2

Estonian Quantity: Implications for Moraic Theory

MARTIN EHALA

1 Introduction[1]

Estonian distinguishes three degrees of quantity: short (Q1), long (Q2) and overlong (Q3). Q1 occurs in short stressed syllables (CV) like the first syllable of *koli* 'trash', and Q2 in long stressed syllables (CVV and CVC) as, for example, the first syllable of *kooli* 'school' g. sg., and *kolli* 'ghost' g. sg. Q3, like Q2, occurs also in long stressed syllables, but with an extra prosodic quantity – the so-called "overlength" – added. This "overlength" is traditionally transcribed by a lengthening sign after the segment that carries its peak: *koo:li* 'school' p. sg., or *kol:li* 'ghost' p. sg. As the minimal triplets *koli – kooli – koo:li* and *koli – kolli – kol:li* show, all three quantity degrees are phonologically distinctive.

Furthermore, in Estonian, quantity also influences the stress placement: a Q3 syllable can form a monosyllabic foot of its own, in which case the secondary stress immediately follows the main stress: *sóo:làsele* 'inhabitant of swamp', sg. abl. Neither Q1 nor Q2 can show this pattern: a stressed Q1

[1] I would like to thank Matti Leiwo, Pirjo Turunen, Anthony Fox and two anonymous reviewers for their comments and criticism, as well as CIMO for a research grant at the University of Jyväskylä that enabled me to undertake this research.

Generative Approaches to Finnic and Saami Linguistics.
Diane Nelson and Satu Manninen (eds.).

or Q2 syllable must be followed by at least one unstressed syllable as seen in *élusàle* 'alive', sg. abl. (Q1), and *sóolasèle* 'salty', sg. abl. (Q2).

Phonologically, the ternary pattern of Estonian quantity has been explained by various means. The majority of accounts have characterised overlength either in moraic terms as lengthening in the Q3 syllables (Bye 1997, Eek and Meister 1997), or in metrical terms as the ability of Q3 syllables to exhaust a foot (Prince 1980, Kager 1994). Both approaches have their problems – moraic accounts lead to excessive use of moras (3 or even 4 moras per syllable), metrical accounts retain representational binarity, but at the cost of recursive foot structure.

However, these problems could be avoided if the marked member of the Estonian quantity system is considered Q2 rather than Q3. As I am going to argue in this paper, such a revision is well motivated both on empirical and theoretical grounds as well as by the evidence from the historical evolution of the Estonian quantity system. Besides this, the Estonian quantity system also provides support for the hypothesis that syllable weight is not a linear function of segmental length. This has direct implications to Moraic Theory (Hayes 1989, 1995) which uses the notion of mora to encode both length and weight distinctions.

2 Phonetic Properties of Estonian Quantity

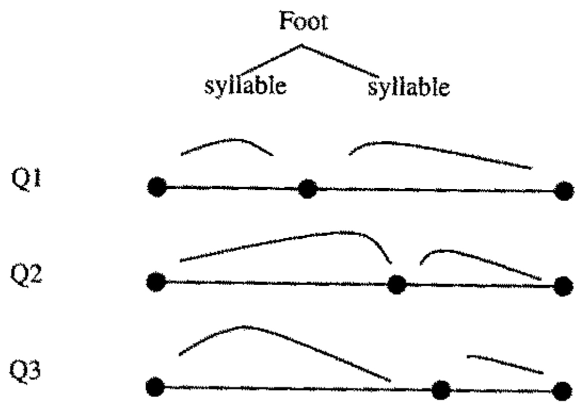
More than a half century of experimental phonetic research on Estonian quantity has proved that it is impossible to derive the Estonian three quantity degrees from the corresponding durational differences of sounds (see Lehiste 1997). Perception experiments have also shown that the acoustic information contained in a single syllable is not sufficient to distinguish the three quantity degrees. Eek and Meister (1997) conducted an experiment to study whether listeners are able to identity the quantity degree of the first syllable of a disyllabic word without the information of the second syllable vowel. The listeners were able to identify the short quantity (Q1), but were not able to differentiate between Q2 and Q3. Only when the listeners were presented information about the quantity of the second syllable vowel were they able to identify Q2 and Q3.

This suggests that the durational ratio of the syllables in a disyllabic foot is the main indicator of quantity in Estonian. However, as Ilse Lehiste (1997) reports, the duration ratios alone are not sufficient to distinguish Q2 and Q3. In a perception experiment, listeners were presented "disyllabic" noise bursts which corresponded in their durational ratios to disyllabic Q1 (2:3), Q2 (3:2) and Q3 (2:1) sequences. For the sake of symmetry 1:2 sequence stimuli were also included. The subjects were able to distinguish the short-long (1:2 and 2:3) sequences from long-short (2:1 and 3:2) sequences, but were unable to differentiate further 1:2 sequences from 2:3 sequences and 2:1 from 3:2. The experiment was replicated by synthesised disyllabic nonsense word stimuli. The results were similar to the previous experiment.

One of the features differentiating the Estonian Q2 and Q3 feet is the fundamental frequency curve (Lehiste 1997, Eek and Meister 1997). In a Q2 disyllabic foot, F0 curve reaches its peak on the end of the first syllable vowel and starts to fall on the second syllable vowel, but the fall starts higher than in Q3 feet. In Q3 feet, the F0 peak is achieved in the first third of the first syllable vowel and is followed by a steady fall that continues also on the second syllable. However, the difference in the F0 curve is insufficient to distinguish between Q2 and Q3 syllables, if the second syllable is not present.

This has lead phoneticians to a conclusion that Estonian quantity degrees are perceptually identified in the domain of a disyllabic foot by a number of interrelated phonetic cues, the most important of which are the durational ratios of syllables in the foot and the fundamental frequency curve. I have represented the patterns of Estonian three quantity degrees in (1) where the length of the lines marks the durational ratio of syllables in the foot and the curves above them indicate the contour of F0. As the Estonian disyllabic feet strive to isochrony, I have represented all of them as durationally equal, though the Q1 foot is the shortest, Q2 foot (surprisingly) the longest and Q3 in between, closer to Q2 than Q1. The durational differences of different foot types, however, are smaller than the durational differences between intra-feet segments (Eek and Meister 1997).

(1)



3 Phonological Accounts of Estonian Quantity

The oldest theory of Estonian quantity comes from Veske (1879). He argued that all Estonian phonemes have three distinctive degrees of length. This theory was dominating in Estonian linguistics about a hundred years, but has gradually lost its currency as experimental phonetic research has made it apparent that the Estonian quantity degrees do not correspond to durational differences of speech sounds. Currently, most of the linguists agree that the Q2-Q3 distinction is a prosodic one. However, there is quite a wide range of opinions about what is the domain of Q2-Q3 distinction and how to express it phonologically.

As a ternary opposition of quantity is typologically rare and theoretically unwelcome, the most common phonological solution has been to separate the three-way distinction to two binary oppositions: segmental length distinguishes between short (Q1) and long (Q2) quantity, and an extra prosodic quantity distinguishes overlong (Q3) from Q2. This is most clearly expressed in (Hint 1997a: 133), repeated here as (2):

(2) SYLLABLES

- **Segmentally**: SHORT
  - **Prosodically**: stressed (Q1)
  - unstressed
- **Segmentally**: LONG
  - **Prosodically**: stressed
    - without prosodical quantity (Q2)
    - with prosodical quantity (Q3)
  - unstressed

This hypothesis defines the Estonian quantity opposition as a syllabic phenomenon.

Alternatively, it is suggested that Estonian quantity must be interpreted phonologically as a feature of the whole foot (Eek and Meister 1997): Q1 and Q2 feet are balanced, and Q3 feet are unbalanced. The terms balanced/unbalanced refer directly to the phonetic differences of feet: the articulatory energy appears more evenly distributed in Q1 and Q2 disyllabic feet than in Q3 feet (see (1)). Schematically this can be presented as (3):

(3)

	FEET		
Prosodically	balanced		unbalanced
Segmentally	short	long	long
	Q1	Q2	Q3

The difference between (2) and (3) lies in the prosodic domain where Q3 is to be phonologically represented. In (2) it is the syllable, in (3) it is the foot. It should be noted that the difference between balanced and unbalanced feet have been so far expressed only by moraic means (Eek and Meister 1997), not much different from those which use moras to represent overlength in syllabic terms (for example Bye 1997). Thus, so far there is no precise representational difference between the syllable-based and foot-based accounts of Estonian quantity.

What is common to these approaches is that both of them assume that Q3 is the marked member of the Estonian quantity system. In moraic terms, this is commonly represented by an extra mora assigned to the syllable, making Q3 syllables trimoraic (/koμμμ.liμ/ 'school' sg. p.). To avoid theoretically unwelcome trimoraic syllables, some authors (Bye 1997, Eek and Meister 1997) have suggested that overlength is to be represented either as a free mora or a degenerate syllable.

Though both of the approaches (2 and 3) manage to account for the patterns of Estonian quantity at the prosodic level of representation, their problems become apparent at the intersection of segmental and prosodic quantity. I will give two short examples based on the analyses of Bye (1997) and Eek and Meister (1997) to demonstrate this.

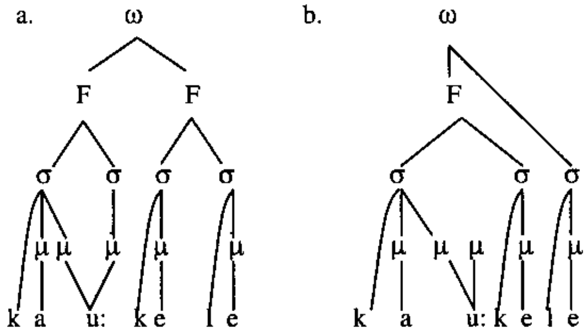
According to Bye (1997), overlength in Q3 syllables is variably expressed either by a degenerate syllable or a free mora. This allows him to explain why Q3 syllables may constitute feet of their own (*káu:kèle* – /[kau:][ke.le]/ 'far'), but need not (*káu:kele* – /[kau:.ke]le/). If Q3 forms a foot of its own, it is represented as a sequence of a syllable and a degenerate syllable (4a); if not, the overlength is expressed by an unassociated mora (4b):

(4) a. $\sigma_{\mu\mu}+\sigma_{\mu}^{degen}$

b. $\sigma_{\mu\mu}+\mu$

Bye (1997) does not specify, how the moras should be connected to the segments, but he assumes that the third mora is connected to the Q3 centre (indicated by lengthening sign ':' in traditional accounts). Thus, I suppose that the two representations for the word *kau:kele* would look like this:

(5)

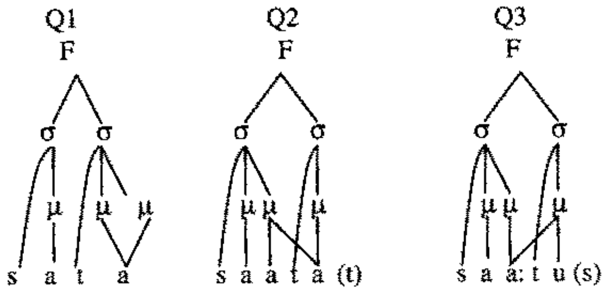


In (5a), the degenerate syllable analysis enables parsing the first syllable and the degenerate syllable as a foot, which allows secondary stress to fall on the second syllable: *káu:kèle*. As the free mora on (5b) is not associated to a syllable node, monosyllabic foot is not possible, which leads to a stress pattern *káu:kele*.

This analysis is correct for most of the cases. However, the question is whether degenerate syllables and free moras are indeed theoretically more welcome than trimoraic syllables. They might, provided that one can dispense trimoraic syllables altogether, but under the conventions of Moraic Theory (Hayes 1989, 1995), the description of Estonian would still require trimoraic syllables, even if Q3 is represented by a degenerate syllable or a free mora. Let us consider two forms: Q2 *aitta* (a name) and Q3 *ait:ta* ('barn' sg. p.). If geminate consonants and short vowels are underlyingly monomoraic (Hayes 1989), the representation of the name *aitta* (Q2) would require three moras in the first syllable (/aμiμtμ.taμ/) though it does not have overlength. Provided that the third mora could be analysed as a degenerate syllable (/aμiμ.tμ.taμ/), this would lead to an unattested stress pattern in trisyllabic forms of this word, such as **áittàle* sg. abl. Furthermore, the Q3 form *ait:ta* ('barn' sg. p.) would then require an additional, fourth mora /aμiμtμμ.taμ/ to distinguish it from Q2 name *aitta* that already has three moras in the first syllable. There is no satisfactory answer, how to accommodate quadrimoraic syllables with the analysis proposed by Bye (1997). Of course, one might also adopt the principle of mora-splitting (Maddieson 1993), but the three devices together (degenerate syllables, free moras and splitting) would seriously reduce the attractiveness of the theory allowing for far more possible combinations than there are attested syllable types.

Eek and Meister (1997) have tried to solve the problem without the notions of degenerate syllables and free moras, just relying on the concept of mora splitting. In their interpretation, all disyllabic feet contain three moras which are distributed to replicate as close as possible the phonetic patterns of Q1, Q2 and Q3 disyllabic feet (see (1)). This would necessarily require splitting moras between the syllables:

(6)



The examples in (6), taken from Eek and Meister (1997: 94) show the representation of Q1 word *sata* 'hundred', Q2 word *saatat* 'you escort' and Q3 word *saa:tus* 'product'. The phonetically half-long vowel in the second syllable of Q1 feet is represented as bimoraic, the difference between long *saatat* and overlong *saa:tus* is expressed by different splitting patterns: in Q2 feet, the first syllable gives some of its weight to the second syllable, and in Q3 vice versa. This correctly expresses the phonetic facts that in Q2 feet, the second syllable vowel is often half-long and in Q3 feet it is extra-short. However, the whole analysis is unacceptable due to crossing association lines.

It might be argued that the problems of combining prosodic and segmental quantity in moraic accounts of Estonian overlength are due to the principles of Moraic Theory, but one should admit that other approaches do not do better.

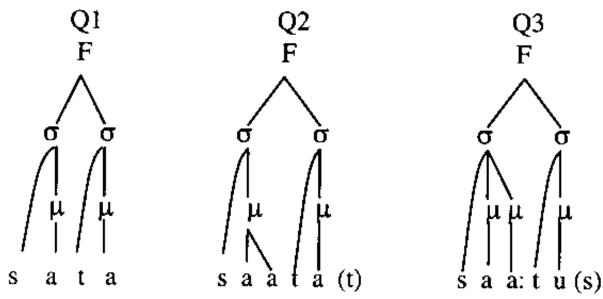
For example in treatments of Prince (1980) and Kager (1994), overlength is accounted for by the ability of Q3 syllables to form a monosyllabic foot of their own. In these accounts Q3 syllables differ from Q2 ones by their ability to exhaust two metrical grid positions (s w for Prince (1980), x x for Kager (1994)), while Q2 and Q1 can exhaust only one grid position. This explanation would be perfect, if Q3 would always constitute a foot. However, as already mentioned, Q3 syllables may, but need not to form a monosyllabic foot. The two stress patterns of the word *kau:kele* illustrate this behaviour well: in *káu:kèle* (/[kau:.][ke.le]/) Q3 is a foot, in *káu:kele*

(/[kau:.ke.]le/) it is not. If the only defining feature of Q3 is its ability to exhaust a foot, the description of forms like *káu:kele* would require recursive foot structure: /[[kau:.]ke.]le/, to avoid two extrametrical syllables. Such a treatment is at odds with the principle of Prosodic Hierarchy (Selkirk 1984, McCarthy and Prince 1986), according to which feet cannot be comprised of feet. Hayes (1995) who adopts essentially the analysis of Prince (1980) overcomes this by stating that trimoraic syllables may optionally be treated as disyllabic sequences. As in his analysis Q3 syllables are trimoraic, the foot exhaustion feature is not defining for Q3 and need not be absolute.

Above I have shortly overviewed some accounts of Estonian quantity, none of which is fully satisfactory. The problems of all these explanations stem from a common assumption that the marked member of the opposition is the overlong Q3. Actually, as we will see in Section 5, the analysis of the historical evolution of the phenomenon suggests that the ternary quantity opposition has not emerged due to over-lengthening of the Q2, but instead by shortening of Q3.

This is the crucial point of my analysis. Provided that the marked member of the opposition is Q2 (historically shortened Q3), one could consider Q1 a normal short and light (monomoraic) quantity and Q3 a normal long and heavy (bimoraic) quantity. Q2 would be something in between: segmentally long, but light by weight (monomoraic). According to this hypothesis, the examples in (6) would have a representation as in (7):

(7)



As we will see in Sections 6 and 7, this analysis would significantly simplify the treatment of both the Estonian quantity as well as the stress system. However, the same analysis would run counter to one of the main principles of the Moraic Theory, namely that syllabic quantity is a linear function of segmental length (shortly discussed in the next Section). The main theoretical argument of the paper is that the representation of weight and segmental length should be separated. This will be discussed in detail in Section 8.

4 Segmental Length and Syllable Weight in Moraic Theory

In the Moraic Theory (Hayes 1989, 1995), the notion of mora has a dual role: it is the basic unit of syllable weight and, at the same time, it is used to express segmental length distinctions. The last is achieved by the following set of conventions:

(8) a) short vowels have one mora
b) long vowels have two moras
c) short consonants are moraless
d) long consonants and geminates have one mora
e) glides are moraless
f) segments in syllable onsets are moraless

From such conventions, it follows that syllable weight is a linear function of segmental length. This means that CV syllables are always light (i.e. monomoraic) and CVV syllables always heavy (bimoraic). CVC syllables are light by default, but in languages where CVC counts as heavy, the coda consonant is assigned a mora by the rule of Weight by Position (Hayes 1989).

For the description of most languages, the double role of moras (to express both weight and length) does not constitute problems. However, complications arise in languages that have geminate consonants, but count CVC syllables as light (see Tranel 1991). If geminate consonants are underlyingly moraic, the CVC syllables closed by geminates are necessarily heavy (bimoraic) while the rest of CVC syllables are light. One possibility to solve the controversy is to separate the representation of segmental length from that of weight, as suggested by Tranel (1991) or to use an extra tier of moras (Hayes 1995).

Complications arise also in languages where syllable weight is not a linear function of segmental length, such as Estonian. As already mentioned, in Estonian, syllabic quantity is not calculated from the moraicity of its constituents (except Q1). For example Q2 syllable *sep* in *seppa* ('smith' sg. g.) is bimoraic according to (8), but Q2 syllable *ait* as in *aitta* (a name) is trimoraic. Interestingly, Q3 syllable *ait:* in *ait:ta* ('barn' sg. p.) is also trimoraic, but Q3 syllable *sep:* as in *sep:pa* ('smith' sg. p.) is bimoraic following (8). This problem could also be solved if weight is to be represented differently from that of length.

The conventions in (8) create also unnecessary superfluity in phonological descriptions. If syllable weight is calculated on the basis of segmental length, all languages that have a distinction of segmental length should also have a distinction of syllable weight. This needs not necessarily be so. Consider an analogy of tones: tones are usually represented in phonological descriptions of these languages that make use of them. The languages that do not use tones do not have a tonal tier. Why should languages that do not use syllable weight, have weight obligatorily expressed in their phonological representations? Certainly it would be more appropriate to express weight distinctions only in these languages where it is relevant. This could be done if encoding the length distinctions could be separated from encoding the weight distinctions.

There are several proposals for how to do it. The most common of these is the original proposal of Heinrich Hock (1985) to use CV-tier to express length distinctions and μ-tier for weight distinctions. Tranel (1991) and others repeat the same proposal. Recently Fox (1998) has suggested to adopt a separate Q-tier which is to be placed between the syllable and foot levels.

The main argument against all these solutions is that this would require an extra representational grid, which would make the theory too unrestricted. This objection can be eliminated if we abandon the CV-tier and represent length distinctions by counting the root nodes (Selkirk 1990). According to her proposal, geminates and long vowels have one set of feature values that are linked to two root nodes, the feature values of single segments are linked to one root node.

It is often argued against separating weight from length, since it would allow more association combinations between segmental length and syllable weight than actually attested, such as bimoraic short segments and monomoraic long segments (Hayes 1989: 293). I think these worries are slightly exaggerated. Concerning bimoraic short segments, I leave the discussion for time being. As about the possibility of monomoraic long segments, these are attested in some languages and the strongest evidence for them comes from Estonian, where they occur in Q2 syllables. This claim is supported by the evidence from the historical evolution of the Estonian quantity system to which I turn below.

5 Historical Evolution of the Estonian Quantity System

There are several hypothesises concerning the historical evolution of Estonian quantity (for an overview, see Tauli 1954). The most widely accepted of them is the Veske-Collinder's theory (Veske 1879, Collinder 1929) which explains the emergence of Q3 by compensatory lengthening after syncopy and apocopy in about XIII–XV century.

In the course of syncopy and apocopy in Estonian, short vowels were lost under certain conditions in the word final position and in the middle of the word. The conditions were as follows:

1. The short vowel was lost from the unstressed open syllable if the preceding stressed syllable was long. Under this condition, short vowels were lost both in the absolute end of disyllabic words (*jal:k* < **jalka* 'foot') and in second syllables of longer words (*tüt:ret* < **tüttäret* 'daughters', *an:tnut* < **antanut* 'given').
2. The word final short vowel in trisyllabic words was lost regardless of the weight of preceding syllables (*orav* < **orava* 'squirrel', *raamatt* < **raamattu* 'book').
3. The word final short vowel in disyllabic words was not lost, if the first syllable was short (*pesa* < **pesa* 'nest').
4. The syncopy and apocopy were not triggered if the syllable concerned was closed (*laulu* < **laulun* <**laulun* 'song' g. sg., *oravat* < **oravat* 'squirrels').
5. Parallel to the syncopy and apocopy, there was also shortening of long vowels and diphthongs in unstressed syllables both in closed as well as in open ones (*met:sa* < **met:san* < **metsaan* 'forest' ill. sg., *met:sa* < **metsaa* 'forest' p. sg.).

Hint (1980) has summarised these conditions elegantly in terms of moras: the vowel of unstressed open syllables was lost, if it was on the distance of two moras from the nearest preceding stress.

According to Veske (1879) and Collinder (1929), the changes mentioned above were accompanied by compensatory lengthening in the immediately preceding stressed heavy syllable. As a result, these syllables acquired overlength whereas the rest of the stressed heavy syllables remained in their original quantity which in the present Estonian system equals to Q2. Thus, overlength evolved in cases like *jal:k* < **jalka*, *tüt:ret* < **tüttäret*, *an:tnut* < **antanut*, *met:sa* < **metsaa*, whereas no compensatory lengthening occurred, if the immediately preceding syllable was stressless (*orav* < **orava*, *raamatt* < **raamattu*).

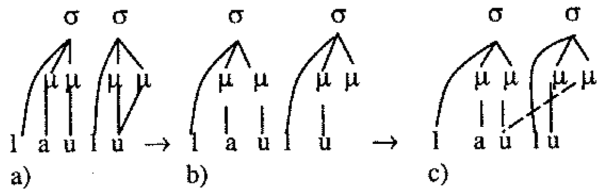
Thought the correlation between the vowel loss or shortening and the modern Q3 in the preceding syllable is quite high, it does not necessarily mean that there was compensatory lengthening involved. First there are a few morphological types where no vowel loss occurred, but the preceding stressed syllables are Q3 nevertheless. One of these types includes verbs of weakening grade alternation such as *ɣp:pima* 'to study'. This type has about 2000 stems which make up nearly one fifth of Estonian verb stems. In these verbs the short vowel from the second open syllable has not been lost, though the first syllable is heavy. Accordingly, there could not have occurred compensatory lengthening, either. Contrary to expectations, the first syllables of these verbs are still Q3 (*ɣp:pima*, *kam:mima* 'to comb', *hoo:lima* 'to care about'). This deviation is commonly explained by analogy with verbs of the type of *tun:tuma* 'to seem' where the Q3 could have emerged by compensatory lengthening (*tun:tuma* < **tuntuuma*, see Kettunen 1929). The problem with this explanation is that the *ɣp:pima*-type is more than twice as numerous than the *tun:tuma*-type that has only about 900 stems. First, it is unlikely that the smaller type has provided the basis for analogy to the larger type. Second, analogical levelling usually does not spread to all cases where it could, but in Estonian there are no verbs of appropriate structure that have escaped this alleged analogy. Thirdly, and most importantly, the *tun:tuma*-type does not have grade alternation: all its forms are in Q3 (*tun:tuma* : *tun:tup* 's/he/it seems'), whereas *ɣp:pima*-verbs have the regular grade alternation pattern, Q2 occurring in forms where the second syllable was historically closed (*ɣppib* 's/he/it studies', Q2), and Q3 where the second syllable was open (*ɣp:pima* 'to study', Q3). If the *tun:tuma* type really had influenced the *ɣp:pima* type, the last would certainly have no grade alternation.

There is a similar problem with the nouns of strengthening grade alternation (*hammas* 'tooth'). This type has Q3 in plural forms, though the original diphthong in the second syllable has not shortened: *ham:pait* < **hampait* 'tooth' p. pl. In this case, it is not possible to appeal to analogy, either, since there is no type that could serve as the basis for it. The question then arises, how could these forms have overlength if no compensatory lengthening has taken place.

Though just two types that do not conform to the compensatory lengthening pattern might not be enough to falsify the theory, I will show later that there is alternative analysis that is able to account also for these deviant types, and is therefore superior to the compensatory lengthening hypothesis.

However, the problems of the Veske-Collinder theory do not end here. While compensatory lengthening after the loss of the vowel is easily modeled in Moraic Theory (Hayes 1989), it is not the case after shortening of the long vowel. The reason is that the syllable node cannot be eroded by Parasitic Delinking (Hayes 1989: 268), since it still contains an overt nuclear segment (8b). And consequently, the loose mora could not acquire a new association without crossing the association lines (8c):

(9)



Of course, it would always be possible to find a way to avoid crossing the lines, but the question is whether it is indeed reasonable to modify the theory to accommodate a type of change for which there is only hypothetical evidence from one language. It is much more reasonable to ask whether the suggested compensatory lengthening took place at all. And there is evidence that it did not.

Let us compare Estonian and Finnish. The languages are closely related, and though they are not mutually intelligible, there are many words that sound and mean the same in both languages. The quantity system of Finnish differs from that of Estonian by lacking the three degrees of quantity that Estonian has developed after the Proto Baltic-Finnic has broken down to separate languages. Instead, Finnish has retained the binary short-long opposition of length which is generally believed to resemble the quantity system of Proto Baltic-Finnic. According to Veske-Collinder hypothesis, this difference between Finnish and Estonian quantity can be explained by the lack of syncopy and apocopy in Finnish – the change that created Estonian Q3 through compensatory lengthening.

If this is the case, then the Estonian Q1 forms should be quantitatively close to Finnish short quantity and the Q2 forms to the long one. The Q3, on the other hand, should be different from that of the Finnish long quantity. While the Finnish short quantity and Q1 are quite close, the Finnish long quantity corresponds perceptually and articulatorly to Q3 rather than to Q2 (see Pajusalu 1994 for details). Similar findings are also reported by Eek and Help (1986: 12–13): "... in contemporary Finnish … the duration of a long vowel or a long ambisyllabic consonant in the main stressed syllable is approximate or even exceeds the duration of the Estonian main stressed syllable with the grave accent (the so-called Q3) …" The results of the phonetic research is supported by the fact that Finnish learners of Estonian do not have difficulties with articulating and perceiving Q1 and Q3, but do find it hard to replicate Q2. All these facts are certainly incompatible with the hypothesis that it is the Q3 that is the late development of Estonian.

This controversy was not unknown to Collinder, either. To explain it, he suggested that there has been an across-the-board lengthening in Finnish long syllables that turned them phonetically similar to Estonian Q3 (see

Ariste 1947 for discussion). This is obviously an *ad hoc* explanation that does not have any independent motivation from the history of Finnish.

Above I pointed out three shortcomings in the Veske-Collinder's theory, for which there is no satisfactory explanation:

1. The lack of vowel loss or shortening in *ɣp:pima-* and *hammas*-types.
2. The impossibility of compensatory lengthening after shorthening of long vowels in the next syllable.
3. Phonetic closedness of Finnish long quantity and Estonian Q3.

Next I will outline an alternative explanation to the emergence of the Estonian quantity system, an explanation which is also able to cope in a natural way with the above mentioned deviations.

This explanation was first suggested by the last century Finnish scholar Emil Nestor Setälä (1896). He argued that the Estonian quantity system dates back to Proto Finno-Ugric and is a part of Lappish and Baltic-Finnic grade alternation. Certainly, Setälä's dating of Estonian quantity alternation so early has raised justified doubts, and at present, his theory has completely lost its currency. However, connecting the origin of the Estonian quantity opposition to the Baltic-Finnic grade alternation has its appeal, and some linguists have also attempted it in the 20th century.

The change which led to the emergence of the Baltic-Finnic grade alternation is believed to have started under Proto-Germanic influence in II millenium BC (see Posti 1953–54). The nature and the conditions of the change were similar to that of the Verner's Law: the second syllable initial stop was weakened if the syllable was closed; if it was open, the stop remained unchanged. As the second syllable was closed in some forms of the stem (such as genitive case) and open in others (partitive), the change created an alternation which in contemporary Baltic-Finnic languages is known as consonant gradation or grade alternation:

(10)	Genitive	Partitive
a.	*jal^{k}an > *jalan > jala (Q1)	*jalkata > *jalkaa > jalka (Q3)
b.	*tark^{k}an > *tarkan > tarka (Q2)	*tarkkata > *tarkkaa > tarkka (Q3)
c.	*sep^{p}an > *sep^{p}an > sep^{p}a (Q2)	*seppata > *seppaa > seppa (Q3)

Above, the stops in genitive case forms weakened, since the second syllable was closed by the genitive ending *-n*. In cases like (9 a and b) the weakening lead to total loss of the stop, in (9c) the geminate was just weakened. (Note that in modern Finnish, the cognate of (14c) has evolved to a full loss of the geminate (Fn *sepän* 'smith' sg.g.), exactly like the cognates of (14 a and b) *jalan* 'foot' sg.g. and *tarkan* 'precice' sg.g.). At a later stage, Estonian lost *n* in a word final position, which lead to a contemporary genitive forms *jala* 'foot', *tarka* 'clever' and *seppa* 'smith'. These are the weak grade forms. In partitive forms, there was no weakening, since the second syllable was open. At the later stage the partitive ending *-ta* shortened to *-a*, and was eventually lost altogether in the course of apocopy. The resulting partitive forms *jalka*, *tarkka* and *seppa* are called the strong grade.

In Finnish, the grade alternation only involves stops; in Estonian, it also includes quantity alternation: Q2 forms are the weak grade and Q3 forms the strong grade. Ariste (1947) has suggested that the Estonian quantity alternation (and the ternary quantity opposition as a whole) has emerged on the analogy of Baltic-Finnic consonant gradation. He argues that the weakening in cases like (9b) and (9c) spread first to words with other geminate consonants (10a) and consonant clusters (10b), and thereafter also to words with long vowels (10c) and diphthongs (10d):

(11) Genitive	Partitive
a. *konnan > ko^{n}nan > ko^{n}na (Q2)	*konnata > konnaa > konna (Q3)
b. *metsan > me^{t}san > me^{t}sa (Q2)	*metsata > metsaa > metsa (Q3)
c. *looman > lo^{o}man > lo^{o}ma (Q2)	*loomata > loomaa > looma (Q3)
d. *laulun > la^{u}lun > la^{u}lu (Q2)	*lauluta > laulua > laulu (Q3)

After this had happened, the initially weakly expressed prosodic alternation strove to become more salient. In this process, both the weak grade weakened even more and the strong one started to lengthen. This lengthening process, as Ariste (1947) claims, caused the loss of partitive case ending *-a*. When this happened, case differences (partitive *versus* genitive) become to be expressed solely by means of quantity alternation. In a sense, Ariste (1947) argues that the quantity alternation emerged due to the system's pressure to find a use for a phonetic variation that has accidentally entered into language.

Hint (1981, 1997) agrees with Ariste (1947) that the consonant loss in the weak grade was accompanied by a prosodic weakening, and that this pattern was taken over to types in (11) by analogy. However, in his opinion it happened much later, in about XVI century, after Estonian has lost its genitive and partitive case endings. After the case endings were lost, the forms of genitive and partitive became homophonic in cases like (11) whereas the forms in (10) were kept apart by the means of consonant gradation. To avoid ambiguity, the prosodic alternation, accompanying grade alternation was taken over as the sole means of differentiating the otherwise homophonous case forms in (11).

Tauli (1954) argues that the weakening in cases like (11) did not emerge due to analogy, but as a part and parcel with the consonant gradation, i.e. as an expansion of the Verner's Law from stops to other segments, too. Though the quantity alternation is not known in other Baltic-Finnic languages (except for some signs of it in Livonian), I agree with Tauli (1954) that it was caused by the same processes responsible for the consonant gradation. There is a good argument for such an interpretation.

First, there is evidence from Finnish dialects that the first consonant in a consonant cluster between the first and second syllable is shorter when the second syllable is closed, than when it is open (Laurosela 1922). The same tendency can be noticed in some varieties of spoken Finnish, and the data in a thorough study of Finnish quantity by Lehtonen (1970: 126–129) indicate that the weakening appears also in standard language though only in case of intervocalic geminates (CVC_GC_GVC). Of course, this weakening in Finnish is by no means phonemic, but just a phonetic particular which is left unnoticed by native speakers. But still, it suggests that the processes that once caused grade alternation (Verner's Law) are still somewhat active in present Finnish. And this can be seen also in present Estonian, for example in the derivational system.

If a disyllabic Q3 stem like *vää:ra* 'wrong' or *pär:li* 'pearl' is added a derivational suffix which closes the second syllable (*-nt*, *-nk*, *-kkas*, *-ntama* etc), Q3 weakens to Q2: *väärant* 'bastard', *pärlentama* 'to shine like a pearl'. Morphologically neither of these stems has quantity alternation, all their forms are in Q3. Even stems which have grade alternation (*anne* : *an:te* 'gift' n. and g. sg.) do not occur with these affixes in their weak grade form (**annekkas*, Q2), but instead in a weakened (Q2) strong grade form (*antekkas* 'gifted'). The norms of the standard Estonian prescribe Q3 in these derived forms (*an:tekkas*), but even efforts of the mother tongue teachers at schools have not had any success in correcting the usage.

These facts suggest that the quantity alternation might well have emerged first as a phonetic phenomenon together with the consonant gradation in Proto Baltic-Finnic, but it only turned phonemic in Estonian after several sound changes caused some vital case endings to be lost, as argued by Hint (1980, 1997). In other Baltic-Finnic languages the alternation was largely levelled out.

Where I do not agree with Tauli (1954) is the emergence of Q3. He argues for a phonetic lengthening in Q3 after the syncopy and apocopy, but not for a compensatory lengthening of the phonemic nature. The existence of a purely phonetic historical change is, however, very hard to disprove. Still, the fact that Finnish long quantity and Estonian Q3 are very close both articulatorly and perceptionally suggests rather that there has been no significant lengthening involved in the emergence of Estonian Q3 syllables.

Assuming that Q2 emerged through weakening rather than Q3 through compensatory lengthening allows us to explain the irregular pattern where

Q3 is not accompanied by apocopy or syncopy (*γp:pima-* and *hammas-*type). First, as I do not assume any lengthening in Q3, I do not assume any causal connection between syncopy/apocopy and the emergence of Q3. Thus, neither the lack of vowel loss or shortening in certain cases or presence of them in other cases should not have had any consequences to the quantity of the preceding syllable: these syllables retained their initial quantity (long = Q3) either cases. And that seems to be so: Q3 occurs invariably in forms that were not weakened under Verner's Law (i.e. in forms where the second syllable was open) regardless of the syncopy or apocopy that might or might not happen in this syllable at the later stage. It is only an historical accident that the syncopy and apocopy later happened in the same environment (in the open second syllable if the preceding stressed syllable was long).

What is more important is that Q2 occurs invariably in forms that were weakened under Verner's Law (i.e. where the second syllable was closed). And this explains why *γp:pima* and *hammas* type still have the regular pattern of quantity alternation, despite the lack of vowel loss or shortening after Q3 syllables: it is not lengthening in Q3 (strong grade) which caused quantity alternation, but weakening in Q2 (weak grade). The fact that *γ p:pima* and *hammas* type have the weak grade regularly in forms where the following syllable was closed, and strong case elsewhere notwithstanding of vowel loss, gives a strong support to this hypothesis. This is also the crucial fact that Veske-Collinder's theory of compensatory lengthening is unable to account for: in *γp:pima* type, the occurrence of Q3 could be explained by analogy to *tun:tuma* type, but not the presence of quantity alternation, since the *tun:tuma* type does not have one.

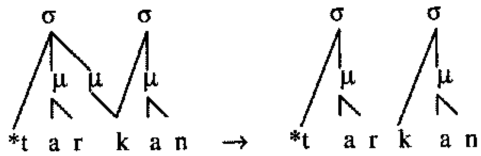
6 Estonian Quantity Reconsidered

The historical evidence, presented above has direct consequences to the synchronic theory of Estonian quantity. The most prominent phonological accounts of Estonian quantity assume that its essence is connected to the notion of overlength, and that it is Q3 which is the marked member in this ternary opposition. However, the evidence from historical evolution of Estonian quantity suggests that it is Q2 that is the marked member of the opposition, whereas Q1 and Q3 remain more or less identical to the Proto Baltic-Finnic short and long quantities respectively. If this is correct, the Estonian quantity system could be analysed without the help of any unusual theoretical concepts such as degenerate syllables or recursive foot structure.

Let us consider first the emergence of Q2 in moraic terms. I assume, following Ariste (1948), Tauli (1956) and Hint (1981, 1997) that the consonant loss in the weak grade of grade alternating words (*tark^{k}an >

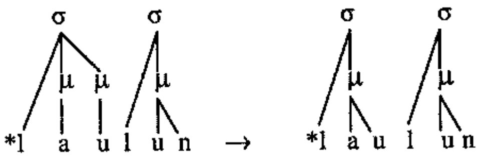
*tarkan > tarka 'clever' sg. g.) was accompanied by a prosodic weakening. Actually, in Moraic Theory (Hayes 1988, 1995) this change is essentially prosodic, since the only feature that distinquishes geminates from single consonants is their moraicity:

(12)



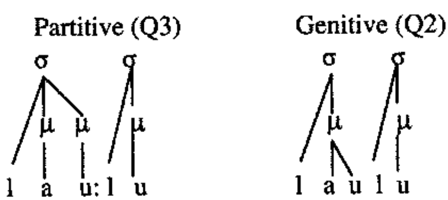
Together with this change, weakening of the first syllable also affected words with other geminates, consonant clusters, long vowels and diphthongs. However, unlike the cases with stops (12), the other types of weakening were purely prosodic, without any segmental losses:

(13)



If the difference between Q2 and Q3 is to be analysed as difference in their moraic count as argued above, one is forced to conclude that the quantity alternation which emerged in cases like (13) made syllable weight phonologically distinctive in Estonian. This becomes apparent if we compare the genitive and partitive forms of the same word *lau:l* 'song'. These forms are distinquished solely by their prosodic difference, partitive *lau:lu* is in a strong grade (Q3), genitive *laulu* in a weak grade (Q2). In moraic terms, it can be expressed as follows:

(14)



Here it should be noted that the lengthening sign in the transcription of Q3 forms is a convention based on the hypothesis of compensatory lengthening in these forms. As the current analysis is based on the opposite hypothesis, it should need a different transcription: Q2 indicated by a shortening sign, for example, *laŭlu*, Q3 without diacritics: *laulu*. However, for the sake of consistency, I will not change the transcription in the middle of the paper.

As the only means that differentiates the case forms in (14) is the weight of the first syllable, it should be concluded, that Estonian has a distinctive syllable weight. Schematically it can be summarised as follows:

(15)

	SYLLABLES		
Segmentally	short (CV)	long (CVC, CVV, CVVC, etc.)	
By weight	light (1μ)	light (1μ)	heavy (2μ)
	Q1	Q2	Q3

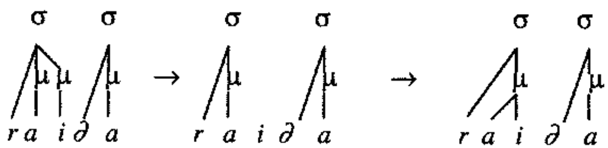
The notion of distinctive syllable weight may appear unusual, but actually it is unavoidable in any phonological accounts of Estonian quantity that describe it in moraic terms. The reason is that it does not matter whether Q2 is treated as light and Q3 as heavy or Q2 as heavy and Q3 as superheavy, the difference between *laulu* g. sg and *lau:lu* p. sg. boils down to the difference in syllable weight.

This interpretation also requires counting long syllables with the same segmental structure sometimes bimoraic and sometimes monomoraic, ie. it requires the acceptance of monomoraic long vowels and diphthongs as well as moraless geminate consonants. This claim is easier to accept if we bear in mind that monomoraic long syllables (Q2) are durationally shorter than corresponding Q3 syllables (see Pajusalu 1994).

Furthermore, there appear also to be other languages that, for stress purposes, count syllables with the same segmental structure sometimes light and sometimes heavy. For example, Hayes (1995: 242) argues, on the basis of the data by Woodbury (1987), that in Norton Sound, a dialect of Central Alaskan Yupik, CVC syllables are heavy only if they are in the initial position in the word. The same phenomenon occurs also in Pacific Yupik (Hayes 1995: 333–334). In Kashmiri, CVC syllable is heavy only if it is the best potentially stressable syllable in the word, otherwise it is light (Morén 1998). In Latin, the CVC syllables are counted as heavy, unless they appear in a weak position in the foot. This can be observed in Latin quantitative verse, as show the following cases provided by Allen (1973: 182–83): ***úter****vostró:rum*, *dé**dit**dó:rno:*, *gú**ber**ná:bunt* (light CVC syllables are presented in boldface). Hayes (1995: 120) accounts for this by a rule of Iambic Shortening which deletes a mora from a heavy syllable if it appears in a weak position in the foot. This rule applies equally to CVV and CVC syllables. In CVV syllables it causes the long vowel to shorten, in CVC syllables no segmental changes follow the mora deletion.

So far we have seen that CVC syllables allow their weight to vary without changes in their segmental structure. Fijian provides examples where this happens also to diphthongs in CVV syllables. Fijian has a rule that shortens long vowels in the penultimate syllable, if the final vowel is short. The same rule also applies to diphthongs. Such diphthongs become phonetically shorter than normal ones and their first components undergo a greater degree of phonetic assimilation with the following segment (Schütz 1985: 545). Hayes (1995: 145–46) calls this rule Trochaic Shortening and treats it as loss of mora. If this rule applies to long vowels they become ordinary short vowels, but diphthongs become monomoraic. For example, addition of the suffix *ða* to the stem *rai* 'see' leads to *răiða* 'see it' with a monomoraic diphthong. Hayes (1995: 145) represents this case as follows:

(16)



In all of these languages, the occurrence of monomoraic long syllables is structurally conditioned, their exceptional quantity is due to the quantity of neighbouring syllables or by their position in respect to the word boundary. Interestingly, there appears to be at least one language besides Estonian that makes use of distinctive syllable weight. It is Cahuilla (see Seiler 1977). In this language, there is a morphophonological rule that intensifies the first syllable in stems beginning with CVCV or CVCCV. In CVCV sequences it manifests as gemination (17a), in CVCCV sequences as lengthening of the coda consonant (17b, examples from Hayes 1995: 139):

(17)	a. *čéxivèn*	'clear'	b. *wélnet*	'mean one'
	čéxxìven	'it is very clear'	*wéllnèt*	'very mean one'

Normally CVC syllables are light in Cahuilla, the intensified ones, however, count as heavy. This can be seen from the secondary stress, which is moved to the second syllable in the intensified forms. Both Seiler (1977) and Hayes (1995) analyse this as mora insertion to the first syllable.

A similar stress pattern is also characteristic to Estonian: weight distinction between Q2 and Q3 syllables is relevant for Estonian stress assignment. As secondary stress can fall to the immediately following syllable, if the preceding syllable is Q3, but cannot, if it is Q2 or Q1, treating Q3 as heavy and Q2 and Q1 as light also simplifies the description of Estonian stress system significantly.

7 Implications for the Estonian Stress System

In Estonian, main stress is generally on the first syllable and secondary stresses follow on a binary or ternary intervals:

(18) *téravàmaltt* or *téravamàltt*

At the end of the word only CVV, CVCC or CVVC syllables can bear a stress, but not CV and CVC syllables. Thus, ternary stress interval is not permitted in cases like **téravamàl* and **sóoyemattèl*. This could be accounted for by assuming consonant extrametricality at the end of a phonological word (Hayes 1995: 322). There is also a tendency to avoid ternary stress if the third syllable in the ternary interval is long. The tendency is the strongest for CVV syllables with diphthongs (long vowels are allowed only in the syllable bearing the main stress), thus *élavàileki* is correct while **élavailèki* is not. If the third syllable is CVC, ternary stress can be used (*válusattèle*), though binary stress is more common in such cases (*válusáttele*).

Estonian stress pattern is further complicated by the fact that Q3 syllables can exhaust a foot. This means that a Q3 syllable may be followed by a secondary stress on the next syllable (*káu:kéle*). However, the requirement to exhaust a foot is not obligatory for Q3, as seen in *káukele*.

The most widely accepted account of Estonian stress is Prince (1980). In his analysis Q3 syllables are assigned two grid positions (strong and weak), all other syllables only one. Since two grid positions exhaust the foot, it allows Prince (1980) to define Q3 structurally as foot.

This gives an elegant account of the stress pattern where a Q3 syllable is followed by a secondary stress on second (19 a) or third syllable (19 b):

(19) a) [káu:][kèle]ki 'far away, too'
[júl:][késse]ki 'bold, too', ill. sg.
[jál:][kàte][lèki] 'foot, too', g. pl
b) [káu:]ke[lèki]
[júl:]kes[séki]
[jál:]ka[tèle]ki

However, in Estonian, secondary stress may fall also to the fourth syllable, like *jál:kadelèki*. Such cases are critical to the analysis of Prince (1980), since they would require two extrametrical syllables between the main and secondary stress: [jál:]kate[lèki]. To account for this and to maintain his claim that the defining property of Q3 is its ability to exhaust a foot, Prince (1980) employs the notion of recursive foot: [[jál:]ka]te[lèki]. This notion has not found independent motivation and is at present abandoned on theoretical grounds.

Hayes's (1995: 316–329) account of Estonian stress adopts essentially the analysis of Prince (1980), but he does not claim that Q3 must always be a foot. Instead he defines Q3 as being trimoraic, and allows trimoraic syllables optionally to be treated as disyllabic sequences. As a disyllabic sequence can constitute a foot, this accounts correctly for the cases like (19 a, b). If the Q3 syllable is not treated as a disyllabic sequence, it makes it possible to account for cases like *jál:katelèki* without the use of recursive foot structure: [jál:ka]te[lèki]. Thus, in Hayes (1995), the correct result is achieved by trading the recursive foot structure for trimoraic syllables and their property to act as disyllabic sequences.

If we treat Q3 syllables as heavy and all other syllables as light, as argued in this paper, it would allow us to adopt the analysis of Prince (1980) without the need for recursive foot structure. The argument is based on the assumption that foot is binary either on moraic or syllabic basis (Prince 1980: 534):

(20) $F \rightarrow u\ u$ (where $u = \sigma, \mu$)

In Estonian, stress is calculated variably on syllabic or moraic basis. This position is well motivated by the fact that Estonian, if it ever has been fully weight sensitive, is shifting back to a syllable-counting stress system. This can clearly seen from the changes in progress in the Estonian morphological system.

In Estonian, affix allomorphy is conditioned by the stress pattern of the word (see Hint 1980, Kager 1995). For example the partitive pl. ending *–it* is possible only in stems where the secondary stress falls on it: *soolase-it* 'salty' p. pl. ([sóola][sèit]), *pae:se-it* 'limestony' p. pl. ([páe:][sèit]). This is due to a phonotactic constraint of Estonian that prohibits diphthongs and long vowels from unstressed syllables. In the word final syllable could not have a secondary stress, partitive plural is formed by vowel alternation: *vallalise* ⟶ *vallalisi* 'single', pl. p. ([válla][lìsi]).

About a half a century ago, trisyllabic words with Q3 in initial syllable allowed only one pattern: *suu:lise* ⟶ *suu:lisi* 'oral' p. pl. ([súu:][lìsi]). Increasingly, such words started to be used with the *–it* ending *suu:liseit* ([súu:li][sèit]), and at present both variants are allowed in literary language. This change can be explained by assuming that stress assignment has shifted from the moraic basis to the syllabic one: in older forms the Q3 syllable was obligatorily a foot, which made the final syllable unstressable, since the second syllable already had a secondary stress ([súu:][lìsi]). In innovative forms, Q3 syllables are treated as any other syllables, i.e. stress placement has become syllabic. This makes the final syllable stressable, as the second syllable does not have one ([súu:li][sèit]). Since this innovation is spreading to another morphological type, there is enough ground to believe that at present, Estonian allows both moraic and syllabic basis for stress assignment.

If this is the case, the critical cases like *jál:kàtelèki*, *jál:katèleki*, and *jál:katelèki* could be accounted for as [jál:][kàte][lèki] (moraic stress assignment), [jál:ka][tèle]ki (syllabic stress) and [jál:ka]te[lèki] (syllabic stress with a ternary stress pattern) respectively.

Assuming that Q3 is heavy and Q2 and Q1 are light makes it possible to do without the use of recursive foot structure, trimoraic syllables or optional disyllabicity of Q3. The reason is that Q3 is not treated as the marked member in the Estonian quantity system, and consequently, does not need to be defined anything else, but a heavy syllable. Following (24), a heavy syllable constitutes a foot on moraic basis, but not on syllabic basis. Since Q2 syllables are treated as light, it explains why they are not able to constitute a monosyllabic foot.

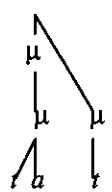
8 Implications for Moraic Theory

In section 3, I outlined some problems that arise from the principle of the Moraic Theory that syllable weight is a linear function of segmental length. To avoid these problems, it is necessary to separate these two phonological domains. A strong support for this comes from the Estonian quantity system: the distinction between Q2 and Q3 is independent of segmental length distinctions.

However, it might be argued that the question is not about length and weight, but weight and syllable prominence. The distinction between weight and prominence is made on the basis of length as opposed to a raw prominence of perceptual salience that is usually due to high tone, low vowels or heaviness of the syllable (see Hayes 1995: 270–76). Thus, it might be argued that the distinction between Q2 and Q3 is also one of prominence rather than quantity: the phonetic cues differentiating between Q2 and Q3 are not solely durational, but include also difference in pitch contour. However, Hayes (1995) argues at length that syllable prominence does not influence foot construction that is based solely on weight properties. Thus, prominence can guide the main stress placement amongst eligible candidates, but not determine the foot size and boundaries. Yet Estonian quantity degrees are namely those that are relevant for constructing the foot structure of words: the possibility of monosyllabic feet is determined by Q3. Provided that Hayes's (1995) distinction between weight and prominence is correct, The Estonian quantity must be treated as a distinction of weight, not that of the prominence.

The other possibility to solve the problem is to use a dual distinction of weight (Hayes 1995). In this case, there are two moraic grids in the prosodic representation. The lower moraic grid represents length contrasts, calculated using the conventions in (8), the higher layers are constructed according to the sonority of the segment involved: the more sonorous a segment is the higher the column of moras it is associated to. Such a double notation allows representing weight contrasts relevant for stress assignment on the higher level and length contrasts on the lower level. For example, the representation of a light CVC syllable, closed by the first half of a geminate would be as follows:

(21)



This representation captures the fact that CVC syllable counts as light in this language (upper grid) while retaining the convention that geminates are underlyingly moraic (lower grid).

The descriptive power of this notation is constrained by the Continuous Column Constraint (Prince 1983: 33, Hayes 1995: 34), which prohibits gaps in the column, i.e the situation where a segment has a mora on the higher level, but not on the lower level. This constraint captures the fact that bimoraic short vowels are not attested typologically: as they would require one mora on the lover level, but two on the higher one, they would violate the constraint. On the other hand, the notation allows for monomoraic long vowels and diphthongs, and there is indeed evidence for the existence of these.

Though the possibility of dual weight can account for the discrepancies between segmental length and syllable weight, it does not separate the notion of length from that of weight. The need for this arises in languages that make the distinction of segmental length, but not the one of weight. Provided that length is expressed by moras, all languages having length must also have weight. This must not be so if the theory strives to descriptive adequacy.

This feature, as well as the ones described above can be accounted for by the two-root theory of length (Selkirk 1990) that encodes length by counting root nodes and weight by counting moras. This means that moras need not to be underlyingly expressed and weight can be defined in the course of derivation just like syllable boundaries.

This is well motivated since underlying representations should not contain information that could be provided by rules (or constraints). Furthermore, this allows for the variation between weight-sensitive and weight-insensitive languages: if a language has a weight distinction, moras are assigned to appropriate segments, if weight distinction is not employed, no moras are assigned. This can be elegantly expressed in terms of Optimality Theory.

The distinction between weight-sensitive and weight-insensitive languages is achieved by ranking faithfulness constraint FAITHμ relative to the family of mora-assigning constraints ASSIGNμ. If FAITHμ >>ASSIGNμ, no moras are allowed (since they are missing underlyingly) and language surfaces as weight-insensitive. If ASSIGNμ >> FAITHμ, the candidates with moras attached will win over moraless candidates. Languages having this ranking are weight-sensitive. However, some languages, such as Estonian, use distinctive syllable weight, and this means that for Estonian moras cannot be assigned in the course of derivation, but need to be specified in underlying representations. These languages would need the ranking FAITHμ >>ASSIGNμ, to prevent changes in moraic tier.

Let us now look more closely how the constraint rankings account for attested variation in syllable weight. For this, it is necessary to specify the more precicely the constraint family ASSIGNμ. I suggest it include 4 constraints, listed in (22):

(22)	$V\mu$	- vowels are moraic
	$*\sigma_{\mu\mu\mu}$	- trimoraic syllables are prohibited
	WxP	- postpeak material must be mora-bearing
	$*C\mu$	- consonants must not be moraic

These four constraints may have the following rankings: $V\mu$ and $*\sigma_{\mu\mu\mu}$ are unordered and universally top ranked within the family ASSIGNμ. The ranking of WxP and $*C\mu$ varies across languages, which gives us two possible rankings:

(23) a) $V\mu$, $*\sigma_{\mu\mu\mu}$ >> WxP >> $*C\mu$
b) $V\mu$, $*\sigma_{\mu\mu\mu}$ >> $*C\mu$ >> WxP

Ranking in (23a) gives us the possible weight combinations for a language where CVC syllable counts as heavy (see the typology in Table 1). First two types, presented in Table 1, are grammatical since they do not violate any ASSIGNμ constraints. The third type ($CV\mu V\mu C$) violates WxP, but it is still grammatical, since it does it to satisfy higher ranking $*\sigma_{\mu\mu\mu}$ and $V\mu$. The fourth type ($CV\mu C\mu$) violates the lowest ranking $*C\mu$ to satisfy WxP, and is therefore grammatical. All other possible combinations are ungrammatical, as they would in some or other way violate against high-ranking constraints. For example, the fifth example violates WxP, and as this violation is not forced by higher-ranking constraints, it is the fatal one. The sixth example violates the lowest ranking $*C\mu$, but since it is not to satisfy any higher-ranking constraints, it is fatal. The rest of the examples in Table 1 are not acceptable, since they violate highest-ranking constraints $V\mu$ and $*\sigma_{\mu\mu\mu}$.

		$V\mu$	$*\sigma_{\mu\mu\mu}$	WxP	$*C\mu$
1	$CV\mu$				
2	$CV\mu V\mu$				
3	$CV\mu V\mu C$			*	
4	$CV\mu C\mu$				*
5	$*CV\mu C$			*!	
6	$*C\mu V\mu$				*!
7	$*CVC\mu$	*!			*
8	$*CV\mu C\mu C\mu$		*!		**

Table 1

If the language uses ranking order (23b) as in Table 2, most of the possible syllable types are the same as presented in Table 1 (types 1–3). The difference concerns the type CVC which surfaces as light, because of the $*C\mu$

which ranks higher than WxP. The last candidate in Tableau 2 is not acceptable, since it has an unforced *Cμ violation.

		Vμ	*$\sigma_{\mu\mu\mu}$	*Cμ	WxP
1	CVμ				
2	CVμVμ				
3	CVμVμC				*
4	CVμC				*
5	*CVμCμ			*!	

Table 2

As seen, the ranking orders in (23) summarise the widely attested cross-linguistic variation of weight in CVC syllables. Surely, the constraints presented above are quite unspecified yet, and work perfectly only in cases where codas have only one consonant. If the coda contains two or more consonants, the constraints as presented in (22), do not specify which one of the consonants receives a mora which one does not. Second, the constraints predict that for CVVC syllables, codas should be weightless. It is not clear yet whether this prediction is empirically correct. Though this syllable type is rare cross-linguistically, Estonian has it and a further study is needed to test this prediction.

9 Conclusion

The evidence, presented in this paper suggests that syllable weight is comparatively autonomous from the segmental (root) tier, and in some languages, it is not even in a direct correlation with segmental length. Weight seems to be autosegmental just like other prosodic features such as, for example, tones. Moraic theory, as presented in Hayes (1988, 1995), however, treats weight as totally dependent on segmental structure, i. e. it has also a duty to represent segmental length. As there are more combinations of segmental length than attested possibilities of syllable weight, it is one of the main reasons why there is yet no satisfactory theory of calculating syllable weight from the segmental structure of syllable. If mora were freed from the duty of expressing segmental length, several desirable consequences would follow:

1. Weight need not be underlying, which means that segmental length would not create unattested weight values.

2. The rule of Weight by Position would become redundant, as different constraint interactions would assign moras according to a general principle.
3. The account of Estonian quantity would become significantly simplified.

Surely, the constraints for weight assignment need to be specified further, and Estonian with its complex length and weight combinations is a good testing ground for elaborating the Moraic Theory.

References

Allen, S. 1973. *Accent and Rhythm.* Cambridge Studies in Linguistics 12. Cambridge: Cambridge University Press.

Ariste, P. 1947. *Foneetilisi probleeme eesti keele alalt.* Tartu: Teaduslik Kirjandus.

Bye, P. 1997. A Generative Perspective on 'Overlength' in Estonian and Saami. In *Estonian Prosody: Papers from a Symposium*, ed. I. Lehiste and J. Ross, 36–68. Tallinn: Eesti Keele Instituut.

Collinder, B. 1929. *Über den finnisch-lappischen Quantitätswechsel.* Uppsala.

Eek, A. 1986. Problems of Estonian Word Prosody. *Estonian Papers of Phonetics* 1984–1985:13–66.

Eek, A. and T. Help. 1986. *Rütminihked eesti keele kujunemisloos.* Tallinn: Eesti NSV Teaduste Akadeemia.

Eek, A. and E. Meister. 1997. Simple Perception Experiments on Estonian Word Prosody: Foot Structure vs. Segmental Quantity. In *Estonian Prosody: Papers from a Symposium*, ed. I. Lehiste and J. Ross, 71–99. Tallinn: Eesti Keele Instituut.

Fox, A. 1988. The Mora, Compensatory Lengthening, and a Two-Tier Theory of Quantity. *Leeds Working Papers in Linguistics and Phonetics* 6: 40–67.

Hayes, B. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20: 253–306.

Hayes, B. 1995. *Metrical stress theory: Principles and case studies.* Chicago: University of Chicago Press.

Hint, M. 1973. *Eesti keele sõnafonoloogia I.* Tallinn: Keele ja Kirjanduse Instituut.

Hint M. 1980. Minevikuline ja tulevikuline aines keelesüsteemis. Prosoodiatüübi nihked ja selle tagajärjed. *Keel ja Kirjandus* 24: 215–223, 270–278, 349–355.

Hint, M. 1997. *Eesti keele astmevahelduse ja prosoodiasüsteemi tüpoloogilised probleemid. Typological problems of Estonian grade alternation and prosodical system. Viron kielen astevaihtelun ja prosodian typologiset ongelmat.* Tallinn-Helsinki: Eesti Keele Sihtasutus.

Hint, M. 1997a. The Estonian Quantity Degrees in Prosody and Morphophonology, In *Estonian Prosody: Papers from a Symposium*, ed. I. Lehiste and J. Ross, 125–135. Tallinn: Eesti Keele Instituut.

Hock, H. H. 1986. Compensatory Lengthening: In Defence of the Concept 'Mora'. *Folia Linguistica* 2:431–460.

Kager, R. 1994. *Ternary Rhythm in Alignment Theory*. Manuscript, University of Utrecht.

Kager, R. 1995. On Affix Allomorphy and Syllable Counting. Paper presented at *Conference on interfaces in phonology*, Berlin, March 23–25, 1995. [Accessible at http://ruccs.rutgers.edu/roa.htm].

Kettunen, L. 1929. *Eestin kielen äännehistoria*. Helsinki.

Laurosela, J. 1922. *Foneettinen tutkimus Etelä-Pohjanmaan murteesta*. Helsinki.

Lehiste, I. 1997. Search for Phonetic Correlates in Estonian Prosody. In *Estonian Prosody: Papers from a Symposium*, ed. I. Lehiste and J. Ross, 11–35. Tallinn: Eesti Keele Instituut.

Maddieson, I. 1993. Splitting the Mora. *UCLA Working Papers in Phonetics* 83: 9–18.

McCarthy, J. and A. S. Prince. 1986. *Prosodic Morphology*. Manuscript, University of Brandeis.

Morén, B. 1998. *The Puzzle of Kashmiri Stress: Implications for Weight Theory*. Ms, Rutgers University. [Accessible at http://ruccs.rutgers.edu/roa.htm].

Odden, D. 1997. Some Theoretical Issues in Estonian Prosody. In *Estonian Prosody: Papers from a Symposium*, ed. I. Lehiste and J. Ross, 165–193. Tallinn: Eesti Keele Instituut.

Pajusalu, K. 1994. Kuidas soomlased õpivad kuulma eesti keele välteid. In *Lähivertailua 7*, ed. Karl Pajusalu and Valma Yli-Vakkuri, 75–88. Turku: Turun Yliopisto.

Posti, L. 1953–54. From Pre-Finnic to Late-Proto-Finnic. Studies on the Development of the Consonantal System. *Finnisch-Ugrische Forschungen* 31: 1–91.

Prince, A. S. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11: 511–562.

Prince, A. S. 1983. Relating to the Grid. *Linguistic Inquiry* 14: 19–100.

Schütz, A. J. 1985. *The Fijian Language*. Honolulu: University of Hawaii Press.

Seiler, H. 1977. *Cahuilla Grammar*. California: Malki Museum Press.

Selkirk, E. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Massachusetts: MIT Press.

Selkirk, E. 1990. A Two Root Theory of Length. *University of Massachusetts Occasional Papers in Linguistics* 14.

Setälä, E. N. 1896. Über Quantitätswechsel im Finnisch-ugrischen. *Journal de la Société Finno-ougrienne* 14.

Tauli, V. 1954. The Origin of the Quantitative System in Estonian. *Journal de la Société Finno-Ougrienne* 57: 1–19.

Tranel, B. 1991. CVC Light Syllables, Geminates and Moraic Theory. *Phonology* 8: 291–302.

Vanoverbergh, M. 1955. *Iloko Grammar*. Detroit: Cellar Book Shop.

Veske, M. 1879. *Eesti keele healte õpetus ja kirjutuse viis*. Tartu.

Woodbury, A. 1987. Meaningful Phonological Processes: A Consideration of Central Alaskan Yupik Eskimo Prosody. *Language* 63: 685–740.