the *n* of the second (= *n:n*) would be described by speakers as being “in Q3”. Thus, the notation VV: signifies “Q3 vowel or diphthong”, and C:C, “Q3 consonant”. The apparently capricious asymmetry in colon-placement will be explained forthwith, in section 1. It will turn out that the sign : indicates what it usually does, phonetic lengthening of the preceding segment. (See below, section 4.2, for details.) Segments in Q1, the short ones, will be written single, as in *kapi* ‘hoof’ (Estonian orthography: *kabi*), using the familiar symbols that most closely approximate their value, except that palatalization of *n*, *t*, *s*, *l* will not be indicated.³

The vowels of Estonian are /i ü e ö ä ã a o u/. Schwa here represents the back or central unrounded mid-vowel, orthographically ö. A large number of *falling* diphthongs

³ As a spur to the auditory imagination, the following impressions might be of use. When noninitial and outside of obstruent clusters, Q1 obstruents are quite lax and, although voiceless, sound voiced to the naïve ear; orthographic *kabi* gives a fair idea of how the English speaker is going to perceive phonological *kapi*. Q2 obstruents sound pretty much like voiceless consonants; *kappi* ‘cupboard, gen. sg.’ resembles English *copy*, except that native speakers uniformly perceive the syllabification *kap/pi*. Q3 consonants sound like the healthy and familiar geminates of Finnish or Italian: *kap:pi* ‘cupboard, part. sg.’.

are obtained by appending /i e a o u/ to the simple vowels. The native consonants are /p t t' s s' k m n n' l l' r h j v/; /f t^s š č/ are also found in recent loans.

Since almost all native words have main stress on the first syllable, the reader may always assume, unless there are indications to the contrary, that this pattern obtains.

1. Rudiments of Quantity

Faced with the three-way surface contrast, a blatant *prima facie* insult to the phonological number *two*, a number of authors have argued that Estonian quantity should be analyzed not as a primitive ternary feature, but as the manifestation of two quite distinct binary oppositions. One of these will be the contrast of single vs. double segments, marking the distinction "short"—"not short". The further differentiation of "not short" into Q2 and Q3 will be accomplished by a separate mark, which receives phonetic interpretation as a lengthening of the appropriate segments. This approach, we shall see, is amply justified. Harms (1962) proposes that syllables containing material in Q3 have a "postposed" stress. Trager (1961) suggests that there is a phoneme of "prolongation of quantity" which can be associated with vowels. Tauli (1954; 1968; 1973) was the first to argue that the Q2/Q3 distinction is a feature of syllables, rather than segments, attributing "heavy" stress to those syllables with Q3 elements, "light" stress to those (stressed) syllables associated with Q1 and Q2 materials. Recently, the Harms-Tauli position on the *syllabic* status of overlength, minus the assumed stress dependence, has been pursued in McCarthy (1976) and Leben (1977).

Representing long segments as geminate sequences is well-motivated. Long consonants closely resemble clusters in their overall distribution and behavior; Lehiste (1966) even reports evidence for occasional rearticulation of long, i.e. geminate, stops.⁴ As in Classical Greek, Japanese, and English, long vowels pattern exactly like the diphthongs. For example, it is a law of Estonian that in native stems only the first syllable may contain a long vowel—or a diphthong; all other syllables must contain short vowels.⁵ This law clearly equates long vowels and diphthongs, presumably as tautosyllabic VV sequences.

An immediate and striking virtue of the nonsegmental approach to overlength is that it allows us to gain some insight into the peculiar-looking distribution of Q3 with respect to Q2 and Q1 (property IV of section 0). The following table of Estonian syllable types, adapted from Lehiste (1965a, 452), summarizes the situation.

⁴ This interpretation of the spectrographic data is challenged in Vihman (1974). As the investigation proceeds, however, we will accumulate plenty of phonological evidence for the equation long = geminate.

⁵ The only exceptions to this law in the larger domain of *word* are diphthongs ending in *i*, which are introduced by affixing such morphemes as /-i-/ 'plural theme'.

(1) *Basic Syllable Types*

I.	II.	III.	
CV-CV	CVV-CV	CVV:-CV	a.
	CVVC-CV	CVV:C:-CV	b.
	CVC-CV	CVC:-CV	c.
		*CVVC:-CV	d.
		*CVV:C-CV	e.

(The dash divides the string into syllables; the notation C-C indicates either a long consonant (Q2) or a cluster, depending on whether the consonants are identical or not. Following Lehiste, I have abstracted away from syllable-initial and syllable-final consonant clusters.)

From this data, Leben has fashioned an elegant distributional argument for construing overlength as a phonological property of syllables rather than segments. I present the argument from Leben (1977) here, modified slightly so as to be rigorously syllabic in character.

Notable in (1) is the failure of Q2 to coexist with Q3 (III d,e). Whereas a Q1 vowel may be followed by any kind of consonant, a Q3 vowel or diphthong may be followed by a single Q1 consonant (which begins the next syllable) or by a Q3 consonant; but not by a simple long consonant or comparable cluster. Similarly, a vowel in Q2 or ordinary diphthong may be followed by a single consonant (Q1), by a double consonant (Q2) or equivalent cluster, but never by a Q3 consonant.

This method of characterizing distribution by sequence constraints is clumsy and incomprehensible. If we abandon it, we see—with Leben—that there is a straightforward horizontal relationship between column II and column III. The permissible syllable types of III are exactly the syllables of II with certain characteristic elongations. Thus, erasing the colons from IIIa,b,c gives IIa,b,c.

If we assume that overlength is not marked underlyingly on the segments where it is heard, but rather on the syllables that contain them, the otherwise puzzling distribution of length degrees becomes a simple function of familiar syllabic variables. Suppose there is a special variety of syllable, call it (temporarily) τ to distinguish it from ordinary σ . The heavy syllable types listed under II— $[\sigma \text{ CVV}]$, $[\sigma \text{ CVVC}]$, $[\sigma \text{ CVC}]$ —can also appear as $[\tau \text{ CVV}]$, $[\tau \text{ CVVC}]$, $[\tau \text{ CVC}]$. In the latter case, they are subject to a rule of phonetic implementation which has the following effects:

- (2) a. $\text{CVV} \rightarrow \text{CVV:}$
 b. $\text{CVVC} \rightarrow \text{CVV:C:}$
 c. $\text{CVC} \rightarrow \text{CVC:}$ } in the env. $[\tau __]$

(2a,b,c) are offered as a convenient map of the process, not a serious set of rules. The actual rule is explored in section 4.2.

It should be emphasized that this ad hoc syllable marking, primitive as it seems, is

superior to having colons salted throughout the string in underlying form, because it does allow us to represent the (now obvious) generalization behind Lehiste's tabulation. But we will seek, and indeed find, a motivated distinction between the two syllable types.

The empirical basis of Leben's argument needs further scrutiny; the crucial facts in column III of (1) are reported rather differently by other scholars. In Eek (1975), a valuable review article, it is held that the structure CVVC-CV, with the first syllable in Q3, can be realized as *any* of IIIb,d,e: CVV:C:CV, CVVC:CV, CVV:CCV. Tauli (1954; 1968) makes a similar point, and he goes on to argue for the syllabic status of overlength based on these data, and quite correctly. He writes (1954, 2–3):

When the long vowel or diphthong is followed by a geminate or consonant cluster . . . it is difficult to say whether the difference between the second and third degrees depends on the length of the vowel or on the following consonant. In fact in the extra long syllables of that type we can prolong either the long vowel (resp. the second component of the diphthong) or the following consonant⁶—the impression is the same in both cases. [Emphasis added/ASP]

Tauli's argument is basically this: the variant realizations, though phonetically distinct, are perceived as equivalent; this follows from a theory in which they all come from a single source. Furthermore, the domain of variance is the *syllable*, and this fact pinpoints the phonological unit that bears the distinguishing mark.

Eek is not clear about whether this variation is dialectal, idiolectal, or “free”; but it does not matter. So long as there is a determinate relationship between the heavy syllable types in column II of (1) and the Q3 types of column III, whether it be Lehiste's restrictive 1–1 map or Tauli's free variation, there will be good evidence that Q3 is phonologically a property of syllables, not segments. A deeper question can now be asked: just what kind of syllabic property is it that we are dealing with?

Before we move on, it will be well to extinguish a troublesome confusion that inheres in the traditional way of describing segments in terms of the categories Q1, Q2, Q3. This arises because long and overlong consonants extend between two syllables, ending one and beginning the next: they are intersyllabic. Long and overlong vowels are of course contained within the confines of a single syllable. So Eek (1975, 29) writes:

In Estonian, quantity is a *segmental* phenomenon as far as the quantity degree is related to a sound belonging to different syllables (e.g. *ka/te*—*ka/te*—*ka:/te*) and a *syllabic* phenomenon as far as the quantity degree is related to a segment consisting of several sounds (e.g. *lau/lu*—*lau:/lu*); contrastive consonant clusters speak in favor of either segmental or syllabic quantity (e.g. *kar/ta*—*kar:/ta*).⁷

Eek's distinction, it should be clear, is somewhat off the mark. All the lengthening

⁶ In Tauli (1968, 527), he allows, with Eek and Lehiste, both the vowel and the consonant to be longer.

⁷ Transcription of Estonian words slightly modified to accord with the standards of this article.

action takes place within the syllable. It is misleading (at best) to refer to the stretch between the vowels in *kat:te* as a Q3 consonant. In reality, it is a cluster of two consonants *that happen to be identical*, of which the first is lengthened (syllable-finally). It is exactly parallel to the cluster *-rt-* in *kar:ta*. Should we call the *r* here “long” or “overlong”?⁸ Rather, perhaps, we should call the entire cluster “overlong”, by virtue of its lengthened first component. But what reason is there to classify consonant clusters? The only point of significance for us is that something special happens to the syllable-final consonant—*r* in one case, *t* in the other—and the consonant beginning the next syllable is quite irrelevant.

In fact, the whole idea of classifying consonants into categories Q1–Q2–Q3 is based on something of a false analogy. The first step on the path of error is to think of the geminate as a “long consonant”, analogous to the “long vowel” of, say, school Latin, thereby losing touch with the fact that geminates pattern like other clusters. From this it is a small step to thinking of the intervocalic articulatory event(s) of *kat:te* as being an “overlong” (single) consonant, one that contrasts with “short” and “long” versions of the same consonant. Thus led on by our analogy and by the desire to find pleasing and symmetrical contrasts along one dimension, we lose sight of the fundamental relationship (structural identity!) between the initial syllables of *kat:te* and *kar:ta*, and we focus on a profitless distinction between inter- and intrasyllabic “quantity”. In an important sense, the “Q3 consonant” is a freak of perspective, a phonetic illusion.

We shall therefore modify our use of the term *Q3* so that it refers to syllables where extra lengthening takes place. There will in fact be no distinction of *quantity* in the phonological representation of Estonian. Both vowels and consonants can cluster, and the members of two-element clusters may be identical. And syllables may be specially designated, giving rise to certain characteristic elongations: Q3. We turn now to the nature of this “special designation”.

2. Stress

The key to the character of the Q3 syllable lies in the peculiar role that Q3 plays in the stress system of the language. Hint (1973) presents a thorough account of the patterning of stress in Estonian—outlined in the English summary to his book, and repeated quite fully, with critical commentary, in Eek (1975). From Hint’s work the following general picture emerges: primary stress falls on the first syllable in native words;⁹ and secondary stresses follow, with one or two syllables lying between stresses in the usual case. (Choice of the one- or two-syllable interstress interval is influenced by syllable structure in ways we shall consider below.)

⁸ Raun and Saareste (1964), for example, use a notation that is ambiguous or contradictory on this point.

⁹ The only exceptions among the native vocabulary are a few words in the feminine suffix *-nna*, e.g. *sapránna* ‘girl-friend (of girl)’. The suffix itself is evidently borrowed from Germanic (Swedish?), along with some of its stress properties. Cf. Swedish *lösínna* ‘lioness’ → Est. *lavínnna*.

The following are typical examples, illustrating both the usual patterns and the kind of optionality that is commonly found:

- | | | |
|--------|----------------------------|----------------------|
| (3) a. | kínnast | 'glove, part. sg.' |
| b. | pálatt | 'piece, part. sg.' |
| c. | pínestav | 'blinding' |
| d. | kávalátt | 'cunning, part. sg.' |
| e. | páhemáit | 'worse, part. pl.' |
| f. | rételile | 'ladder, all. sg.' |
| g. | pínestávale-pínestavále | 'blinding, ill. sg.' |
| h. | pínestávasse-pínestavásse | 'blinding, ill. sg.' |
| i. | hílisémattéle-hílisemáttle | 'later, all. pl.' |

None of the examples in (3) contains an overlong syllable. When we introduce overlength into the calculation, we find a striking contravention of the requirement, evident in (4), that stresses be nonadjacent:

- | | | |
|--------|---------------------------------|---|
| (4) a. | áa:s:tátt | 'year, part. sg.' (cf. pálatt, kínnast) |
| b. | káu:kéle-káu:kele ¹⁰ | 'far away' |
| c. | júl:késse | 'bold, ill. sg.' |
| d. | jál:kétest-jál:ketést | 'track, imprint, el. pl.' |
| e. | tőö:s:tústesse-tőö:s:tustésse | 'industry, ill. pl.' |
| f. | téo:t:táttuttéltt | 'supporter, backer, abl. pl.' |
| g. | ákronóo:míté-ákronóo:mite | 'agriculturalist, gen. sg.' |

The data of (4) show that an overlong syllable may be immediately followed by a (secondary) stress.

What is it about the Q3 syllable that allows it this unusual freedom? The answer becomes apparent, I believe, if we interpret the phenomena of Estonian stress from the viewpoint of the metrical theory of stress, such as is elaborated in Liberman (1975), Prince (1976), Liberman and Prince (1977). Within this theory, distinctions of relative prominence are taken as a reflex of a (binary-branching) relational structure that organizes a string of syllables into a hierarchy of phonological units.

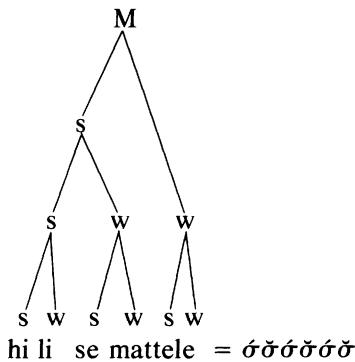
The data of (2), along with the description of Hint (1973, 208–209), which appeals to notions quite similar to those of metrical theory, suggest that the lowest-level—syllable-dominating—metrical units of Estonian are of the form [s w] or [[s s w] w]. As for the character of the higher-order grouping that joins up these basic units, there is little evidence available; let us simply assume that higher-order structure is left-branch-

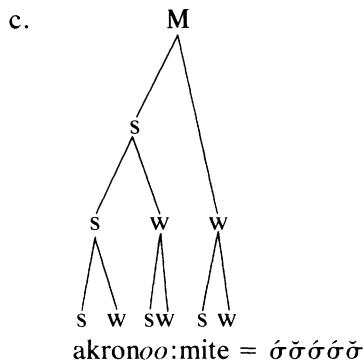
¹⁰ Not all variants are equally favored, as Eek (1975, 15, fn. 16) and Hint (1973, 154–165, *passim*) are careful to note. However, all are clearly possible. Eek remarks that although the form *káu:kele* is “prevalent”, the form *káu:kéle* is a “potential variant”, and contrasts markedly with **pánkéle*, with a non-Q3 first syllable, a stressing that would “lead to strangeness”.

A METRICAL THEORY FOR ESTONIAN QUANTITY

ing. The rule of prominence is this: label every pair [N M] of sister nodes in metrical structure as [s w]. These provisions will result in structures like the following ones:

(5) a.





Contrasts like that between *jul:kesse* and *pimestav* now receive an explanation. Since the first syllable of *pimestav* is not overlong, it is associated with the *s* of the minimal [s w] metrical unit, automatically rendering the second syllable *w* and *hors-de-combat* with respect to stressing. In *jul:kesse*, on the other hand, the overlong first syllable absorbs the minimal [s w] unit all by itself, and the second syllable may therefore receive stress if conditions are propitious, as they are.

In what follows, we will see that the bulk of Q3's curious properties follow directly from this interpretation of "special designation" that the overlong syllable bears.

2.1. The Estonian Stress Pattern

2.1.1. Applying Stress: Relations and Categories. In Liberman and Prince (1977), it is postulated that metrical structure is created as a concomitant to the application (and iterative reapplication) of stress rules within the word. Here, for reasons that will become clear, we prefer to develop a theory of the sort first put forth in an informally circulated paper, Prince (1976), and since pursued by various authors. The idea is that a stress pattern should be conceived of directly as a relation between metrical subtrees and syllable structure, and defined as such, without the intermediary of rules that “apply” stress. In Prince (1976), this relationship was said to be described for a given language by means of a “template”, a fully-labeled tree diagram that represented the syllabic constraints on the lowest level of metrical organization. Iteration of this template across the word was presumed to yield the full pattern. Several familiar languages were described in this fashion: Latin, Greek, Classical Arabic, Literary Arabic, and St. Lawrence Island Eskimo. A simple example of a template, in this sense, would be $[[_w \text{m}] [_s \text{m}]]$, m standing for V or CV, which describes the rudimentary unit out of which the alternating stress pattern of Southern Paiute is built.¹¹

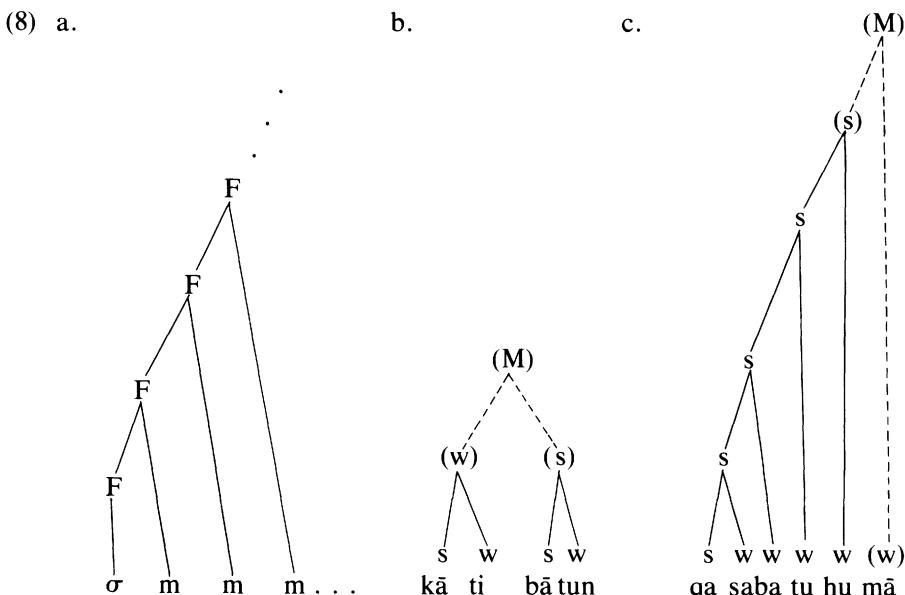
A clarification of this idea is achieved when we distinguish the relational *s*, *w*-labeling from the constituent structure that supports it. The labeling, evidently, can be

¹¹ As described in Sapir (1930). See Wheeler (1979) for an illuminating account of stress in this and related languages.

given by a rule concerned only with sister constituents [N M]. The trees themselves can be described by phrase structure rules of a limited sort. For Classical Arabic, leaving aside details, one might give a template like this: [[_s] [_w m]]. The desired interpretation is that, whereas the right node is weak and must immediately dominate a CV syllable, the left node is strong and may dominate anything—any kind of syllable, or further metrical structure. This is to say: labeling is [s w], and the basic trees are of a left-branching character, as indicated by such rules as these:

- (7) a. $F \rightarrow F\ m$
 b. $F \rightarrow \sigma$

Rules (7a,b) license trees on the pattern of (8a), which can be applied to the analysis of words as in (8b,c):



(Higher-level structure than that described in (7) has been dotted in; labeling, parenthesized here, is the same as in English: right node *s* iff it branches. Note that this describes lower-level labeling as well.)

The rules of (7) will in general allow a multiplicity of possible analyses. In (8c), for example, the string *qaşabatuhu-* is structured as a uniformly left-branching tree; but various substrings might also be so structured: for example, *qaşa* and *batuhu* are perfectly suitable individually for such treatment; this would put main stress on *ba*, incorrectly. Since each word has a single determinate stress pattern, there is some principle of selection at play here. It says that the maximal permissible analysis must be chosen, as in (8c) where the largest possible tree drives main stress to the initial

syllable. This principle corresponds to the "disjunctive ordering" of the standard theory. (The reader is referred to McCarthy (1979a) for a nuanced and illuminating study of Arabic dialects that proceeds basically along these lines.)

Perhaps the most striking formal attraction of such an approach is that it pushes forward the program, adumbrated in Liberman and Prince (1977), of eliminating essential variables—zero-subscript—from the vocabulary of phonological rules. In Classical Arabic, a stress may fall on a syllable lying an (in principle) unlimited distance from the preceding stress or the word-end; in a theory of the type given in Chomsky and Halle (1968), this fact would call for a rule like this:

$$(9) \quad \sigma \rightarrow [+str] / __ ([\sigma C V])_0 \left\{ \begin{array}{c} \# \# \\ \sigma \\ [+str] \end{array} \right\}$$

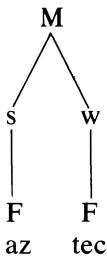
The force of the zero-subscripted term in rule (9) is carried by the *recursiveness* of rule (7a), a property that is "local" in an abstract but nonetheless significant sense. It should be noted that the kind of configuration exhibited in Classical Arabic is attested, with variations of detail, in Eastern Cheremis, Komi, Tahitian, and a number of other, even less well-studied languages.

One large problem inheres in this "fully metrical" approach to stress patterns: relational structure per se is incapable of registering all the significant distinctions of prominence. In certain languages, English among them, a bisyllabic metrical frame [s w] corresponds to two distinct patterns, depending on whether the w-node dominates a stressed syllable (*AzTEC*, *triAD*, *veto*, *doMESticity*) or a completely stressless one (*attIC*, *triAL*, *mottO*, *orCHEstration*). In Liberman (1975) and Liberman and Prince (1977), this was seen as grounds for maintaining, even within relational theory, the categorical distinction between [+str] and [-str]. But to do so quite undermines the program of giving stress patterns directly in metrical terms, and resurrects the stress rule in the old sense.

The theory of relations, then, needs a category; but it need not be, like a feature, localized on vowels or syllables. The natural way to suprasegmentalize the (+, -) stress distinction within a structural theory is to introduce a category dominating structure, such that any occurrence of the category must *contain* one and only one "stressed" syllable. Such an entity may be called a *foot* (and has been by E. Selkirk, who has put forth this idea in recent work: see Selkirk (1977) and (1980)). Within a foot (F), the metrically strongest syllable is interpreted as *stressed* and all the weaker ones as stressless. In the degenerate case, where F does not branch to s,w-structure, the single syllable it dominates will be *faute-de-mieux* the strongest; therefore "stressed".

Adding the category *F* to *M* in the metrical arsenal gives rise to slightly ramified structures which fully represent all gradations of stressing:

(10) a.



b.



In (10a), the final, weak syllable is a foot—[_F t^hɛk)—and so the vowel in it does not reduce; in (10b), the final syllable is merely a weak appurtenance to a foot, and its vowel is therefore reduced, its initial consonant lax and flapped.¹²

The argument for doing things this way is straightforward: *foot* is an ineliminable category of phonological description (the domain of rules, both segmental and prosodic); the feature [±stress] has no independent life outside its role in setting up boundaries to feet; therefore, [±stress] may be dispensed with as the artifact of an unnecessary calculation, if feet are computed directly. (This is, of course, the *structure* of an argument; the premises must be established empirically.)

Changing to a foot-based theory entails a certain amount of technical redevelopment; for example, rules of destressing turn into rules depriving syllables of *foot* status. Selkirk (1980) deals with these matters in great detail. An important issue is the fate of the constraint that bars metrical *s* from immediately dominating [–stress], for this plays a variety of explanatory roles in the description of English (and presumably other languages). For example, it prevents the so-called “rhythm rule” from changing the [w s] of *maroon* to [s w] in, say, *maroon sweater*; immutable *maroon* thus contrasts in rhythmic behavior with *thirteen* and *Dundee*.¹³ Within foot theory, it is plausible to give the sense of this restriction in the following implicational form:

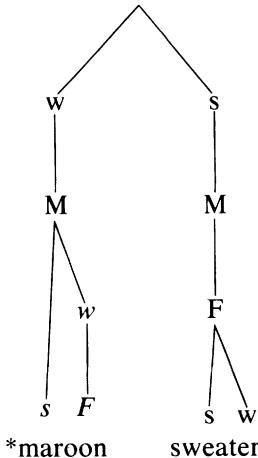
(11) If a *w*-node dominates an *F*, then its sister *s* must also dominate an *F*.

The idea behind this is that *F* is intrinsically stronger than a naked syllable. If the two are metrically paired, *F* must be—and remain—the stronger. Restriction (11) can be seen in operation in the forms of (12):

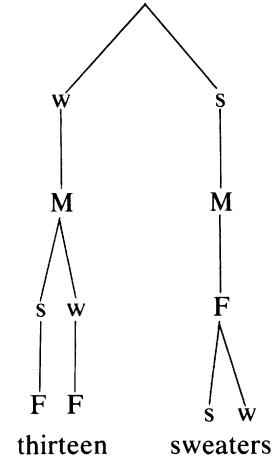
¹² See Kiparsky (1979), Selkirk (1978) for the treatment of English stop allophony in terms of rules bounded by the foot-domain. See Selkirk (1980) for vowel reduction.

¹³ See Jones (1972, 253–254). The phenomenon of rhythmic stress shift is discussed fairly extensively in Liberman and Prince (1977, 311–323). See also Kiparsky (1979) for further observations on English, and Nespor and Vogel (1979), Bing (1980), McCarthy (1979b), Solan (1979), and Wheeler (1979) for rhythmic effects in other languages.

(12) a.



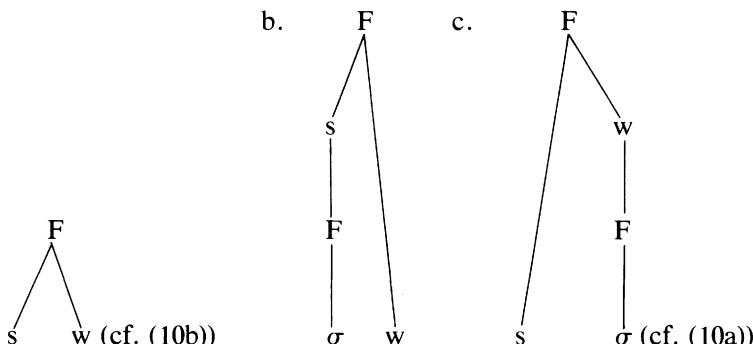
b.



Principle (11) also explains why the foot-structure resulting from a rule like $F \rightarrow F \sigma$ will always be labeled [s w].

It is important to observe that, even assuming (11), the foot theory is not a representational equivalent of the [\pm stress] theory that it superannuates. If we distinguish between a weak syllable that is a foot (= stressed) and one that is merely a proper part of a foot (= unstressed), we are also forced, willy-nilly, to distinguish similarly among strong syllables. The new notation leads us to imagine three distinct variants of the [s w] relation:

(13) a.



The contrast between (13c) and (13a or b) is just that between *áztèc* and *áttic*. But (13a) and (13b) must have identical stress patterns. A *tertium quid* seems to have materialized in the person of (13b). But *pro quo?* one asks, sharply. Would-be wielders of Occam's Razor will have to return to their dens unblooded, however. As we will show in detail, the distinction between strong syllable as foot and strong syllable as foot-part is amply and obviously a major fact of Estonian.

2.1.2. Estonian Stress. The challenge, then, is to describe the character and limitations of the foot in Estonian. Hint carries through his analysis of the stress system by dividing words into what he calls “phonological component words”—*feet*, to us. (Eek, in his redaction of Hint’s work, calls them “rhythmic units”.) A major difference between our conception and Hint’s is our limitation to binary branching. Hint admits unanalyzed tripartite feet and breaks words down directly into multipartite sequences of “component words”. Nevertheless, there is little difficulty in interpreting his conclusions, and those of Eek, as restrictions on our notion of “foot”.

Putting aside (momentarily) the problems of final syllables and of Q3, we find that the usual foot consists of two or three syllables. In a bisyllabic foot, the constituent syllables may be of any type whatsoever, as the following examples illustrate:

- (14) a. [F hili] [F semat] [F tele] ‘later, all. pl.’
- b. [F əppet] [F tuste] [F leki] ‘lessons, too, all. pl.’
- c. [F vara] [F seimat] [F tele] ‘earliest, all. pl.’
- d. [F valut] [F tavait] ‘aching, part. pl.’
- e. [F noorik] [F kule] ‘bride, all. sg.’

A trisyllabic foot may be achieved by adding a syllable to any bisyllabic foot-form, but the added syllable must be “short” (Hint (1973, 206); Eek (1975, 17)). A “short” syllable is of the form C_0V , except at the end of the word, where C_0VC also counts. No foot contains more than three syllables. Thus, a word like (14a) may be stressed according to the alternative analysis [F hilise] [F mattele]; but neither *appettusteleki* (14b) nor *varaseimattele* (14c) can be, because in each case the first foot would illegitimately end with a heavy syllable: *[F əppetus], *[F varasei]. These principles of foot structure are represented in the following rules (where *u* = syllable here, *m* = “short” syllable):

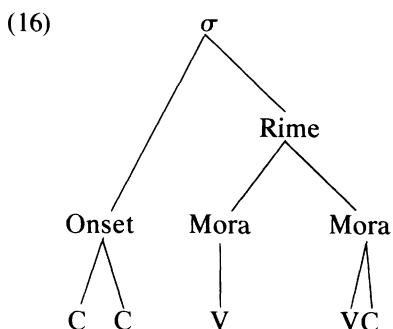
- (15) a. $F \rightarrow u\ u$
- b. $F \rightarrow F\ m$
- c. F contains no more than 3 syllables.

Rule (15a) allows for the simple bisyllabic foot; rule (15b) lets longer feet arise, based on the simple ones, as in [F hili] se] [F matte] le]; rule (15c) limits the magnitude of complex feet coming about through recursion. (Our reasons for choosing the locution (15c), rather than, say, using a nonrecursive category *F'* on the lefthand side of (15b), will soon become apparent.)

Above, we suggested that the overlong syllable is essentially a heavy syllable that comprises in itself a “minimal metrical unit” [s w]; we now recognize this unit as the foot. But feet are characteristically bipartite. What is it about the heavy syllable that uniquely fits it to be construed as a foot? There is a tradition in the linguistics of stress and quantity—well-represented in the work of Trubetzkoy (for example (1969))—that contemplates the frequently found equivalence *one heavy = two light syllables* in terms of the notion *mora*. A mora, James McCawley has somewhere written, is *one* of what

heavy syllables have *two* of. In these terms, we can say that a foot may consist of two syllables, as we have seen, or minimally, of two (tautosyllabic) moras, in somewhat special circumstances.

The mora receives a more concrete interpretation than jesting McCawley's when we take the syllable to be a hierarchically-structured entity, as I think J.-R. Vergnaud was the first to observe, at least in the present age. Dividing the syllable into *onset + rime*— $C_0 + V_1C_0$ —a clear difference between heavy and light syllables emerges. The light syllable, $[_0 C_0] [r V]$, has a simple rime, consisting of one unit; the canonical heavy syllables $[_0 C_0] [r V V]$ and $[_0 C_0] [r V C]$ have complex, two-unit rimes. We should not too hastily identify the *segments* of the rime as moras, for fuller rimes are possible, for example $[r V V C]$, $[r V C C]$, but they do not produce “trimoraic” syllables. But if we accept the hypothesis that syllable structure is also strictly binary-branching,¹⁴ then these rimes must be further divided—as V-VC, VV-C, V-CC, or VC-C, whichever is appropriate. Thus, any complex rime will be divided at its topmost level into two bits; these we can identify as the moras of traditional description. In short, “mora” means ‘immediate constituent of the rime’. By way of exemplification, we offer the following diagram of the complex syllable CCVVC.

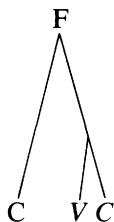


A crucial observation is that *onset* plays no role in determining the form of stress patterns; in linguistic rhythm, *rime* is all that matters. Of the dozens of stress systems that are known with reasonable certainty, not a single one depends on *onset* structure or refers to it in any way. Such forthright unanimity is most impressive in this world of tattered implications and fading subregularities. Linguistic theory ought to encounter the fact boldly. A notation like C_0 will not do: it is designed to allow for choices, as C vs. CC vs. CCC vs. . . . vs. C_0 ; but there is no choice. In this article, we will postulate that syllable structure outside the rime is simply invisible to prosodic rules, although it remains present and may be accessed by rules of other types. By this fiat, stress rules will not be able to refer to extrarime structure, and in determining relations between

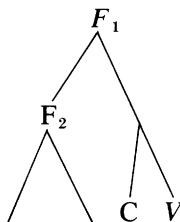
¹⁴ This idea originates with Kiparsky (1978). For development of it, see Kiparsky (1979), Selkirk (1978), McCarthy (1976; 1979a; 1979b), Halle and Vergnaud (1978).

prosodic categories, intervening nonrime material will not count. When describing or defining metrical trees, then, we will say that one prosodic node (e.g. *Foot*) immediately dominates another (e.g. *Mora*) even if irrelevant structure stands between the two.¹⁵ Diagrams (17a,b) represent typical configurations:

(17) a.



b.



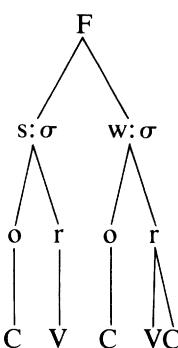
Of (17a), it is legitimate to say that the daughters of *F* are the two moras of the rime; of (17b), that *F*₁ immediately dominates both *F*₂ and the single mora embedded on its right.

A further convention: we shall regard complex nodes, e.g. *s:F:σ* “strong mono-

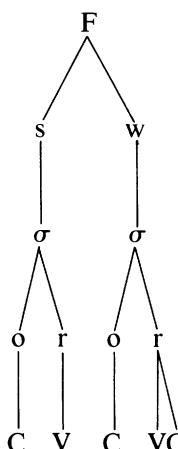
syllabic foot”, as equivalent to exhaustive domination: *F*, and vice versa. All of the

structures in (18) are rendered prosodically equivalent by our assumptions:

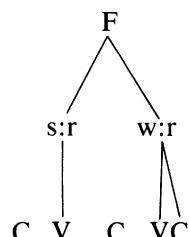
(18) a.



b.

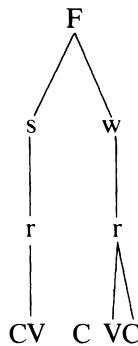


c.



¹⁵ An interestingly different and more extensively formalized version of the basic idea—rime-relevance, onset-irrelevance—is given in Halle and Vergnaud (1978). McCarthy (1979a) is a published article that makes good use of their formalism.

d.



e.

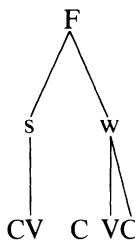


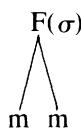
Figure (18a) is the full representation of a typical foot. In (18b), the complex nodes have been expanded. In (18c), reference to onset branching has been omitted: this is the prosodic point of view made graphic. In (18d), complex nodes are expanded. In (18e), the obvious is completely omitted: this is the shorthand that will be used from here on.

With these principles in hand, we can interpret the rules of (15) as giving the full range of basic foot patterns. The term *u* ranges over *syllable* and *mora*, prosodic units below F in the hierarchy of inclusion that extends upward to M and beyond. Thus, (17a) and (18) are instances of the rule $F \rightarrow u\ u$. The term *m* may now be unveiled as *mora*; (17b) is described by the rule $F \rightarrow F\ m$ (15b). The set of rules (15a,b) defines exactly the following five foot-types:

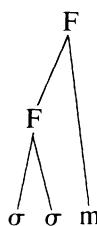
(19) a.



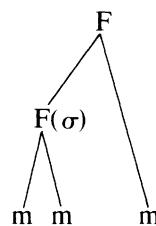
b.



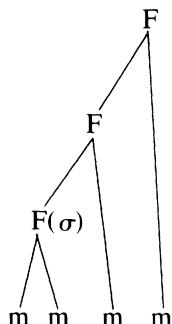
c.



d.



e.

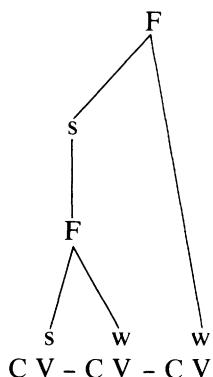


The parenthesized σ of (19b,d,e) is a heuristic indicating that in the crucial cases the

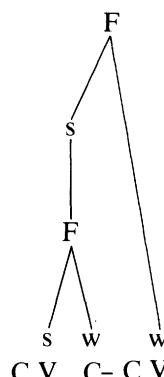
marked feet will be syllables; notice that $[_F m m]$, (19b), can easily be disyllabic, a case which is also covered by $[_F \sigma \sigma]$, (19a), when both σ are monomoraic. In (19e), the marked foot cannot be disyllabic, otherwise the topmost F would contain 4 syllables.

In interpreting the use of m in rules like those of (15) and, equivalently, in diagrams like (19b,c,d,e), it is presupposed that metrical tree structure cannot contradict the intrinsic branching of syllables (i.e. rimes). Thus, a tree like (19d), left-branching with terminal m , can be planted as in (20a,b) but not in (20c), where a single heavy syllable is split up by the metrical constituency.

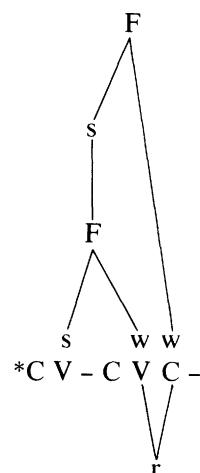
(20) a.



b.



c.



At the lowest significant level, then, metrical form is determined jointly by foot rules, such as (15), and by the givens of syllable structure. In Prince (1976), this hypothesis about the relationship of metrical and syllabic organization is called “the principle of syllabic integrity”.

The foot-types of (19) are exemplified in these words:

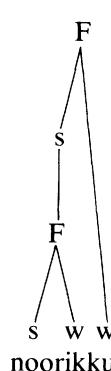
(21) a.



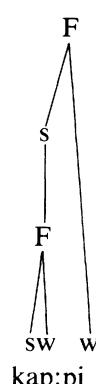
b.

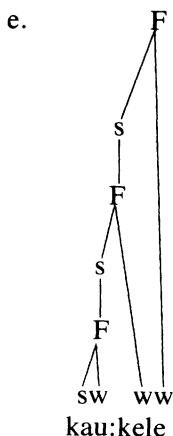


c.



d.





Note that these words are in “narrow” transcription, the relevant lengthenings having been inserted (on the basis of foot-structure).

The rules allow, via recursion, a trisyllabic foot beginning with a Q3 syllable, as in (21e). But the recursive rule $F \rightarrow F m$ is motivated by non-Q3 cases like *nóorikku* (21c), which allow only a “short” additional syllable to be attached to the basic foot. The prediction is that *both* syllables following the overlong one must be monomoraic, as in *káu:kele* (21e). This is evidently correct; Eek notes explicitly that trisyllabic feet may be formed “with Q3 initial syllables if the second and third syllables are short . . .” (1975, 17).

However, the rules predict as well that even in a bisyllabic foot initiated by a Q3 syllable, the second syllable must be light; for this configuration also arises only through the rule $F \rightarrow F m$. This prediction is borne out fully, by Hint’s account, only in bisyllabic words; in longer words, where another foot follows, it turns out that a Q3 syllable can join up with any kind of syllable to form a bisyllabic foot. Typical examples, from Hint (1973, 160), are *tőö:s:tüs-tésse* (4e), *áu:sät-téle* ‘honest, all. pl.’, *vát:mët-tékä* ‘key, com. pl.’, *tőö:k:käi-léki* ‘hardworking too, all. pl.’. (Of course, these all allow the predicted alternate stressing $\dot{\sigma}:\dot{\sigma}\dot{\sigma}$, with a word-initial monosyllable foot.)

Since this fact does not follow from anything previously established, we simply record it in a special rule:

$$(22) \quad [_{\mathbf{F}} \sigma] \sigma \rightarrow [_{\mathbf{F}} 1 2] / \begin{matrix} & \\ 1 & 2 \end{matrix} \mathbf{F}$$

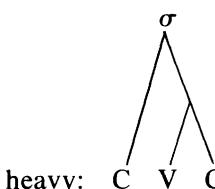
Rule (22) says that the sequence $\sigma:\sigma$ may be construed as a foot medially in metrical structure, i.e. before another F.

Final syllables pose some interesting descriptive problems. They are necessarily stressless when of the form CV or CVC, but may be stressed if they contain a complex

vowel or end in a consonant cluster. When stressed, a final syllable must, in the present theory, be a foot unto itself—for there is no following syllable that it can be metrically compared with, as *s* to *w* under F. Here again we see that a syllable which is F must be heavy—bimoraic. But it is puzzling that the otherwise quite heavy syllable CVC patterns with the syllable CV as light at word-end. Some explanation is surely to be found for this not uncommon situation. Arabic, for example, is similar, as McCarthy (1979a,b) shows; but, lacking explanation, let us concede the fact, and seek an appropriate representation for it.

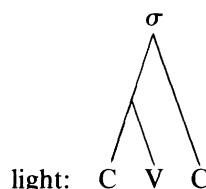
The long and short of it is that a word-final consonant does not act, prosodically, as if it were a part of the word-final syllable. It is plausible to assume that this reflects a real structural anomaly. We can experimentally adapt a suggestion of McCarthy's here, postulating that the CVC light syllable has different geometry from its more familiar heavy isomer, that it is *levo-* rather than *dextro-* ramatory, as it were:

(23) a.



heavy: C V C

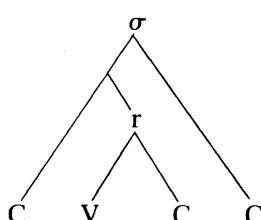
b.



light: C V C

The heaviness of (23a) comes from its branched rime; in (23b), it is perhaps not so clear what the rime constituent is, but either of the obvious choices—V or final C—gives a unit rime, a single mora. McCarthy proposes (1979a) that the rime is the right daughter of σ , therefore C here, but we suggest instead that it is the *sister of the onset*, therefore V. This means that we shall be dealing with heavy syllables of the form (24), where only the constituent *r* is accessible to rules of prosody:

(24)

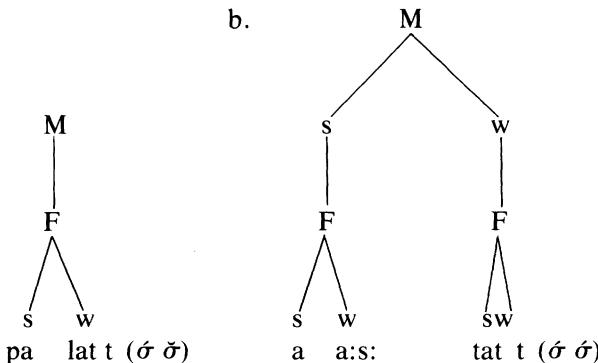


Our proposal, then, is that the last consonant of an Estonian word is adjoined outside the rime, where it will be prosodically inert.

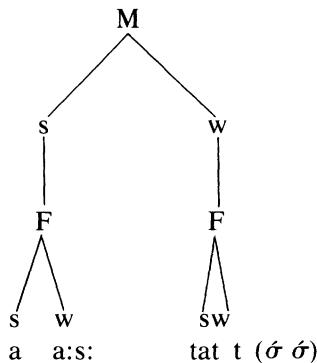
Under this conception, stressing (and nonstressing) of final syllables in typical

examples from (3) and (4) will emerge as shown in (25):

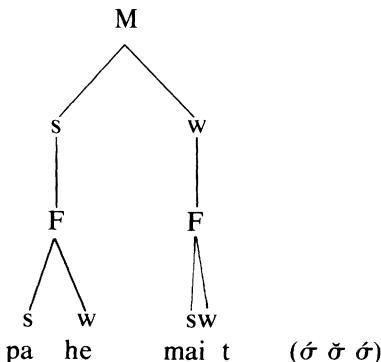
(25) a.



b.



c.



(Only the prosodically relevant branching has been drawn in.)

A question now arises: if stressed final syllables are *feet*, why are they not *overlong*? Foot-hood is evidently a necessary but not sufficient condition for receiving Q3 treatment. However, it is for the most part quite easy to spot the syllabic feet that are overlong: basically, the Q3 syllable is the strongest foot of the M-domain. Weak feet within M, such as the stressed finals of (25b,c), do not figure in this calculation. There are, of course, some subtleties to be managed, and these will be taken up in section 4 below, when we turn to the details of the morphophonology of overlength.

A final major question about the intent of the proposed rule system: when can the *u* of $F \rightarrow u u$ be interpreted as *syllable*, when as *mora*? The allocation of overlength (F) to syllables, where $u = m$ for both *u*'s, is governed by morphological considerations. In the other cases, something like a principle of maximality seems to be operating: take $u = \sigma$ when you can. This means that, overlength aside, we do not expect to find adjacent stresses caused by the occurrence of a monosyllabic foot; the simple foot will usually be bisyllabic, and this (maximal) option will always be available in the syllabically fertile medial regions. Only at word-end, after we have run out of options, does the single syllable (when heavy) emerge as the maximum foot. So in *pimestattutte*, we

do not find *[_F pimes] [_F tat] [_F tutte] (**pímes-táttúte*), but only [_F pimes] [_F[_F tattut] te]. And in *páhemáit*, because trisyllabic F must end on a monomoraic syllable, by F → F m, the word's first foot must lapse after *pahe-*, making a foot, perforce, of -*maít*: [_F pahe] [_F mait]. A survey of the extensive word-lists in Hint and Eek shows that this view is essentially correct: they cite no cases where a non-Q3 internal syllable forms a foot, under conditions of normal, nonemphatic pronunciation.¹⁶ The explanation probably lies in the fact that “maximality” functions here to avoid “stress clash”¹⁷ by ensuring a decent interval—commonly a syllable—between peaks of prominence.

To sum up, then. We have found that the basic foot-patterns of Estonian are described by a few simple rules, and that each word (M) must be parsed into a continuous string of feet (probably a universal of prosodic theory, on a par with the law that each word must be a continuous string of syllables). All labeling is [s w]. Such rules can be thought of as well-formedness conditions on metrical structure; they define the array of permissible Estonian word-trees, just as the phrase structure rules of the base define the set of syntactic trees. From these provisions the central characteristics of the system follow: words have initial main stress; secondary stresses appear every two or three syllables, with an optionality that is circumscribed by the constraints on foot structure; and, because feet are defined on moras as well as syllables, a single heavy syllable can under certain conditions constitute a foot.

At this point, we are ready to see how the properties of Q3 emerge from its role in the stress system. But several points of detail remain to be treated, and we shall do this directly, before returning to the main theme in section 3 below.

2.1.3. Three Points and the Rules, Summarized. (i) *Proclitic Syllables.* In atypical circumstances, a word may begin with a single unstressed syllable. The nativized suffix -*nna* is responsible for this configuration in words like *sapránna* ‘girlfriend’, *løvínna* ‘lioness’. A number of loanwords are of this type: for example, *projékti*, *mahórkale* ‘cheap tobacco, all. sg.’, *kaleítoskóo:p:p*. We note this possibility in the following rule, which is, in a sense, an auxiliary or patch-up procedure that deals with a marked situation, rather than a part of the basic system:

$$(26) \quad \sigma F \rightarrow [F \ 1 \ 2] / [M \ \underline{\quad}] \\ \quad \quad \quad \quad \quad \quad 1 \ 2$$

(ii) *Fixed Secondary Stress.* Hint observes that for certain words, e.g. *sémínár* ‘seminary, teacher’s college’, *múu:séum*, there is an unexpectedly narrow range of stressing possibilities under affixation. One finds only *séminárisse*, not **séminarísse*, which is equally admissible by the usual rules. One also finds *múu:seúmi*,

¹⁶ Under special emphasis, back-to-back stresses can appear, especially in words with medial diphthongs. Typical is *ea:k:kaimait* ‘most aged, part. pl.’, which is normally stressed [_F ea:k] [_F kaimait] (ó:óó), but which can be dissected emphatically to [_F ea:k:] [_F kai] [_F mait] (ó:óó).

¹⁷ For formalization of this notion, see Liberman and Prince (1977). For discussion of the phenomenon, see the references of fn. 13.

múu:seúmisse, but not **múu:séumi* or **múu:seumísse*. Secondary stress is fixed on the final syllable of the nominative—where it may *not* appear on the surface, because that syllable (*-ar*, *-um*) is monomoraic when final, therefore not fit to support a foot. Evidently, some sort of lexical specification is required. *What* sort depends in large part on the lexical form attributed to Estonian nouns. It is plausible to assume that the nominative singular is generally derived by deletion of a stem-final vowel that shows up in the other cases and before derivational suffixes. Stems like /kappi/, /raamattu/, /pimesta+va/, /seminari/ will underlie nominatives *kap:p* ‘cupboard’, *raamatt* ‘book’, *pimestav* ‘blinding’, *seminar*. We could then assume that in the cases at hand, the final two syllables of the underlying stem-form are lexically specified as a constituent: /semi[nari]/, /muuse[umi]/. If this constituent makes it to the surface, it will be accounted an *F* by the rules describing foot-types. If not—due to nominative vowel deletion—it will simply disappear, since it bears no label that might remain as a mark of its existence.¹⁸

Interestingly, the derivational suffix *-tar* (i.e. /ttari/) is like the cited loanwords in its stress behavior. It is also worthy of note that the items of this class all end in the sequence VR, R a voiced consonant (Hint (1973, 214)), and that new loans of this shape productively enter the “fixed-secondary stress” class: e.g. *bätminton*. A lexical redundancy of some generality may be at issue.

(iii) *Preference*. In a system of optionality, preferences can be expected to emerge. The basis of such preferences, whether morphological, phonological, or conventional, deserves scrutiny. We have already suggested that a notion of “stress clash” may play a role in determining preferred patterns. Another interesting fact is that stress is more likely to fall immediately after a Q3 diphthong ending in a nonhigh vowel, a so-called “late diphthong” (*hilisdiftong*), than after other types of Q3 syllables. The reader is referred to Hint (1973, 154–165) and Eek (1975, 11–26) for an exemplified discussion of the issues.

(iv) *Rule Summary*. For the reader’s convenience, all the rules discussed are repeated here:

- (27) a. *Basic Foot-Defining Rules*
 - i. $F \rightarrow u\ u$ (where $u = \sigma, m$)
 - ii. $F \rightarrow F\ m$

¹⁸ In what follows, we will assume, for purposes of morphophonemic discussion, the existence of nominative vowel deletion. However, it is possible to regard the evanescent vowels as themes, morphemes appearing everywhere except the nominative singular. (This gives us four classes of stems, by theme-vowel: *a*, *i*, *u*, and *i* alternating with *e*; the *i* class is the productive one, into which new loans are entered.) On this view, the significant observation would be that the (now) stem-final syllable is stressed whenever it appears before the theme; thus, we can say that the theme-vowel is *prestressing* here, like *-ic* in English. Included in the lexical representation of the *theme* would be an annotation similar to a syntactic subcategorization frame: $[_F\ \sigma C_0 __ C_0]$. The intended sense is that an item so marked must appear in the second syllable of a foot.

- iii. $[_F \sigma] \sigma \rightarrow [_F 1 2] / \underline{\quad} F$
 $\begin{matrix} 1 & 2 \end{matrix}$
- iv. No foot may end on more than two *w* syllables.
- b. *Auxiliary Rule for Proclisis*
 $\sigma F \rightarrow [_F 1 2] / [_M \underline{\quad}]$
 $\begin{matrix} 1 & 2 \end{matrix}$
- c. *Higher Order Branching*
 Feet are arranged into a left-branching tree.
- d. *Labeling*
 All labeling is [s w].
- e. *General Condition on Metrical Composition of Word*
 The contents of the domain M must be exhaustively divided into feet.

3. Properties of Q3

Our essential hypothesis is that Q3 syllables are *feet* in metrical structure. Recall that “quantity” at the segmental level is a matter of gemination, and that the lengthenings of Q3 are projected by rule from the hierarchical representation. Although various details remain to be discussed, our hypothesis has enough structure now that we can see how the major peculiarities of Q3 turn on the properties of metrical representation.

(I) *Association with Stressed Syllable*. Q3 only occurs in or about a stressed syllable. Each F must contain one stressed syllable: the metrically strongest syllable in F. If there is but one syllable in F, it is stressed. Since Q3 develops in the configuration $[_F \sigma]$, it occurs only in stressed syllables.

(II) *Monosyllables*. All lexical category monosyllables are overlong. Every *M*—phonological word—consists of Fs. A monosyllable must therefore be of the form $[_M[_F \sigma]]$. In this configuration, σ will be overlong. Notice too that the exigencies of F will restrict the segmental composition of such σ . Estonian stress, independent of Q3, requires that F consist of two units: syllables or moras. A monosyllable foot must be bimoraic. Consequently, it is predicted that C₀V and C₀VC, both monomoraic at word-end, are not fit segmental timber to build a word from. (Recall that the last consonant of a word does not contribute to the prosodic weight of the final syllable.) This prediction is correct: such acceptable syllables as, say, *kä* and *kät*, cannot be words. The stem for ‘hand’ is apparently /kät/ (see Milsark (1971)), but it can surface only with attachments: e.g. *käsi* /kät+i/, nom. sg.; *kät:t* /kät+t/, part. sg.

More generally, any final syllable that bears main word-stress must be overlong. This also follows from our hypothesis. Any stressed final syllable must be a foot unto itself; if it is the strongest foot of the M domain, it is subject to the rule of overlengthening. Because of the nature of Estonian stress, this generalized constraint will only show up in recent loans, e.g. *avenüü*: where the law of initial main-stress is abrogated;

but it is all the more impressive that even in relatively unassimilated loans the principle of Q3 distribution is manifestly inflexible.

(III) *Adjacent Stresses*. As is usual in the world's languages, stresses like to be a syllable or two apart from each other; but a Q3 syllable (always, of course, stressed) seems to invite an immediately following stress. This is indeed the fact which was brought forth to motivate the representation of the Q3 syllable as foot. One foot starts where another leaves off, and every Estonian foot begins with its strongest element, the "stressed syllable" it contains. Thus, the configuration $[_F \sigma] [_F \sigma \dots]$ results in stress adjacency.

(IV) *Sequencing of Q2 and Q3*. As Leben observed, Lehiste's constraints on the sequencing of long and overlong segments, discussed in section 1, will follow if overlength is a phonological property of syllables. We also noted that the other prevalent interpretation of the facts, as permissible variation in the realization of Q3, strongly supports the syllabic interpretation. We treat overlength as a metrical property of syllables, and therefore the good of Leben's and Tauli's position accrues to us as well.

(V) *Intonation*. In discussing the "falling intonation", as applied to single-word phrases, Lehiste (1960) reports that an "abrupt fall" takes place on the second syllable when the first is not overlong; but the fall takes place on an overlong first syllable. It appears that the nuclear HL of the falling intonation is mapped onto two syllables when Q3 is not at issue, one tone to one syllable, but is mapped wholly onto a single overlong syllable. Here again, as in the stress system, a Q3 syllable is treated as the equivalent of two ordinary syllables. In our theory, the basis of such equivalence is the metrical foot. Thus, it is plausible to hypothesize that the nuclear (falling) tone is realized over the unit *foot*, rather than syllable; the foot-hood of the Q3 syllable causes it to capture the whole contour. To put it slightly differently, we could say that the tones of an intonation melody (H and L, in this case) map onto metrical terminal nodes—daughters of F—rather than syllables. With no more than a modest understanding of the intonation system, we cannot hope to establish the truth of this conjecture; nevertheless, it is well worth noting that the present theory does promise a natural account of what has hitherto been viewed as simply another isolated quirk of Q3.

(VI) *Grade Alternation*. "Short" and "long" consonants can appear as overlong in the Strong Grade; so can "long" vowels and diphthongs; but "short"—single—vowels do not turn into Q3 vowels under the influence of gradation. We will take up the details in section 4 below, but the basic proposal is straightforward enough to be presented here: the principal effect of the Strong Grade, we will argue, is to assign *F* to syllables. Syllables with complex rimes—VV, VC, VVC, VCC—will end up overlong, but monomoraic syllables will be untouched, since they cannot be *F* at the surface, necessarily binary as it is. Furthermore, only *w* in *F* is lengthened. Therefore, since no Q1 vowel can be transmuted into Q3 by foot assignment, gradation must do without this alternation. Similarly, foot assignment will never change Q1 into Q2.

(VII) *Syllable Conflation*. When a bisyllabic sequence #C \acute{V} CV... turns mono-

syllabic with the loss of the medial consonant, an occasional effect of the Weak Grade (see section 4), the resulting syllable is overlong. Such an initial sequence must be metrically analyzed as #[_F CVCV]. Loss of medial C with attendant melding of syllables gives #[_F_σ CVV] . . . , a canonical environment for (over)lengthening to take place in. The present theory would be hard put to explain a *lack* of overlength under such conditions, and the appearance of Q3 in the Weak Grade thus becomes inevitable, rather than merely puzzling.

That such a variety of effects follows from our single hypothesis is strong confirmation of its value. It is hard to imagine that a segmental treatment of quantity could view this system of relations as anything but a ragbag of phonological odds and ends. Some clarity is achieved when we recognize that the Q2/Q3 contrast differs in nature from the Q1/Q2 contrast; that Q3 is defined on syllables, Q1 and Q2 on segments. But the full range of connections is laid bare only when the Q3 syllable is given its place within the kind of hierarchical theory of phonetic representation assumed here.

4. Toward a General Theory of Quantity for Estonian

In this section, we sketch the role of Q3 in grammar, and move on to consider its phonetics. Going beyond Q3, we examine two other kinds of durational phenomena which turn out to resemble Q3 rather strikingly when analyzed from the present point of view. An approach to unifying the three kinds of “lengthening” effects is proposed, making key use of the constructs of metrical theory.

4.1. *The Grammar of Q3*

Overlength is the servant of grammar, of grammatical category or subcategory, and the occurrence of Q3 is largely governed by *gradation*, the central process of Estonian morphophonemics. In the historically related and more familiar *consonant gradation* of Finnish, the original phonetic basis for the phenomenon remains quite visible, despite some minor “analogical” lapses. When the structure VC_aVC $\left\{ \begin{matrix} \# \\ C \end{matrix} \right\}$ arises inside a word, as through suffixation to a morpheme ending VC_aV, certain changes take place in the character of C_a: geminate obstruents simplify; single obstruents lax, assimilate, or disappear. Morphemes affected by these “weakening” processes are traditionally said to be in the *Weak Grade*; when not, the *Strong Grade*. In modern Estonian, due to sweeping historical changes, there no longer exists any consistent indication of the historical closed-syllable environment for weakening, and the Weak Grade/Strong Grade distinction cannot be phonologically defined. Gradation does remain deeply entrenched in the patterns of inflectional morphology, however, and it is still quite productive, in certain of its forms, with recent loans.

In the singular of gradating words, for example, nominative and partitive cases belong to one grade class, and the remaining cases—genitive, illative, inessive, elative,

allative, adessive, ablative, translative, terminative, essive, abessive, and comitative—all belong to the other.¹⁹ So from the stem /tarkka/ ‘wise, prudent’, we have *tark:ka* (partitive: Strong Grade) but *tarka* (genitive: Weak Grade). Compare the Finnish cognate /tarkka/ ‘exact, precise’, which has *tarkka* (nominative: Strong Grade) and *tarkan* (genitive: Weak Grade), where physical *-n* transparently provokes the weakening.

Gradation in Estonian is generalized to affect both vowels and consonants, and in its centrally productive forms, Q3 alternates with Q1 or Q2. Typical examples of the Q3–Q2 alternation are:

(28)	A. <i>Strong Grade</i>	B. <i>Weak Grade</i>
a.	pat:tu (part. sg.)	pattu (gen. sg.)
b.	kap:pi (part. sg.)	kappi (gen. sg.)
c.	luk:ku (part. sg.)	lukku (gen. sg.)
d.	lin:na (part. sg.)	linna (gen. sg.)
e.	tan:ki (part. sg.)	tanki (gen. sg.)
f.	sil:ma (part. sg.)	silma (gen. sg.)
g.	jät:kama (inf.)	jätkanut (p.p.)
h.	lap:si (part. pl.)	lapse (gen. sg.)
i.	as:tuma (inf.)	astun (1 sg. pres.)
j.	mak:su (part. sg.)	maksu (gen. sg.)
k.	kaa:lu (part. sg.)	kaalu (gen. sg.)
l.	söö:ma (inf.)	söönut (p.p.)
m.	tei:se (ill. sg.)	teise (gen. sg.)
n.	vaa:k:suma (inf.)	vaaksup (3 sg.)

Here is gradation bare of segmental gawdry. Strong Grade is signaled only by the presence of Q3 in the first, main-stressed syllable. For these words, Strong and Weak Grade stems share a segmental representation, and differ suprasegmentally on whether the first syllable is a foot or is merely part of one. The contrast is between (for example) Strong Grade [[_F pat] tu] and Weak Grade [_F pattu].

We know from scrutiny of the stress system that F generally aims to be bisyllabic; it appears, therefore, that a special grammatical condition, proper to the Strong Grade, licenses (indeed, forces) the monosyllabic feet of (28A).

(29) *Basic Grade Principle*

For M (= “word”) in the Strong Grade, the strongest syllable is F.

The BGP (29) is a completely productive law of morphology, spelling out the consequences of category membership. The BGP is given here as a condition on representations, a mere statement of fact; we deliberately avoid, as tangential (if not orthogonal)

¹⁹ This generalization is excerpted from Raun and Saareste (1964, 25). See also Oinas (1975, 331–336) for an extremely useful compilation of data relating to gradation classes.

to our ends, questions about the details of formalization. (Is the BGP a condition on lexical forms? Should it be instead a rule of phonology—"assign F"? Etc.) We will assume that the BGP is true of very early stages of phonological derivation.

Because the BGP establishes a phonological distinction between the grade categories, it opens up the interesting possibility that Weak Grade rules affecting segments are *directly* conditioned by prosodic rather than “grammatical” factors. As it happens, this can be true of only one of the Weak Grade processes: the degemination evident in table (30):

(30)	A. Strong Grade	B. Weak Grade	
a.	paa:t:ti (ill. sg.)	paati (gen. sg.)	'boat'
b.	kəi:k:ke (part. sg.)	kəike (gen. sg.)	'all'
c.	poi:s:si (part. sg.)	poisi (gen. sg.)	'boy'
d.	tark:ka (part. sg.)	tarka (gen. sg.)	'wise'
e.	kimp:pu (part. sg.)	kimpu (gen. sg.)	'bundle'
f.	arkument:ti (part. sg.)	arkumenti (gen. sg.)	
g.	bii:f:fstee:k:ki (part. sg.)	bii:f:fsteeki (gen. sg.)	

Here the grades differ in segmental representation. Each Strong Grade alternant must contain a geminate that serves as the basis for Q3; that geminate corresponds in the Weak Grade to a single consonant. This cannot be the result of gemination in the Strong Grade, because there are minimal structural contrasts that would thwart any such rule. So, with a long vowel in the first syllable, we have e.g. *sii:ti* 'silk, part. sg.' as against *paa:t:ti* (these cannot both be /CVVCV/); with a diphthong, *tei:se* (28m) vs. *poi:s:si* (30c); and with VC-, *tan:ki* vs. *tank:ki* 'tank, part. sg.' (not both /CVCCV/). Simplifying the geminates of (30A) runs smoothly, however, and we are led to assume basic stem-forms such as /paatti/, /käikke/, /poissi/, /tarkka/, etc., all subject to the following rule:

(31) Prosodic Degemination

$C_i C_i \rightarrow C_i / [F : \dots V[+seg] \longrightarrow V, \dots]$

Condition: F is minimal, i.e. does not contain another F .

The condition effectively limits the rule to the Weak Grade, since the geminate will always straddle a foot boundary in the Strong Grade, e.g. [f_LF paat] ti]. The rather clumsy environmental restriction V[+seg] ____ is motivated by the fact that words like (28a-d), e.g. /pattu/, retain the geminate in both grades. (Perhaps the real generalization is that rule (31) cannot change the weight of a syllable; cannot make a complex rime into a simple one.) Prosodic Degemination (31) apparently does not misgenerate, despite the generality of its domain (minimal F); the complex sequences it affects do not, it seems, arise outside of the grading position. Thus (31) may represent a general rule of the language, rather than a morphological peculiarity.

The other weakenings are not similarly graced. Raun and Saareste (1964) label them

“unproductive”; we cannot expect them to apply to new loanwords, neologisms, and so on, even though they do serve to relate quite a few forms in the old vocabulary. Our assumption will be that this set of rules applies only to words marked with the diacritic feature [+A], mnemonic for “archaic”. Because new words enter the lexicon as simply as possible, without unpredictable abstract marks, the class of [+A] items cannot swell.

Prominent among the “unproductive” weakenings is the total assimilation of obstruents /p t s/ to preceding homorganic sonorants. The alternations are *n:t ~ nn*, *m:p ~ mm*, *r:t ~ rr*, *r:s ~ rr*, *l:t ~ ll*. The language lacks *nsV*, *lsV*—in fact, surface *rsV* would come from /rtV/ in an abstract morphophonemics via a minor rule. The sequence [ŋk], present phonetically, always comes from /nk/, without homorganicity. Examples like these are typical:

(32)	A. Strong Grade	B. Weak Grade
a.	lin:tu (part. sg.)	linnu (gen. sg.) ‘bird’
b.	kum:pa (part. sg.)	kumma (gen. sg.) ‘which of two’
c.	kor:ta (part. sg.)	korra (gen. sg.) ‘order, floor’
d.	kər:si (part. pl.)	kərre (gen. sg.) ‘stalk, straw’
e.	sil:ta (part. sg.)	silla (gen. sg.) ‘bridge’

Observe the effects of the ineluctable BGP in column A. Assimilation is as in (33):

(33) *Total Assimilation*

[M _F . . .	V	C	C	V . . .]	. . .
		[+son]	[αplace]		
		[αplace]			
SC:	1	2	3		
SD:	1	2	2		

Condition: F minimal, M = [+A]

Total Assimilation applies before Prosodic Degemination, because assimilated segments will degeminate whenever possible. Stems /keeltu/ ‘prohibition’, /suunta/ ‘direction’, /pöörtu-/ ‘turn (verb)’, when cast into the Weak Grade, do not surface as *keellu, *suunna, *pöörrun, but rather as *keelu*, *suuna*, *pöörun*, with the potential geminate simplified after the heavy syllable, as always.

In the environment *front vowel + l,r — nonhigh vowel, /k/ turns to /j/*. Typical alternants are *mär:ka ~ märja* (‘wet’ sg. part./gen.), *kül:ke ~ külje* (‘side’ sg. part./gen.).

(34) *K-J*

k → j / [M _F . . .	V	[+son]	—	V . . .]	. . .
		[-back]	[-nas]		[-hi]

Condition: F minimal, M = [+A]

Elsewhere, Weak Grade /k/ vanishes after /s,l,r,h,V/. We find, for example, *las:kma*

~ *lasep* ('let' inf./3 sg. pres., < German *lassen*), *jal:ka* ~ *jala* 'foot', *ar:ka* ~ *ara* 'timid', *nah:ka* ~ *naha* 'skin', *mäke* ~ *mäe:* 'hill' (all sg. part./gen.).

(35) *K Deletion*

$k \rightarrow \phi / [M_F \dots [+cont] _ V \dots] \dots]$

Condition: F minimal, M = [+A]

Such part. sg./gen. sg. contrasts as *tupa* ~ *toa:* 'room', *tii:pa* ~ *tiiva* 'wing', *hal:pa* ~ *halva* 'bad' show what happens to nonassimilating /p/.

(36) *Weak P*

a. $p \rightarrow \phi / [M_F \dots _ \dots] \dots // u _ V$

b. $p \rightarrow v / [M_F \dots [+son] _ V \dots] \dots]$

Condition: F minimal, M = [+A]

(Rule (36a) goes both ways: u $_ V$ and $V _ u$).²⁰

Nonassimilating /t/ weakens to *j* in some words, disappears in the rest: *səta* ~ *səja* 'war' (sg. part./gen.), *vea:p* 'he draws' from /veta+p/, cf. inf. *vetama*.

(37) *Weak T*

a. $t \rightarrow j / [M_F \dots V _ V \dots] \dots$ in some stems

b. $t \rightarrow \phi / [M_F \dots V _ V \dots] \dots$

Condition: F minimal, M = [+A]

These last rules tacitly depend upon a feature of the stated BGP that must appear curious when brought to light. To wit: we have Strong Grade [$F_1[F_2 CV] CV$] with a monomoraic foot F_2 , unprecedented and undesired in the stress system at large. Since nonbranching feet are not well-formed, if we persist with this version of the BGP, we shall have to add a clean-up rule that erases the label "F" from the offending syllables:

(38) *Foot Fixing*

$$\begin{array}{c} \overline{} \\ | \\ ``F'' \rightarrow \phi / m \end{array}$$

Of course it is not absolutely necessary to use F as the hallmark of grade in every case; we could refer directly to Weak Grade in the unproductive rules, and limit the BGP so that it only deals with branching feet given by the motivated definition $F \rightarrow u u$ —this would be quite adequate for the productive alternations. Choice between this approach and the one assumed expositarily has no consequences for the thematic issues of this article, and we will not argue it out here.

Glancing back over the rules Total Assimilation, K–J, K Deletion, Weak P, and

²⁰ Cf. Tauli (1973, 36).

Weak T, it becomes obvious that further compactification is possible, even desirable. For example, the rules delete *p*, *t*, and *k* in similar environments. More generally, all of the rules operate in the same minimal foot domain, and all affect the initial consonant—the onset—of the foot’s second syllable. It is reasonable to propose that this commonality be factored out; perhaps the phonology of gradation should be made to follow on a rule such as this:

(39) *Fundamental Gradation*

$$C \rightarrow [+weak] / [{}_M {}_F \dots __ V \dots] \dots]$$

Condition: F minimal, M = [+A]

The further rules would then operate in simplified environments on those segments diacritically marked “weak” by (39).²¹

To achieve our present ends, however, we need only recognize the existence of the *Basic Grade Principle*—that a main-stressed syllable is a foot in the Strong Grade—and appreciate its pervasive effects on the phonetics of grade in Estonian. Details of formal morphophonemics can be safely deferred to another occasion. The theory of phonetic representation that we aim to establish does not crucially depend on commitments to any particular theory of morphophonemics, much less to any particular analysis or rule system. We draw support, here as elsewhere, from the evident pattern of relationships that any analysis must render.

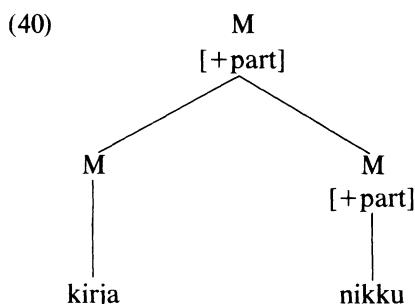
4.1.1. Remark on Place of Alternation. Just one place in an Estonian word will change with grade. The examples produced so far show gradation in (or just after) the first syllable of the stem. But the alternating segment(s) can be found elsewhere in two classes of words: (a) those formed with certain derivational suffixes, e.g. /konta/, /nikku/, /likku/, indeed any suffix ending in /-ikku/; and (b) foreign loans in which a heavy syllable occurs nonnatively near the end of the word, e.g. /arkumentti/, /hulikaani/. In neither case is the grading position correlated with the location of primary stress; stress in the native words is on the first syllable, as always, and may also be there in the loans, e.g. árkument:ti vs. hulikáa:ni. How is the place of alternation to be found?

A crucial observation is that the derivational morphemes of class (a) attach to stems externally, i.e. across “word boundary”. Good low-level evidence of this is the failure of /n/ to go to [ŋ] before the suffix /konta/; place-of-articulation assimilation is the rule inside words (and the only source of ŋ), failing only before /konta/ and /ki/, an (obviously external) clitic meaning ‘also’.²² In our terms, such morphemes can be treated as establishing an M-domain for themselves within the larger M that is the whole word; we will have, for example, [{}_M {}_M kirja] [{}_M nikku]] as the underlying stem-form for the word

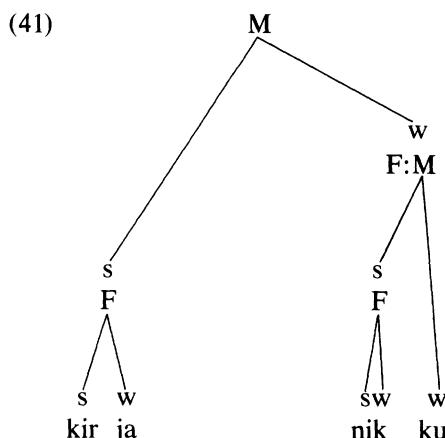
²¹ Rule (39) could also be used to set up degemination. But this cannot be correct synchronically (although it is clearly the case that the diachronic analogue of (39) lies behind all the historical segment weakenings). Degemination has been preserved as a live thing; the other effects, those attendant upon (39), are stuffed and mounted in the lexical museum.

²² Cf. Tauli (1973, 13).

'writer'. (Contrast [_M kirja+u+tta+ma], with internal affixation, for *kirjuttama* 'to write'.) Now, the BGP needs an assist before it can inform us that the partitive singular of 'writer' is *kirjanik:ku*, not **kir:janikku* or **kir:janik:ku*. Aid comes in the form of the notion "head of construction", as extended to morphology by Edwin Williams. Given a word in the partitive case, say, one might naively assume that all of its constituent wordlets ought to inherit the morphological perquisites that come with the whole word's case. In fact, only the "head"—typically the rightmost subword—actually does so. Compare the facts of compound inflection in English: [schoolchild] + Pl emerges as *schoolchildren* = [school] [child+Pl]; only the head of the compound, *child*, bears the tangible mark of plurality. Similarly, in Estonian, only the final *M*-constituent, the head, shows the phonological marks of grade category. With Williams's mechanism of feature percolation to the head, the partitive case of /*kirjanikku*/ is represented morphologically as in (40):



The morpheme /nikku/ is Strong Grade when partitive, invoking the BGP, resulting in the prosodic structure (41):



The reader is referred to Williams (1979) for detailed development of the theory.

Foreign loans like /arkumentti/, /hulikaani/, /akronoomi/, run afoul of narrow restrictions on the phonological form of words. Natively, only the first syllable of a stem may contain VV or VCC; and stems are no longer than three syllables, bisyllables being by far the most common. We conclude that inflexible phonological desiderata force an internal analysis upon the offending loanwords, despite the fact that they are not composed of identifiable morphemes, giving [_M arku [_M mentti]], [_M huli [_M kaani]], and so on. (We put aside the question as to whether *arku-*, *huli-* should be crypto-prefixes, as shown here, or full M.) In such structures, the “phonological words” *mentti* and *kaani* will necessarily be interpreted as “heads”, and will therefore be the sole inheritors of grade category, just as in example (40). It is interesting to note that Hint, working within assumptions similar in spirit to our own, has observed that “the words like *arkumentti* are parallels to . . . native morphological compounds. . . .” (Hint (1973, 210)). It appears that the analogy with compounds (and semicomounds like *kirjanikku*) is taken quite seriously.

4.1.2. Q3 Off the Beaten Path. Lehiste (1966) finds instances of Q3 in unexpected places: she perceives the geminate *t*'s of *vas:tamat(:)ta*, *mak:smat(:)ta* as overlong (p. 16). The morphology is /vasta+ma+tta/ ‘abessive case of infinitive “to answer”’, /maksu+ma+tta/ ‘to pay, abessive’. If /-tta/ were well-behaved, we would find merely Q2. Hint (1973) does not accept Lehiste's judgment, but Eek (1975, 10) regards it as “an essential phenomenon of the Estonian language”. He cites the contrasts *popisette* ‘you (pl.) mumble’ vs. *popiset:ta* ‘impers. neg. pres.’, and *oravatte* ‘squirrel, gen. pl.’ vs. *oravat:ta* ‘abessive sg.’. It seems safe to conclude that there is in fact a well-established dialect with these distinctions. What is to be made of them?

Notice first that the phenomenon is morphologically restricted. It appears to be limited to the affixes /-tta/ ‘abessive’ and /-tta/ ‘impers. neg. pres.’,²³ failing to appear with e.g. /-tta-/ ‘causative’ and /-tte/ ‘pl.’. Significantly, the abessive marker behaves syntactically like a postpositional clitic rather than a case affix. Milsark (1971, 28) remarks:

All of the true case affixes are distributed over all of a set of conjoined nouns, and cause agreement of modifying adjectives. [The abessive ending], however, only occur[s] with the last of a series of conjoined nouns; the rest are in the genitive case. Furthermore any modifying adjectives used with nouns in [the abessive] “case” are in genitive case.

From this evidence of syntactic externality, it is but a small step to the conclusion that the abessive clitic can be a “word”—[_M tta]—which attracts preceding CV into itself and has a grade category (Strong). The same representation suffices for the homophonic impersonal negative ending. Eek notes that the pronunciation *oravat:ta* is optional; this

²³ It is not inconceivable that the morphemes are the same, or closely related. “Abessive” = ‘without’ and /-tta/ is limited to the *negative* impersonal.

seems to indicate a vacillation on the part of speakers between the analysis [_M[_M orava][_M tta]], with Q3, and [_M[_M orava] tta], without it.

4.1.3. On the History of Gradation. The original phonological environment for gradation is recoverable because of the transparency of Finnish: in [_σ CVC₀][_σ C_aVC], something happens to C_a. Deeper questions remain, however. Why is this a natural or likely environment for weakening? Why should consonant weakening in one syllable correlate, in Estonian, with the emergence of overlength in the immediately preceding syllable? The centrality of the BGP (29) suggests that the motivation for the segmental processes of gradation may have been prosodic in character.

Suppose that in the proto-language the only bisyllabic minimal feet were of the form [_F σ CVC], or better perhaps [_F σ m̄̄], a formula which would have the same extension because of the lack of long vowels outside the first syllable. (Word-internal VV sequences arising through affixation can be assumed to be bisyllabic at this stage.) This situation could come about in various ways, depending on the disposition of the other feet. For example, all other feet could be simply monosyllabic. In this case, we could imagine that the preproto-language was purely "syllable-timed"—each syllable a foot; transition to the proto-language occurs by grouping σ m̄̄ sequences into feet. Or the preproto-language could have had an alternating stress pattern built from bisyllabic feet [_F σ σ]; the change would be [_F σ m] → [_{F[F} σ] m], which would leave a residue of minimal bisyllabic feet [_F σ m̄̄]. (Notice that in either case we are assuming that [_F m] is possible: the restriction of F to binarity must come later.)

Whatever the facts of other feet turn out to be, as long as bisyllabic feet are appropriately restricted, consonants can commence to weaken in the environment [_F . . . V . . .], F minimal. Thus, the gradation context becomes exactly "inside a minimal foot"; this we recognize as a place favored for weakening in the world's languages. Stops in American English, for example, suffer lenition foot-internally; hence the familiar contrast in a word such as *pity* between tense, aspirated foot-initial [p^h] and lax, flapped foot-medial [ɾ].²⁴ English also distinguishes minimal feet, though not so absolutely as Finnic. A word like [_{F[F} uni] ty] allows [t^h] and [ɾ] freely, while *pit^hy* is quite unnatural. The domain of obligatory laxing is the minimal foot; outside it, optionality prevails.

From this, we conclude that the prosodic account successfully relates the gradation context to known facts about lenition. It also lays the groundwork for the three-way "quantity" distinction. Thus, the two questions we began with are simultaneously answered.

Modern Estonian derives from the proto-language by various segmental depletions—syncope, apocope, loss of final nasals, etc.—but foot-structure is essentially preserved. Standard Finnish, on the other hand, has lost the prosodic distinction be-

²⁴ See Kahn (1976); Kiparsky (1979) for a foot-based account, Selkirk (1978) for another.

tween Weak and Strong Grade: there is no phonetic difference in the first syllable between e.g. *matka* 'trip, nom. sg.' and *matkan* 'gen. sg.'. A reasonable conjecture is that in Standard Finnish every stressed heavy syllable is a foot. The contrast between $[_F[_F \text{CVC}] \text{CV}]$ and $[_F \text{CVC CVC}]$ has been leveled in favor of $[_F[_F \text{CVC}] \sigma]$, hence no Q3—or rather, no Q2/Q3 distinction; the prediction is that syllable-final consonants (and long vowels) will correspond to Q3 in Estonian. (See section 4.3.5 for more discussion of the foot in Finnish.) In sum, then, our proposal is that Estonian is segmentally innovative, but prosodically conservative, while Finnish is just the opposite.

The idea that prosodic factors of one kind or another are connected with gradation appears in various forms in traditional Finnic linguistics.²⁵ More recently, a novel approach to gradation has come from John McCarthy (1976). His idea is that the crucial motive for lenition is the *ambisyllabicity* of C_a in VC_0C_aVC ; in place of our condition on foot-structure, he has a rule ambisyllabifying C_a backwards to the preceding syllable. (The technical apparatus is from Kahn (1976).) For McCarthy, Weak Grade devolves from syllable structure; for us, from foot-structure. Although the two proposals are conceptually similar, they are not, surprisingly, equivalent. In a *sequence* of closed syllables $\text{CV}-C_a\text{VC}-C_b\text{VC}-C_c\text{VC}-\dots$, every onset C_a, C_b, C_c could ambisyllabify and then weaken by gradation. But no minimal foot can contain C_a, C_b , and C_c . At best, we would have $[_F \text{CV}-C_a\text{VC}][_F C_b\text{VC}-C_c\text{VC}]$, with two feet, where C_b is not in weakening position, even though it is followed by a closed rime. Any foot-domain theory must predict that gradation will alternate as stress does, whereas ambisyllabicity theory allows for the gradation of entire sequences. Which is correct?

In modern Finnish, all dialects, across-the-board application is the rule, e.g. *luke+nto+n* \Rightarrow *luennon* 'lecture, gen. sg.'. But it need not be the case that this reflects the original situation. If there was change of stem-form within paradigms caused by the differences in bisyllabic foot-grouping, we can expect strong pressure to level, especially when the all-important prosodic (foot-structural) distinction between Strong and Weak is lost. If we assume that bisyllabic grouping—foot formation—starts from the *end* of the word, such stem-changes must result. So, from proto **/*luke+nto+n*/ we would get **[_F *lu*][_F *kenton*], which would grade to *[_F *lu*][_F *kennon*]. From **/*luke+nto/* (nom.), we must get **[_F *luknen*]to—the structure *lu*[_F *kento*] is impermissible because it has a bisyllabic foot with a weak open syllable. The outcome here under gradation is *[_F *luen*]to. We end up with a rather complex intraparadigmatic alternation **luento* ~ **lukennon*, which levels to modern *luento* ~ *luennon*.

What we need to know is whether such alternations were actually once part of the language. Evidence in our favor could come from a variety of sources: dialect forms, doublets, etc. If it can be shown that gradation is historically dependent upon the predicted binary grouping, we will have found a subtle and compelling argument for the present theory.

²⁵ I must regretfully disclaim scholarship in this literature. The reader is directed to Eek (1975) and Tauli (1954) for discussion and further references.

4.2. *The Phonetic Expression of Q3*

At the underlying level, the heavy syllables of Estonian come in four basic types: CVV, CVC, CVVC, CVCC. Closer scrutiny of the “superheavy” syllables CVVC, CVCC shows that canonical Estonian adheres rigidly to the principle that sonority declines steadily through the rime. There are no rimes VRR (e.g. *ilm*) or VCC (e.g. *ast*); each step away from the nuclear vowel must be a step down on the ladder of sonority. (All sonorant consonants are held to be equivalent, as are all obstruents.) Notice that, a fortiori, no underlying syllable ends with a geminate. For most words, these findings are adequately summarized in the following formula:

- (42) *Canonical Rime Template*
 $V ([+son]) ([−son])$

The CRT (42) is a well-formedness condition constraining the input to the system of phonological rules.

There are a number of exceptions to the CRT having the form VVR, e.g. /suunta/ ‘direction’. Perhaps these represent a generalization of the CRT in the direction of pure relationality: even in VVR, there is decrease in relative sonority as we proceed outward from the first V. See Kiparsky (1979) for a relational theory of sonority that may be more appropriate than the (categorical) feature notation of (42). There are also a few forms like /kaartte/ ‘card’ with rime VVRC. These we regard as unsystematic exceptions to the CRT.

Phonology blurs the surface picture somewhat. The chief villains are the morphologically-conditioned rules of vowel deletion. Because of them, we find non-canonical-looking *lin:n* ‘town’, *paa:t:t* ‘boat’, *ilmne* ‘obvious’, *kärt:smik:k* ‘innkeeper’, all nom. sg. from underlying /linna/, /paatti/, /ilme + ne/, /kärtsi # mikku/. The phonological processes lying behind these disturbances of placid simplicity are quite clear, and the basic structural generalization of the CRT (42) is unsullied.

How then is Q3 expressed? In a rime VC, of course, it is C that is lengthened. With VV, we must examine the behavior of diphthongs to see how the components of a cluster fare under lengthening. According to K. Ojamaa, who has reviewed all relevant literature in both English and Estonian, “there is general agreement that with overlong diphthongs it is the second segment that is lengthened in Q3 words. . .” (Ojamaa (1977, 37)). And indeed, it does not require the knowing ears of a native speaker to hear the second component as clearer, tenser, longer in Q3 than in Q2. Phonetically, of course, things are not so categorical as perception leads us to believe. Lehiste (1970) has measured spectrograms and found that in Q3 *both* components are longer than in Q2. With reference to Lehiste’s results, McCarthy (1976, 5) writes, “Her measurements show that the increase in duration of the second component is about 6 csec., or 100%. For the first component it is 4 csec., or 50%. These figures suggest that the lengthening of the first component is a secondary effect.” McCarthy’s point emerges

quite clearly in table (43), constructed from data in Lehiste (1970). It shows the increase in length characteristic of several vowels, depending on whether they are first (V_1) or second (V_2) in a diphthongal cluster.

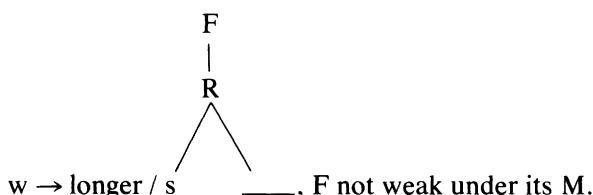
(43) *Percentage Increase Q3(V_1V_2) over Q2(V_1V_2)*

	as V_1	as V_2
i	22%	91%
e	54	123
a	36	68
u	66	113

On the average, the vowel in first position increases 40%, while the vowel in second position increases 87%, fully twice as much. Put another way, we find that in Lehiste's data the ratio V_1/V_2 is 3:2 in Q2 (i.e. 57:43), and 1:1 in Q3 (i.e. 49:51). We therefore agree with McCarthy that Lehiste's measurements support the standard view that V_2 is the locus of Q3 effects. We conclude that VV lengthens essentially the same way that VC does.

In the "superheavy" rime structure VVC, lengthening of both VV and C is reported in Lehiste (1966), giving VV:C:. From this behavior, McCarthy infers that a rime VVC consists of subconstituents V-VC, and that the rule of lengthening-in-Q3 applies not to segments per se but rather to the second (or weak) constituent of the rime as a unit. In our terms, McCarthy's rule becomes (44):

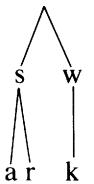
(44) *Q3 Lengthening*



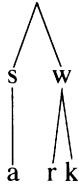
As noted above, the variant realizations VV:C and VVC: are allowed, even preferred, by some speakers (cf. Eek (1975)). This could be the result of a nonstandard version of syllable structure: [_r VV-C] for VVC:, and [_r V-V] C, with C outside the rime, for VV:C. It could also come from a different manner of "percolating" the feature of extra length down from the weak mora node to its constituent segments.

Some light may be shed on the issue of variation by the behavior of VRC rimes. With stems like /tarkka/, we find Strong Grade *tark:ka*, and never **tar:k:ka* or **tar:kka*. But from rule (44), we might well expect **tar:k:ka* to be the correct analogue of e.g. *poi:s:si*. We might suppose that we are witnessing a difference in syllable structure, (45a) instead of the expected (45b).

(45) a.



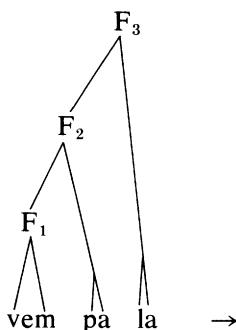
b.



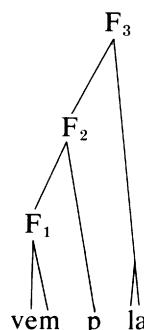
But similar effects show up with rimes VVR; from /suunta/ ‘direction’ comes *suu:nta*, not **suu:n:ta* or **suun:ta*. The apparent generalization is that a sonorant *m n l r* lengthens when it constitutes the entire weak mora, and resists lengthening when merely a proper subpart of it. Thus, the expected structure V-RC, V-VR provides a sound basis for describing the facts, and no single alternative structuring of VVR, VRC will serve as well. Rimes can be uniformly right-branching, without further complication. For speakers who realize VVC as VVC:, such as the one whose productions are measured in Eek (1975), we can assume that the restriction on sonorant-lengthening has been appropriately generalized to vowels.

A final class of illuminating examples is provided by words like *ən:tsa* ‘happy’, *kan:tle* ‘Estonian zither’, *vem:pla* ‘cudgel’, all gen. sg. In each case, the lengthening rime has the shape VRC, as in *tark:ka*, but the Q3 form is VR:C, unexpectedly. A little prospecting in the lexicon shows that the underlying stems are /əntisa/, /kantele/, /vempala/—cf. *ənnis*, *kannel*, *vemmal* (all nom. sg. and Weak Grade). The first rime of the stem—*ən*, *an*, *em*—turns out in each case to be VR, leading to [F VR:] in Q3, as desired.²⁶ We conclude that underlying syllabification is what determines the expression of Q3. We do not, however, conclude that the rule Q3 Lengthening (44) is an early rule, applying before the morphologically conditioned vowel-droppings. Rather, we hold that (44) is a late rule of phonetic interpretation, operating on the phonological surface. To allow for this, we need merely make the minimal assumption about vowel-dropping; namely, that it drops a vowel and has no further consequences. Consider the stem /vempala/. Dropping the first *a* does not affect the constitution of the crucial first rime *-em*, if there is no special rule to resyllabify *p* into it; hence *vem:pla*. More graphically:

(46) a.



b.



²⁶ There is a pleasant and illuminating contrast between *tönt:si* ‘dull, part. sg.’ from /töntsi/ and *ən:tsa* ‘happy, gen. sg.’, from /əntisa/.

The fact that constituent F_2 is adjudged a syllable in (46b) certainly follows from the universal definition of “syllable”, and requires no Estonian-specific convention or rule. Lengthening of *vem* to *vem*: follows from application of rule (44) to representation (46b).²⁷

4.3. Toward a General Theory of Quantity in Estonian

4.3.1. Introductory. There are two other durational phenomena that have been made familiar by describers of the language: the consonant of “ambiguous quantity” and the “half-long vowel”. These are “phonetic” developments, and quite uncoupled from grammar. But when construed from the point of view of metrical theory, they reveal a surprising kinship with Q3—a kinship that is hard to conceive under nonmetrical assumptions. If the proposal sketched here is correct, metrical theory will form the basis for a novel, interestingly generalized conception of the linguistics of quantity and duration.

4.3.2. Ambiguous Quantity. Work reported in Lehiste (1966) implies that a syllable-(or rime-)final consonant C' is significantly longer when it is also *foot*-final, i.e. when in the structure $[_F \sigma CVC']$. Thus, we expect to find *astun* ‘I enroll’ with foot-medial *s*, but *pimes'tav* ‘blinding’, with a longer *s* foot-finally. Lehiste’s concern is not in fact with clusters in general but principally with single and geminate obstruents *p*, *t*, *k*, *s*, which she refers to as consonants “in” a certain degree of quantity. In *pattu* we have *p* in Q1, *t* in Q2; in *pat:tu*, *t* in Q3. Lehiste finds that not all long consonants can be comfortably assigned to a unique quantity grade, either on the basis of refined native-speaker impression or on the basis of raw measured duration. For example, in the word *huvittav* ‘interesting’, she calls the *tt* “ambiguously long”, because it does not fit well into Q2 or Q3. In this it contrasts with *tt* of *parantamattu* (= $[_F \text{paranta}][_F \text{mattu}]$) ‘incorrigible’, which clearly belongs to Q2. What we have notated *CC* is “ambiguously long” when it ends a foot, as in $[_F \text{huvit}]\text{tav}$, but decisively “in Q2” when foot-medial. We generalize the claim to all clusters, and with a nod to the careful transcriptions in Raun and Saareste (1964)—e.g. *armas'tat* ‘you love’ (p. 18)—we propose that the “ambiguous” lengthening affects any foot-final consonant.

Lehiste does not, of course, use the notion “foot”. Although she implies (1966, 39) that “rhythmic structure” is significant, she presents her claim in terms of whether

²⁷ An alternation like *puntar* ~ *punt:tra* ‘bundle, nom./gen. sg.’ might seem to call our account into question. Observe, however, that the Weak Grade nominative is not **punnar*, as would certainly be derived from putative */*puntara*/. The underlying form is therefore /*puntara*/ . In the nominative, we have $[_F \text{punt} \text{ta}] \text{ra} \rightarrow [_F \text{punt} \text{ta}] \text{r}$ (vowel drop) $\rightarrow [_F \text{pun} \text{ta}] \text{r}$ (Prosodic Degemination (31)). In the genitive, which is Strong Grade, we get $[_F \text{punt}] \text{tara} \rightarrow [_F \text{punt}] \text{t} \text{ra}$ by vowel drop, yielding *punt:tra*. Alternations like *nätra* ~ *nät:ra* ‘weak, part./gen. sg.’, from /*nätera*/, are similar.

It is worthwhile noting that our analysis does *not* depend on applying the BGP prior to the rule(s) of vowel deletion, as is tacitly assumed in the text. All we really need is that there be no resyllabification backwards into the preceding rime. As long as the underlying onset of the second syllable stays out of the first rime, it cannot be affected by the foot-based (hence: rime-based) rule of Q3-Lengthening (44).

neighboring syllables are odd or even, counting from the beginning of the word, and whether the preceding syllable is overlong or not. However, whenever stress placement diverges from strict binary alternation, as in the words *vütsakkuséttust* ‘impolite, el. sg.’, *párantamáttu* ‘incorrigible’, she finds that it is what we call foot-position that counts, not arithmetic (cf. p. 14). Furthermore, it is clear that the overlong syllable enters the calculation only because it can disturb the binary flow of stress, being a foot itself. Lehiste’s description, though couched in different terms, is consistent with our own.

Interestingly, Lehiste also regards the *t:t* of such words as *paa:t:tu* as being of “ambiguous duration”. She writes (1966, 40):

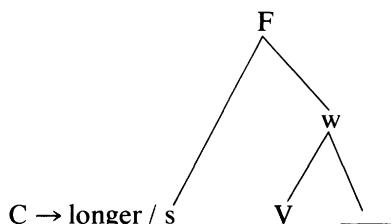
A comparison of words containing plosives in ambiguous quantity between the second and third syllable with words whose first syllables contain a vowel in quantity 3, followed by a consonant in ambiguously long quantity, suggests that at the level of word patterns, the two short syllables of such words as *koguta* [= *kokutta*/ASP] or *huvitav* [= *huvittav*/ASP] are rhythmically equivalent to the overlong first syllables of such words as *saata* [= *saa:t:ta*/ASP] and *kooki* [= *koo:k:ki*/ASP]. The actual manifestations of the plosives in words with these two patterns are not only similar with regard to average duration, but also in respect of range, scattering of values, and the occasional presence of spikes indicating rearticulation.

However, she sees no way to relate the structures CVCVC and CVVC preserving both intuitive accuracy and formal directness. If the geminate of *huvittav* is assigned to Q3, the universally perceived association of Q3 with stress will have to be abandoned. If, on the other hand, the geminate is assigned to Q2, then Lehiste reasons as follows (1966, 41, 42):

In words such as *huvitav* [= *huvittav*/ASP], it would become necessary to explain the realization of /t/ in quantity 2 with reference to the preceding disyllabic sequence, which would entail cumbersome rules: /t/ is realized as having a relatively longer allophone in quantity 2 after a disyllabic sequence whose first syllable is in syllabic quantity 1 or 2, and as having a relatively shorter allophone of quantity 2 after a disyllabic sequence whose first syllable is in syllabic quantity 3.

But a change of description is a change of world. Cumbersome reference to sequence and quantity resolves into the simple notion “foot-final”. And for us, even the canonical instances of the “Q3 consonant” are already reflexes, as it were, of “Q2”, i.e. clustering. We arrive at something like the following rule:

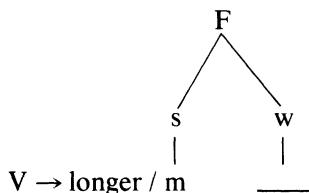
(47) “*Ambiguous*” Lengthening



Native speaker reluctance to identify the geminate of, say, *huvittav* as Q3 should be no bar to rule (47). It appears that the Q3 syllable basically impresses the perceiver as somehow having “heavier” stress than normal, rather than as having this or that segment longer.²⁸ Judgments of segment quantity seem to derive from the syllabic impression. When the elements of the complex percept are separated, it is not surprising that uncertainty should result.

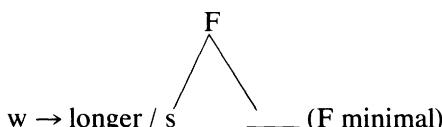
4.3.3. The “Half-Long Vowel”. Careful phoneticians of Estonian have traditionally observed that short vowels are pronounced longer in some positions than in others; the lengthened variant is called the “half-long vowel”. In the word *sekamini* ‘topsy-turvy’, for example, the even vowels are consistently about 50% longer than the odd ones: *seka'mini* is how it’s said. Lehiste (1965b) sets out to ascertain the exact phonetic conditions under which the half-long vowel appears. Her finding, interpreted in the present light, is that a short vowel in absolute foot-final position is lengthened if the first syllable of the foot is short (monomoraic). In pictures:

(48) *Half-Lengthening*



4.3.4. A Generalization. The three quantity-expressing rules (44), (47), and (48) all subserve the same goal, lengthening of *w* in the structure [F s w]. It is not too rash, then, to suppose that there is a single rule at play here, not three, and that the one rule fixes the basic temporal pattern of the Estonian foot for all of its uses.

(49) *Foot Timing*



Rule (49) makes obvious a certain arbitrariness that has been glossed over so far: we could just as well write “*s* → shorter”. But perhaps neither is really correct. What might be the status of the terms “longer”, “shorter”? We have written the quantity rules in a familiar notation, but we have already abandoned some of the familiar ideas linked to that notation—in particular, the idea that quantity is represented phonologically

²⁸ Tauli (1954, 8–9) writes: “It is also remarkable that to Estonians without any linguistic training and with but a scant knowledge of grammar, a syllable in the third degree of length seems to have a stronger stress than the same syllable in the second degree. . . .” It does not take a native ear to hear things this way: the impression of “heavier stress” is a vivid one, and certainly comports with full foot status.

as a feature of segments. Should we reintroduce the feature [\pm long] at the phonetic level? It is not clear that this will be very helpful. Commonly, rules of phonology function to adjust (or determine) the phonetic category that a segment belongs to; with the phenomena controlled by Foot Timing (49), the issue is not category—though it is tempting to talk as if it were—but rather temporal realization of a completely categorized structure. Now, there is a sizable distance between the quantized descriptions of systematic phonetics and the milliseconds of acoustical measurement. The gap must be spanned, in part, by a theory of duration that takes systematic phonetics as input and makes temporal sense of its units. Foot Timing—or the generalization it labors to express—may well belong beyond systematic phonetics, in the theory of duration, where it can directly govern the apportioning of abstract time. If so, the phonetic code will not contain the macrons, colons, and flying dots we have used diacritically throughout this article. As useful as they are in communicating impressions from linguist to linguist, these will play no role in the calculus of durations, and therefore no systematic role at all. The phonetic representation of what we write as *pat:tu* will just be [_F_L_F *pat*] *tu*.²⁹

The idea of Foot Timing (49) is that there is a single bed of Procrustes onto which all feet, whatever their make-up, are temporally stretched. In arriving at this as the terminus of our account, we have, of course, pushed certain evident details aside. Rule (49) does not indicate that “half-lengthening” takes place only after a monomoraic syllable; nor does it restrict “ambiguous lengthening” to affect only the *C* of [_{r:w} VC]; nor does it distinguish the monosyllable foot of, say [_F_L_F *pat*] *tu*] from that of [_F *kava*][_F *latt*] (cf. (3d)). Rules (44), (47), (48) are stated so as to handle these matters, and with the aid of a certain amount of propositional logic, the relevant conditions could be attached to rule (49) as well. But this would merely compromise the integrity of (49). If the idea has much to it, these apparent distinctions of detail among the various subcases of Foot Timing ought to find an explanation, or disappear entirely in the light of new understanding. We leave (49) clean, as a spur to the further research it demands.

4.3.5. Finnish, etc. The prosodic system of Finnish is obviously very similar to that of Estonian. Finnish main stress falls always on the first syllable, and the pattern of secondary stresses closely resembles the one given for Estonian above, with some differences of detail. Interestingly, it appears that Finnish also follows Estonian in some of the quantitative subtleties we have just been considering (though, of course, Finnish lacks the three-way length contrast). For example, Eek (1975, 37) reports, “J. Laurosela has shown that in the South-Pohjanmaa dialect of Finnish, [long?] vowels and diphthongs preceding a closed syllable . . . are shorter than before a following open syllable. K. Donner has pointed out the same phenomenon in the Salmi dialect.” The environment for this purely phonetically-conditioned dialectal analogue to the Q3/Q2 distinction is just the one assumed in the historical account of gradation in section 4.1.3 above. In [_F_L_F CVV] CV], the first syllable—a foot—is not surprisingly longer than in “Weak

²⁹ Details of foot-internal structure omitted.

Grade” [_F CVVCVC]. Furthermore, Lehiste (1965b) produces measurements establishing that the “half-long vowel” appears in Finnish under the same conditions as Estonian. In the word *lákánánani*, for example, Lehiste finds vowel durations 74/107/84/116/69 msec., which are strictly comparable to those reported for Estonian in Lehiste (1968), with the “half-long vowel” once again on the order of 50% longer than the pure short one.

Carlson (1978) contains facts that are even more striking. Carlson observes that “emphatic” pronunciation involves a characteristic pattern of lengthenings. Consider the following examples, drawn from the vocabulary of swear-words:

(50)	i. <i>Normal</i>	ii. <i>Emphatic</i>	iii. <i>Gloss</i>
a.	vitut	vitu:t!	membra muliebria naturalia!
b.	vittu	vit:tu!	membrum muliebre naturale!
c.	paska	pas:ka!	stercus!
d.	perkele	per:kele!	diabolum!
e.	saatana	saa:tana!	satanam!
f.	mulkku	mul:k:ku!	membrum maris naturale!

We recognize this without hesitation as a version of Foot Timing: lengthen *w* in [_F s *w*]. If the first syllable is monomoraic, it is emphasized by elongating the second syllable’s rime; in a complex first syllable, the entire weak mora (cf. *muL:K:ku*) is expanded.

A plausible inference is that Finnish and Estonian use the same timing pattern to realize the minimal foot: a relatively shorter or fixed strong unit is followed by a relatively longer weak unit. In calm discourse, this gives rise to the half-long vowel (and no doubt much else besides).³⁰ *Emphasis* is parasitic upon the mundane, and makes the already long even longer. This is to be expected. Trubetzkoy somewhere observes that when a long-short contrast exists, only the long elements are free to be extended for expressive purposes. The contrast need not be phonemic: in English, for example, the vowel of a word like *node* may be drawn out limitlessly, but not that of a word like *note*; here the relevant contrast is induced by the very low-level rule of lengthening-before-voiced-consonant.

From the point of view of the standard wisdom about English, the Finnic foot looks slightly peculiar: stress should correlate immediately with lengthening, we feel, and *s*, not *w*, should be the greater member. Why might the Finnic languages resist a straightforward lengthening-under-stress? Carlson considers the related question of why only *w*

³⁰ A basic prediction would be that in Standard Finnish, the heavy stressed syllable—which is a foot in every case, by the conjecture of section 4.1.3—should closely resemble the Q3 syllable of Estonian in durational structure. Impressionistically, this seems very plausible, and the relevant data should not be too difficult to obtain through standard techniques. As a side-note of potential interest, we observe that the *t* of, say, Finnish *matka* ‘trip’ appears to be much longer than the *t* of English *Atkins*, which is consonant with the idea that Finnish has “lengthen *w*” and English does not.

is expanded in “emphatic” feet, and makes the following proposal (1978, 7):

These facts [of emphasis] have a natural functional explanation. Given that tense [= emphatic] articulation is one in which articulatory targets are reached in full or even exaggerated, and that Finnish has a clear and pervasive quantity distinction between short and long segments, tense articulation cannot lengthen short segments, but at best shorten them. On the other hand, quantity distinctions are neutralized in the syllable coda, and this can be used for expressive purposes.

Like most functional accounts, this one is relevant principally to the problem of evolution—it explains why a certain formal rule could have flourished in the system; it does not supplant the rule, or even predict its effects in every case. (In diphthongs V_1V_2 , for example, there is no quantity contrast in the V_1 position, yet under emphasis only V_2 is lengthened; the formal rule applies here, without specific functional motivation.) Since the emphatic style is based on prior phonetic givens, Carlson’s reasoning must more properly be applied to the evolution of the durational system itself. Following him, we can argue: if stress is to be expressed by a characteristic timing of foot-constituents, then only w and not s can be lengthened, for longer s would run afoul of the short/long distinction in an impressive number of cases. (It might be necessary to add that w can always be expanded without fear, since there are no long vowels after the first syllable.) Stating the argument this way immediately raises a number of questions. Is the Finnic foot-pattern widespread in languages with pervasive quantity distinctions, as we might expect? Does it appear in languages without quantity distinctions per se? (If not, why not?) Is there a contrasting foot-type “lengthen s ” (e.g. English)? If so, how can length stand for stress (strength) in one case and lack of stress (weakness) in the other, with never a hint of confusion? Are there other patterns of foot-internal timing besides the two that are obviously suggested by the [sw] formalism? Is *foot* always the domain of a timing pattern? We are of course asking for data that may not be available, construed in a theory of duration that certainly does not exist at present. But the theory may be worth constructing, the data worth finding: for these are the natural questions that arise out of the metrical approach to phonetic representation.

A final speculation on the Finnic foot: observe that $[_F \sigma - CVC]$ comports rather well with the “long w ” pattern, inasmuch as the foot’s w -syllable is already “long” by virtue of its complex rime. Now, we suggested above that the ultimate source of the gradation environment was a prosodic structuring of the word in which the only bisyllabic minimal feet were of the form $[\sigma - CVC]$. Rejection of the foot $[\sigma - CV]$, with its “short” w -syllable, may be a kind of structural instantiation of the basic “long w ” pattern. In the subsequent history of the language, the foot $[_F \sigma - CV]$ asserted (or reasserted) itself, perhaps because the requirement of binarity was placed on F , a natural concomitant of stress-timing, but “long w ” (or “short s ”) survived in the detailed realization of these F -structures.

5. Other Views

In this section we will consider (and, alas, reject) two alternative theories of Q3: not merely because we wish to give due respect to our predecessors, but also because their work illuminates the goals and achievements of metrical theory in a particularly direct way.

5.1. Lehiste

In a series of important studies, I. Lehiste has developed the idea that the distribution of the three quantities is based on a phonological hierarchy of segment, syllable, two-syllable unit, and word. There is an obvious similarity here to the terms of our own description, but Lehiste conceives of the hierarchy very differently from the way we do, and, correspondingly, makes very different use of it.

Most striking, perhaps, is that her bisyllabic unit—the analogue of *foot*—is constructed from the arithmetic of odd/even and never explicitly related to claims about stress. (Perhaps this reflects an acoustic phonetician's unease with the acoustically hard-to-define notion "stress".) A major argument for the bisyllabic unit is derived from its role in a proposed constraint that restricts the distribution of phonemic quantities. The essence of the proposal (cf. e.g. Lehiste (1965a)) is this:

- (51) In a sequence $syl_{2n-1} syl_{2n} \dots$, the second syllable may be overlong only if the first one is.

Constraint (51) aims to generalize from the facts of first and second syllables. It is clear that if the second syllable is a foot, the first must also be, by the general laws of stressing; and, contrapositively, if the first syllable is merely part of a foot, the second must belong to the same foot. We have, in the first case, $\#[_F \sigma][_F \sigma] \dots \#$ and, in the second, $\#[_F \sigma \sigma] \dots \#$. Beyond the first two syllables—notice that we are guaranteed that the initial syllable begins a foot—the utility of arithmetic becomes obscure. Words can be found which do not conform to (51), e.g. *nominattii:vi*, *akkusattii:vi*, in which the third and fourth syllables stand in an anti-Lehistan relationship. More significantly, it turns out that there are, in fact, no cases where (51) interferes with the predictions of the foot-based account; no cases where overlength fails to surface in an even syllable that morphology and prosody require to be overlong. At best, then, where the constraint is correct, its claims follow directly from the interpretation of overlength in terms of foot-structure. Once overlength is defined appropriately, the need for (51) vanishes. We conclude that the very real insight behind (51) is best construed in terms of the notion "foot", and that Lehiste's "bisyllabic unit" is a reasonable approximation thereto, as near as one can get without acknowledging prosody.

Even if Lehiste's hierarchy were de-arithmetized (prosodized, as it were), its role in phonetic description would remain fundamentally different from ours. Lehiste conceives of the hierarchy in what we might call a "bottom-up" fashion. At the lowest

level, segments are classified as Q1, Q2, Q3, and perhaps “ambiguous Q”; syllables can then be classified as Q1, Q2, or Q3 on the basis of the *segments* they contain (notice that constraint (51) refers to “syllabic quantity”); and, finally, words can then be classified on the basis of the *syllables* they contain. Quantity information, like time itself, flows in one direction only. Lehiste (1978) contains a clear statement of the program:

I interpret the quantity system of Estonian as building from segmental quantity to the quantity of disyllabic units by going through two levels at which a reduction of information takes place. Going from the segment to the syllable, there is a many-to-one mapping of segmental quantity into syllabic quantity as a result of which the number of contrastive quantity structures is significantly reduced. Going from the syllable to the disyllabic sequence, a further reduction in information takes place . . . (p. 79)

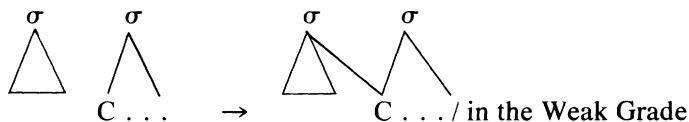
Our view, on the other hand, is that each level in the hierarchy contributes independently, that higher-order categories directly determine phonetic properties that have been (mis-)interpreted as evidence for a richer set of lower-order categories (segments). The segmental distinction “single vs. clustered” (= Q1 vs. Q2/Q3), needed in any case, compounds with the prosodic distinction “foot vs. subpart of foot” (\Rightarrow Q3 vs. Q1/Q2), needed in any case to describe stress patterns, and forth from this marriage the particulars of quantity come. Because of interaction, the theory of categories remains quite minimal.

Behind Lehiste’s theory we can perhaps discern a kind of Structuralist phonemics, with “separation of levels” projected into phonology itself; and along with it, a view of speech perception as a necessarily bottom-up operation, segmenting the acoustic signal into phones, phonemes, syllables, groups of syllables, etc., in that order. But just as the present approach challenges the “separation of prosodic levels” within phonetic description, so does it implicitly challenge the bottom-up view of speech perception. If the perceiver’s goal is to recover the phonetic representation [F[F pat] tu], for example, then identifying foot-structure is one of the primary tasks, on a par with segmentation (into segments). Notice that higher-order analysis need not wait until lower-level analysis is complete; if, as we have suggested, the *foot* of Estonian has a stereotypical temporal pattern, it may well be possible to extract valuable clues (cues?) to foot-structure from rather gross durational information; this in parallel with or even in advance of detailed segmental classification. With Liberman (1978), we venture to suggest that hierarchical structure may figure very directly in speech perception. And not just perception: a theory of production based on control of hierarchical structure will be quite different from one that deals with a simple concatenation of segments. The change of linguistic notation we have argued for turns out to be—perhaps surprisingly—a change of *notion* that could renovate our understanding of speech processes.

5.2. Resyllabification

For McCarthy (1976), as for us, overlength is of suprasegmental provenance. Taking a cue from the path-breaking work of Kahn (1976), he proposes that a change in syllable structure establishes the essential contextual distinction. The rule is this:

(52) Ambisyllabification



(We have considered this rule as a candidate source for gradation in section 4.1.3 above.) Overlengthening can now take place in syllables that end crisply; those joined to what follows by the link of ambisyllabicity are unaffected. On this view, then, we have $\hat{p} \hat{a} : \hat{t} u$ but $\hat{p} \hat{a} t u$. Because the theory correctly portrays overlength as syllabic in origin, it will have many of the desirable entailments. Its major shortcoming is that it does not relate quantity to the stress system, which is, in our view, the central fact of the situation. As soon as the role of prosodic structure is acknowledged, the argument for Ambisyllabification must dwindle. What function could remain to ambisyllabicity per se? It survives only as an intermediary between cause and effect, unwanted and unnecessary. To the notions “segment”, “syllable”, “foot”, and their rules of combination, nothing need be added to arrive at a general understanding of the quantity phenomena.

It seems worthwhile to try to extend this result. To ambisyllabify is to resyllabify, to make a special adjustment in syllabic composition. (McCarthy's rule could as well move the affected consonant entirely back into the first syllable.) In general, such adjustments function to set up distinctions that various segment-modifying rules depend upon. Given the success of metrical theory in treating apparent cases of rules-dependent-on-resyllabification as being, more straightforwardly, rules of the foot-domain, we are led to suggest that special rules of resyllabification simply do not exist. This is no trade-off of one notation for another: the idea of “foot-domain” is intrinsically more contentful than that of “resyllabification rule”, since foot-structure is predetermined by the stress system, while resyllabification, of itself, could be just about anything.

It is also true that resyllabification has a strong intuitive basis as well as a grammatical function. American speakers are not very happy with the division $\hat{p} \hat{i} - t y$, for example, and Kahn's $p \hat{i} t y$ has real appeal. But it is a mistake to too hastily assume that speakers have a privileged, unconfounded insight into the exact location of syllable boundaries. Foot, not syllable, is the principal determinant of English phonology—

where, indeed, does *syllable* do work for us?—and so it is not surprising that foot boundaries can be found unerringly, and syllable boundaries found best when they coincide with foot boundaries.

One form of resyllabification that does seem unavoidable occurs across the word boundary in many (if not all?) languages: C#V comes out as [_oCV . . .]. Kiparsky (1979) proposes that this results from reapplication at the phrasal level of the (universal) rule requiring syllables to have onsets whenever possible. Sharpening our hypothesis, then, we claim that, aside from reassertion of the basic laws, there can be no readjustment of syllable structure.³¹

6. Conclusion: “neuer Anfang”

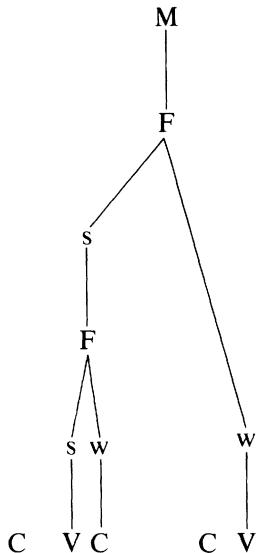
Conceived most broadly, this has been an inquiry into the empirical extension of metrical prosodic theory. A possible (if uninvited) reading of Liberman and Prince (1977) would be that it offers—merely—a pleasant new algorithm for distributing the familiar phonetic properties of stress (0,1,2,3, . . .). A deeper possibility is that metrical theory involves a fundamental revision in the notion of phonetic representation: it changes our ideas about what the “familiar phonetic properties” are.

Major support for this second view comes from the evident centrality of the notion “foot”. Fundamental to stress, it is also the cornerstone of the quantitative edifice. As we have seen, the quantity system is essentially the product of multiplying a segmento-syllabic distinction (heavy/light) by a prosodic distinction (foot/nonfoot). Further, “foot” plays a role in gradation, and may well have been (historically) the primary determinant of the process. Our result is that the effect of metrical structuring is felt not only in the domain of stress, where the theory originated, but equally in the quantity system and in the system of segment-modifying rules.

A second supporting result is the generalization relating Q3 to the “half-long vowel” and the “consonant of ambiguous quantity”: *lengthen w in F*. It appears that there is a single stereotypical pattern for the timing of all minimal F. From this we wish to draw the further conclusion that the implementation of “quantity” lies within the durational system—the transduction into (and from) speech—and not in the phonological system per se. The phonetic code is the hierarchical structuring of simple segments without our heuristic diacriticism, so that the phonetic representation for what we have written as e.g. CVC:CV is essentially as shown in (53):

³¹ The reader is directed to Feinstein (1979) for further discussion of resyllabification, along with a proposal that may require some slight modification of the principle we have given.

(53)



If the code manipulated in speech processes really does have this character and not, as is often thought, that of an alphabetic concatenation, our approach to perception and production will be commensurately modified.

Finally, we wish to emphasize that the central argument of this article depends on very few premises, is almost crudely simple. We have developed a fair amount of conceptual apparatus beyond our minimal needs, though: not only for purposes of descriptive concreteness, but also because the proffered hypotheses and devices are natural within the metrical perspective and therefore worthy of exploration. For example, we treat final syllables as stereoisomers of normal syllables. And we describe the Estonian stress pattern in terms of well-formedness conditions, rather than procedural rules of assignment ex nihilo, and in so doing dispense with iteration and directionality of application. These proposals, and many others presented here, could fade in the ultimate light of truth, and our argument survive. But some, or all, could flourish in that same light, so let them stand forth to face it.

References

- Bing, J. M. (1980) "Linguistic Rhythm and Grammatical Structure in Afghan Persian," *Linguistic Inquiry* 11.3 (this issue).
- Carlson, L. (1978) "Word Stress in Finnish," ditto, MIT, Cambridge, Massachusetts.
- Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.
- Eek, A. (1975) "Observations on the Durations of Some Word Structures: II," *Estonian Papers in Phonetics* 4, 7-53.
- Feinstein, M. (1979) "Prenasalization and Syllable Structure," *Linguistic Inquiry* 10, 243-278.

A METRICAL THEORY FOR ESTONIAN QUANTITY

- Halle, M. and J.-R. Vergnaud (1978) "Metrical Structure in Phonology," ditto, MIT, Cambridge, Massachusetts.
- Harms, R. (1962) *Estonian Grammar*, Indiana University Press, Bloomington, Indiana.
- Hint, M. (1973) *Eesti Keele Sõnafonoloogia I*, Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut, Tallinn.
- Jones, D. (1972) *An Outline of English Phonetics*, Heffer, Cambridge.
- Kahn, D. (1976) *Syllable-based Generalizations in English Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Kiparsky, P. (1978) "Issues in Phonological Theory," in J. Weinstock, ed., *The Nordic Languages and Modern Linguistics*, vol. 3, University of Texas Press, Austin, Texas.
- Kiparsky, P. (1979) "Metrical Structure Assignment is Cyclic," *Linguistic Inquiry* 10, 421-441.
- Leben, W. (1977) "Estonian Quantity," ditto, MIT, Cambridge, Massachusetts.
- Lehiste, I. (1960) "Segmental and Syllabic Quantity in Estonian," *American Studies in Uralic Linguistics*, 21-82, Indiana University, Bloomington, Indiana.
- Lehiste, I. (1965a) "The Function of Quantity in Finnish and Estonian," *Language* 41, 447-456.
- Lehiste, I. (1965b) "Juncture," *Proceedings of the Fifth International Congress of the Phonetic Sciences*, S. Karger, Basel/New York.
- Lehiste, I. (1966) *Consonant Quantity and Phonological Units in Estonian*, Indiana University Press, Bloomington, Indiana.
- Lehiste, I. (1968) "Vowel Quantity in Word and Utterance in Estonian," *Congressus Secundus Internationalis Fennno-Ugristarum*, Societas Fennno-Ugrica, Helsinki.
- Lehiste, I. (1970) "Diphthongs vs. Vowel Sequences in Estonian," *Proceedings of the Sixth International Congress of the Phonetic Sciences*, 539-544, Munich.
- Lehiste, I. (1978) "The Syllable as a Structural Unit in Estonian," in A. Bell and J. B. Hooper, eds., *Syllables and Segments*, North-Holland, Poughkeepsie, New York.
- Liberman, M. (1975) *The Intonational System of English*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Liberman, M. (1978) "Prosodic Structure and Speech Recognition," talk delivered at first Sloan conference on "The Mental Representation of Phonology," November 1978, Amherst, Massachusetts.
- Liberman, M. and A. Prince (1977) "On Stress and Linguistic Rhythm," *Linguistic Inquiry* 8, 249-336.
- McCarthy, J. (1976) "On Hierarchic Representations within Syllables," ditto, MIT, Cambridge, Massachusetts.
- McCarthy, J. (1979a) "On Stress and Syllabification," *Linguistic Inquiry* 10, 443-465.
- McCarthy, J. (1979b) *Formal Problems in Semitic Phonology and Morphology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Milsark, G. (1971) "Toward an Analysis of Quantity Alternations in Estonian," ditto, MIT, Cambridge, Massachusetts.
- Nespor, M. and I. Vogel (1979) "Clash Avoidance in Italian," *Linguistic Inquiry* 10, 467-482.
- Oinas, F. (1975) *Basic Course in Estonian*, Indiana University Press, Bloomington, Indiana.
- Ojamaa, K. (1977) *Temporal Aspects of Phonological Quantity in Estonian*, Doctoral dissertation, University of Connecticut, Storrs, Connecticut.
- Prince, A. (1976) "'Applying' Stress," ditto, University of Massachusetts at Amherst, Massachusetts.
- Raun, A. and A. Saareste (1964) *Introduction to Estonian Linguistics*, Ural-Altaïsche Bibliothek, Wiesbaden.

- Sapir, E. (1930) *Southern Paiute, a Shoshonean Language* (Proceedings of the American Academy of Arts and Sciences 65, nos. 1-3), New York.
- Selkirk, E. (1977) "French Schwa," ditto, University of Massachusetts at Amherst, Massachusetts.
- Selkirk, E. (1978) "Part II: The Phonological Structure of English," ditto, University of Massachusetts at Amherst, Massachusetts.
- Selkirk, E. (1980) "The Role of Prosodic Categories in English Word Stress," *Linguistic Inquiry* 11.3 (this issue).
- Solan, L. (1979) "Spanish Stress," in J. Lowenstamm, ed., *University of Massachusetts Working Papers in Linguistics*, vol. IV, Amherst, Massachusetts.
- Tauli, V. (1954) "The Origin of the Quantitative System in Estonian," *Journal de la Société Finno-Ougrienne* 57, 1-19.
- Tauli, V. (1968) "On Quantity and Stress in Estonian," *Congressus Secundus Internationalis Fennio-Ugristarum*, Societas Fennio-Ugrica, Helsinki.
- Tauli, V. (1973) *Standard Estonian Grammar*, Almqvist & Wiksell, Uppsala.
- Trager, G. (1961) Review of American Studies in Uralic Linguistics, SIL XV, 24-27.
- Trubetzkoy, N. (1969) *Principles of Phonology*, University of California Press, Berkeley, California.
- Vihman, E. (1974) "Estonian Quantity Re-Viewed," *Foundations of Language* 11, 415-432.
- Wheeler, D. (1979) "Stress in Southern Paiute and Tubatulabal," in M. Stein and E. Engdahl, eds., *Studies Presented to Emmon Bach by his Students*, Amherst, Massachusetts.
- Williams, E. (1979) "Morphology," ditto, University of Massachusetts at Amherst, Massachusetts.

*Department of Linguistics
South College
University of Massachusetts
Amherst, Massachusetts 01003*