PEAK ALIGNMENT IN FALLING ACCENTS IN ESTONIAN

Mareike Plüschke

Institute of Phonetics and Speech Processing (IPS), LMU Munich, Germany
plueschke@phonetik.uni-muenchen.de

ABSTRACT

This study investigates the extent to which tonal alignment contributes to the differentiation between quantity degrees in Estonian. For this purpose, the temporal displacement of tone targets was measured in read sentences consisting of accented words followed by different numbers of syllables. The accented word was produced with a falling melody and varied over the three possible vowel quantities in Estonian: short, long, and overlong. In agreement with studies for other languages, the accented word's pitch-accent was timed to occur earlier when there were fewer syllables in the tail. In addition, the peak of the overlong quantity degree was proportionally timed earliest, the peak of the short quantity degree latest and the peak of the long quantity degree in between these two. The general conclusion is that quantity and the post-focal context in Estonian influences the way in which the tone targets are segmentally aligned.

Keywords: tonal alignment, Estonian, quantity

1. INTRODUCTION

The way in which the intonation contour and the segmental string are aligned is a much debated issue and their alignment has been shown to be language-dependent [6]. A largely unresolved issue is the extent to which the alignment between the constituents of an intonational melody, and in particular its pitch-accents, are influenced by phonological quantity. The aim of this paper is to address this question through an analysis of Estonian, which is unusual in having a three-way quantity contrast for both vowels and consonants.

The basic quantity distinction in Estonian is between short (Q1), long (Q2) and overlong (Q3) vowels and consonants (see Table 1 for examples of vowel quantity). The most commonly occurring pitch-accent in Estonian is H*+L, then H+L*, with bitonal L*+H pitch-accent having been reported to occur in colloquial speech [2]. Quantity-dependent peak timing differences, to be discussed below, have only been found for H*+L accents ([3-5]).

Previous acoustic and perceptual studies have shown that pitch contributes to quantity distinction in Estonian (e.g. [3] and [8]). It has been observed (e.g. [3]) that the turning point of the pitch accent in Q3 words occurred in the first half of the stressed vowel, while for Q2 and Q1 words it occurred later, near the syllable boundary. A further timing difference reported in [3] consisted in an earlier F0-drop in Q2 words compared to Q1 words. Furthermore, the F0 fall was completed in the initial syllable in Q3-words, but continued in the following unstressed syllables for matched Q2 and Q1 words (c.f. [2, 3]). Finally, pitch differences were shown to play a crucial role in perceptually distinguishing between Q2 and Q3 (cf. [3, 8]).

The reason for the quantity dependent differences in peak alignment is not entirely clear. This study proposes that these pitch differences could be a by-product of the alignment of the peak with the end of the first mora of the stressed vowel. [7] found that there was a correlation between segment duration and moraic representation. If the peak is aligned with the first mora, then the absolute duration between the vowel onset and the peak should be similar in all three quantity degrees. Whether this is indeed the case is not entirely clear from the literature. While [3] claims that there was not much difference in the duration between the vowel onset and the turning point of the pitch accent, a more detailed look at their data (their Fig. 5) actually shows that only for Q1 and Q3 this duration was similar. For Q2 it was longer.

Nevertheless the proportional peak alignment, which would take into account the durational differences between the vowels, should have the order $Q3 < Q2 < Q1$ as found by [3] where $<$ denotes that the peak of Q3 is timed to occur proportionally earlier than the peak of Q2 and so on. In Q1 the peak is near the end of the vowel at the end of the first mora. In Q2 the vowel is longer than in Q1 and the peak occurs proportionally earlier (presumably aligned with the first mora). In Q3 it is still longer and the peak occurs still earlier than in Q2, also presumably aligned with the first mora.

Several cross-linguistic studies have shown that peak alignment is sensitive to the length of the post-focal material. More specifically, it has been observed that the tone target of accented words is aligned later in the word when more post-focal syllables follow it (e.g. for English [15] and German [9] nuclear accents; for English [14] and Greek [1] prenuclear accents). This effect was tested in the present study by analyzing tonal alignment for the three quantities in two contexts that differed depending on the number of post-focal syllables.

2. METHOD

2.1. Subjects and materials

Nine speakers (six female) of Standard Estonian, between 21 and 31 years old, participated in the experiment. The recordings were made with a high quality head set microphone in a professional recording studio in Tartu, Estonia. The subjects were paid for their participation.

The material consisted of four target word triplets (Table 1). The words of each triplet differed only in the vowel quantity of the first syllable.

Table 1: Target words

short (words in genitive and nominative sg.)	long (words in genitive sg.)	overlong (words in partitive sg.)
mõna *('the spell')*	mõõna *('the low tide')*	mõõna *('the low tide')*
lõma *(made-up word)*	lõõma *('the flame')*	lõõma *('the flame')*
mini *('very small')*	miini *('the mine')*	miini *('the mine')*
Mimi *(a name)*	miimi *('the mime')*	miimi *('the mime')*

The target words were embedded in carrier phrases (Table 2) designed to elicit an H*+L accent with the target word in narrow focus position. With the exception of one item, all were existing lexical items in Estonian. The target words were each embedded in two carrier sentences, one containing no unstressed syllables after the target word and one with two unstressed syllables following the target word (these contexts will be referred to as short and long tails respectively). The carrier phrases had to be differently constructed for the overlong vowel contexts because of differences in the target words' grammatical category.

Table 2: Carrier sentences

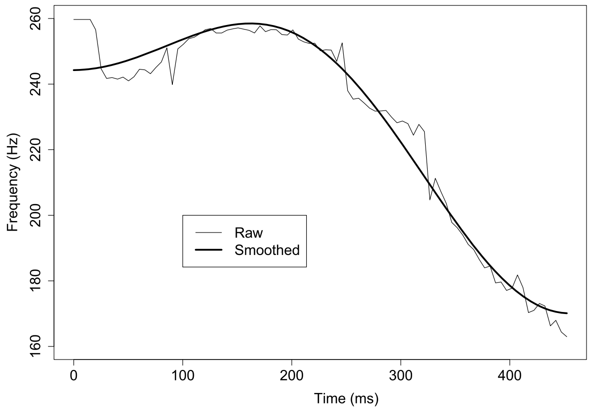
Quantity degree	Carrier sentence with short tail	Carrier sentence with long tail
short/ long	Sa leidsid ___. *(You – found - ___)*	Sa leidsid ___gi ju. *(You – found – ___,too - indeed)*
over-long	Sa nägid ___. *(You – saw – ___)*	Sa nägid ___gi ju. *(You – saw – ___, too - indeed)*

The carrier sentences were presented as answers to questions to elicit narrow focus on the target word. Each carrier sentence was read once. Not every target word was read by every subject. After eliminating recordings containing errors (e.g. the subject produced the vowel with a different quantity than the one indicated), a total of 176 utterances remained for the analysis.

2.2. Analysis

The data were automatically divided into phonetic segments using the *Munich Automatic Segmentation System MAUS* [13] and subsequently manually corrected. An auditory analysis of the data showed that 139 target words were produced with an H*+L pitch accent and 37 with H+L* accents. The results in this paper are based on productions with H*+L. 42 productions with H*+L had to be excluded, because they were not suitable for the automatic procedure used.

Figure 1: The raw and the smoothed contour of the long quantity target word "miimi" spoken by the speaker KK.



The F0 contour of the target word was smoothed using the discrete-cosine-transformation (DCT) with three coefficients following a similar procedure used in [10] (Fig. 1). After the smoothing, the F0 maximum of the whole word was detected automatically and manually checked. The absolute peak alignment, T_{abs} was calculated from (1)

(1) $T_{abs} = T - V_{on}$

where T is the time of the tone target (i.e. the time of occurrence of the f0-peak of H* in the bitonal H*+L accent) and where V_{on} is the acoustic vowel onset.

The proportional peak alignment, T_{prop}, was calculated relative to the acoustic onset and offset of each vowel from (2)

(2) $T_{prop} = T_{abs} / (V_{off} - V_{on})$

where V_{off} is the acoustic vowel offset (see also [3, 14, 15] for a similar procedure).

The statistical analyses reported below were carried out using a mixed model in R. Significance was defined based on an anti-conservative measure of 60 degrees of freedom in the denominator and at an alpha level of 0.01 (see [11], note 1).

3. RESULTS

Three hypotheses were tested: Firstly, that the absolute peak alignment T_{abs} was similar for all three quantities. Secondly, that the proportional peak alignment T_{prop} would be Q3 < Q2 < Q1. Thirdly, that the peak would be left-shifted in the short tail context, compared to the long tail context.

The first hypothesis could not be entirely confirmed. T_{abs} was very similar in Q1 and Q3, but it was a bit later in Q2 (Fig. 3). The second hypothesis was confirmed by the data (Fig. 2). The order of T_{prop} was indeed Q3 < Q2 < Q1. As to the third hypothesis, it was also confirmed: Fig. 2 and Fig. 3 show that the peak was aligned earlier in short tail contexts than in long tail contexts both in the absolute and proportional measures.

A mixed model with the dependent variable T_{prop} and the independent factors Quantity (3 levels: Q1, Q2, Q3) and Tail (2 levels: long, short) and speaker as a random factor statistically confirmed these observations: thus there was a significant effect for both Quantity ($F[2,60] = 110.8$, $p < 0.001$) and Tail ($F[1,60] = 11.7$, $p < 0.01$). A post-hoc analysis showed that the peak of Q1 was significantly later than the peak of Q2 ($t[60] = -3.9$, $p < 0.001$). The peak of Q3 was significantly earlier than the peak of Q2 ($t[60] = -3.0$, $p < 0.001$). The Quantity x Tail interaction was non-significant.

A mixed model with the dependent variable T_{abs} and the same factors as above showed a significant effect for Quantity ($F[2,60] = 13.7$, $p < 0.001$), but not for Tail. The difference between the peaks of Q1 and Q2 and between the peaks of Q2 and Q3 was non-significant. The Quantity x Tail interaction was non-significant.

Figure 2: The proportional time for the tone target, T_{prop}, shown separately for the short (Q1), long (Q2) and overlong Q3) quantities combined with the short (ST) and long tail (LT) contexts.

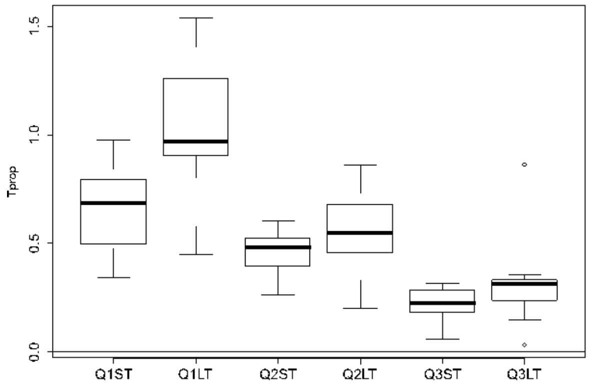
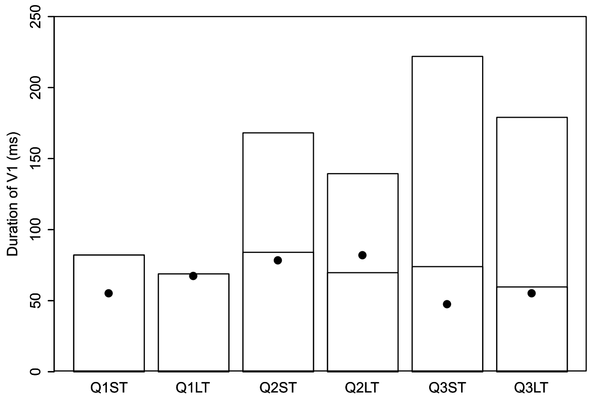


Figure 3: The bars show the duration of V1 separately for the short (Q1), long (Q2) and overlong (Q3) quantities combined with the short (ST) and long tail (LT) contexts. The circles mark the mean value of T_{abs}. In Q2 the lines mark the half of the vowel and in Q3 the lines mark one third after the vowel onset.



4. DISCUSSION

The results of this study have shown that the tonal alignment of a falling nuclear H*+L pitch-accent was affected both by quantity and by the number of syllables following the accented syllable in the tail.

The claim that the peak is aligned with the first mora of the first vowel can be neither confirmed nor rejected. The absolute duration between the vowel onset and the peak T_{abs} was very similar for Q1 and Q3 and longer in Q2, similar to previous findings ([3], their Fig. 5). Nevertheless in this study T_{abs} in Q2 was not as long as in that study [3]. It is not clear whether T_{abs} is the right measure for finding a connection between the peak and the mora, because moras do not need to have the same duration in different conditions, here quantities. It could be therefore more informative to analyze the relationship between the peak and the end of the vowel in Q1, the half of the vowel for Q2 and the third of the vowel in Q3, each standing for the first mora offset (Fig. 3). This will be done in future analyses.

The findings for the second hypothesis are consistent with other studies of Estonian (e.g. [3]) showing that the tone target is aligned proportionally earlier for vowels of greater quantity. This previous research showed that this effect was present in phrase-initial, -internal and -final words. The novel finding of the present study is that this effect seems to be quite robust independently of post-focal context (long vs. short tail contexts). In turn, the context effect found for Estonian is quite similar to that reported for other languages such as German [9], English [14, 15] and Greek [1]. The proportional peak shift due to the varied prosodic context seems to be independent of quantity, because the amount of the proportional peak shift was similar in all quantity degrees.

An unresolved issue remains whether this relationship between early peak timing and long vowel quantity is categorical and phonological as in Dutch [12] or whether it is simply a by-product of vowel duration.

To summarize, this study has shown that the post-focal prosodic context in Estonian influences the peak placement of nuclear accents. Furthermore the absolute peak placement was similar for Q1 and Q3, and slightly different for Q2. The proportional order of the peak placement was Q3 < Q2 < Q1.

Future work will investigate the influence of the post-focal prosodic context on prenuclear H*+L accents as well as nuclear and prenuclear H+L* and L*+H accents. It will also investigate whether other parts of the accent such as the elbow (the point of the fall where the direction is changing) are anchored to a certain position in the segmental string.

5. REFERENCES

[1] Arvaniti, A., Ladd, D.R., Mennen, I. 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.

[2] Asu, E.L. 2004. *The Phonetics and Phonology of Estonian Intonation*. Ph.D. Thesis, University of Camebridge.

[3] Asu, E.L., Lippus, P.; Teras, P., Tuisk, T. 2009**.** The realization of Estonian quantity characteristics in spontaneous speech. *Proceedings of the Xth conference,* Helsinki, 49-56.

[4] Asu, E.L., Nolan, F. 1999. The effect of intonation on pitch cues to the Estonian quantity contrast. *Proc. 14th ICPhS San Francisco 3*, 1873-1876.

[5] Asu, E.L., Nolan, F. 2007. The analysis of low accentuation in Estonian. *Language and Speech* 50(4), 567-588.

[6] Atterer, M., Ladd, D.R. 2004. On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics* 32, 177-197.

[7] Broselow, E., Chen, S.-I., Huffman, M. 1997: Syllable weight: convergence of phonology and phonetics. *Phonology* 14, 47-82.

[8] Lippus, P., Pajusalu, K., Allik, J. 2011. The role of pitch cue in the perception of the Estonian long quantity. In: Frota, S., Elordieta, G., Prieto, P. (eds.), *Prosodic Categories: Production, Perception and Comprehension*, 231-242.

[9] Peters, J. 1999. The timing of nuclear high accents in German dialects. *Proc. 14th ICPhS* San Francisco 3, 1877-1880.

[10] Rathcke, T., Harrington, J. 2006. Is there a distinction between H+!H* and H+L* in Standard German? Evidence from an acoustic and auditory analysis. *SP-2006*, paper 151.

[11] Reubold, U., Harrington, J., Kleber, F. 2010. Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52, 638-651.

[12] Schepman, A., Lickley, R., Ladd, D.R. 2006. Effects of vowel length and "right context" on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34, 1-28.

[13] Schiel, F. 1999. Automatic phonetic transcription of non-prompted speech. *Proc. 14. ICPhS* San Francisco, 607-610.

[14] Silverman, K., Pierrehumbert, J. 1990. The Timing of Prenuclear High Accents in English. In *Papers in Laboratory Phonology I.* Cambridge UK, 72-106.

[15] Steele, S.A. 1986. Nuclear accent F0 peak location: Effects of rate, vowel, and number of following syllables. *J. Acoust. Soc. Am.* 80, S51.