ACOUSTIC CORRELATES OF EMPHASIS IN ESTONIAN

Meelis Mihkla, Heete Sahkai, Mari-Liis Kalvik

Institute of the Estonian Language
meelis@eki.ee, heete@eki.ee, mariliis@eki.ee

ABSTRACT

The paper tests for several potential acoustic correlates of the emphatic accent differentiating between broad focus and clause-final narrow focus in Estonian. The examined correlates were related to F0, duration, intensity and spectral emphasis. Interestingly, duration was found to be by far the strongest correlate of emphasis, whereas cross-linguistically emphasis has been found to be expressed primarily with F0. Since duration relationships in the disyllabic foot play an important role in Estonian word prosody, being the primary correlate of the Estonian three-way quantity system, the study additionally examined the interaction of emphatic lengthening with word quantity. It was found that emphatic lengthening does not affect the quantity features, but depends on the quantity structure of the word, affecting primarily the segment that bears the quantity.[1]

Keywords: focus, emphasis, duration, Estonian quantity.

1. INTRODUCTION

In many languages emphasis plays a meaningful role, differentiating between Information Structural categories (e.g. [8]:257). In Estonian, it has been found to differentiate between broad information focus and narrow information/contrastive focus (narrow contrastive focus was not found to be distinguishable from narrow information focus), cf. [12, 13, 14]. When the narrowly focused word is not clause-final, it is signalled by several features, including phonological ones like the placement of the primary accent, a consistent pitch accent type and possibly also separate phrasing. However, when the narrowly focused word is clause-final, the only differentiating feature is accent strength. In particular, a sentence with sentence-final narrow focus was found to differ from a segmentally identical sentence produced with broad focus in terms of a lengthening of the narrowly focused word and a shortening of the preceding accented word(s). Interestingly however, narrow focus was not significantly differentiated by a larger F0 range (cf. [15] for a similar result), as seems to be generally the case cross-linguistically (it has even been

suggested that contrastive focus, as opposed to broad focus, is only a matter of F0, see [6]). The first goal of the present study is therefore to verify the finding that clause-final narrow focus is not signalled by an increased F0 range, and to test for two other possible F0-related correlates of emphasis: peak alignment, and difference between the F0 maximum of the emphatic accent and that of the preceding accent.

The second goal of the study is to verify the previous finding that lengthening is the primary correlate of emphasis in Estonian, and to examine how emphatic lengthening interacts with the duration ratio of the foot. The duration ratio of the stressed and unstressed syllable of the disyllabic foot is an important feature of Estonian word prosody, as it is the primary correlate of the Estonian three-way quantity system (with peak alignment as the secondary feature; see e.g. [2, 4, 9, 10]). The duration ratio of the short quantity falls in the interval 0.4–1.0, for the long quantity the ratios are typically larger than 1.0, and for the overlong quantity larger than 2.0. Several previous studies have examined the interaction of the duration ratios characterising the quantity degrees with pre-boundary lengthening. [7] found that the characteristic duration ratios of the three quantity degrees are preserved under pre-boundary lengthening. [11] further found that pre-boundary lengthening affects primarily the main bearer of the quantity, i.e. the vowel of the stressed syllable in vowel quantity words and the consonant on the syllable boundary in consonant quantity words. We hypothesise that emphatic lengthening, like pre-boundary lengthening, does not affect the quantity features, and that it is realised differently in the quantity degrees as well as in vowel *vs.* consonant quantity words.

The final goal of the study is to examine two further possible correlates of emphasis: intensity and spectral emphasis.

2. DATA AND METHODS

The data consist of sentences read in a dialogue setting by four subjects, three women and one man. Each sentence was read twice, once in response to a question eliciting broad focus and once in response to a question eliciting narrow focus on the last word. The questions were read by the experimenter, sitting

opposite the subject. The sentence-final test words were six minimal quantity triplets representing both vowel-quantity words (CV[V]CV, e.g. /kalu/ 'fish-part.pl', /kaalu/ 'scales-gen.sg', /kaːlu/ 'scales-part.sg') and consonant-quantity words (CVC[C]V, e.g. /linu/ 'sheet-part.pl', /linnu/ 'bird-gen.sg', /linːnu/ 'city-part.pl'). Altogether, the data contained 144 sentences. For the purposes of normalisation and reference, the material also contained the same sentences read with a narrow focus on the first word.

The material was segmented and annotated with Praat [3] and the values of a set of acoustic parameters were measured, using a Praat script. These values were then used to generate the values of a set of derived parameters, e.g. the duration changes of the target words and their segments. In order to be able to compare the duration changes of the different target words in the two focus conditions and to control for potential variations in speech rate, the durations of the different versions of the sentences were first normalised and then the relative duration changes of the words in the different focus conditions were calculated.

Altogether eight features were chosen as candidates of the acoustic correlates of emphasis: (1) relative lengthening of the target word; (2) F0 range of the target word; (3) relative position of the turning point of the F0 curve in the rhyme of the stressed syllable of the target word; (4) difference between the F0 maximum of the target word and that of the preceding accented word (always the first word of the sentence); (5) difference between the mean intensity of the stressed and unstressed syllable of the test word; (6) difference between the mean intensity of the test word and that of the preceding accented word in the sentence; (7) difference between the mean spectral emphasis of the stressed and unstressed syllable of the test word; (8) difference between the mean spectral emphasis of the test word and that of the preceding accented word in the sentence. In addition, the duration ratio of the rhyme of the stressed and unstressed syllable of the target word was examined, in order to test for a potential interaction between the duration-related features of emphasis and word quantity. Similarly, the position of the F0 peak, chosen as one of the candidate features of emphasis, is also a secondary feature of word quantity; consequently, this parameter too will permit to examine the interaction of emphasis and quantity.

Five of the examined parameters relate to the target words and four are sentence-level parameters. In other words, we aim to establish to what extent emphasis is signalled in the acoustic features of the word itself and to what extent in the global features of the sentence. Previous results relating to duration

changes showed that broad focus and clause-final narrow focus differed not only in the lengthening of the narrowly focused word but also in the shortening of the preceding accented word [12].

We examined first each parameter separately and then used binary logistic regression to determine the significance of all the parameters in combination and their potential interactions.

3. RESULTS AND DISCUSSION

Fig. 1 describes the distributions of four of the features that turned out to be significant with respect to the focus factor. Boxplot A shows that narrowly focused clause-final words are in average 15.5 % longer than clause-final words in broad-focus sentences. The lengthening of the target word is also a strongly significant feature of emphasis: [$F_{(1, 142)}$ = 162.64, $p < 0.0001$]. This confirms the earlier finding that lengthening is an important correlate of emphasis.

Figure 1: Boxplots of relative lengthenings and F0 ranges (in st) in target words, F0 peak differences (in st) in sentences, and mean intensity ranges (in dB) in target words.
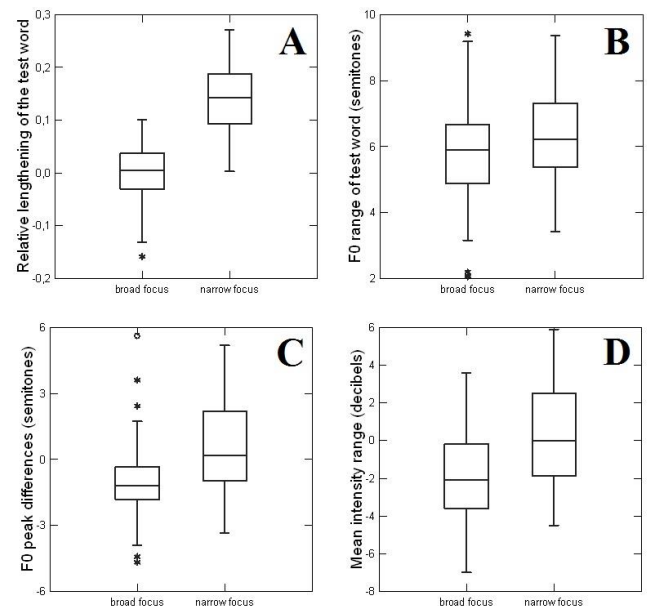
There is a weak interaction between lengthening and word quantity [$F_{(4, 139)}$ = 3.11, $p = 0.0637$], suggesting that the duration changes of the segments of the target words may differ depending on quantity. Table 1 shows the p-values indicating the significance of the lengthening of each segment in the words of different quantity structure and degree in the two focus conditions. Narrowly focused vowel quantity words display a more significant lengthening in the vowel of the stressed syllable

Table 1: The influence of emphasis (final narrow focus *vs.* broad focus) on the segment lengthening of target words in different quantity degrees (significant p-values are in bold).

Segments	Target words with vowel quantity			Target words with consonant quantity		
	Q1 (short) $C_1V_2C_2V_2$	**Q2** (long) $C_1V_1V_2C_2V$	**Q3** (overlong) $C_1V_1V_1{:}C_2V_2$	**Q1** (short) $C_1V_1C_2V_2$	**Q2** (long) $C_1V_1C_2C_2V_2$	**Q3** (overlong) $C_1V_1C_2{:}C_2V_2$
C_1	**0.0272***	0.2716	0.2769	**<0.0001*****	**0.0061***	**0.0013****
V_1	**0.0007****	**<0.0001*****	**<0.0001*****	**0.0006****	**0.0362***	0.5366
C_2	0.0896	0.6768	0.0953	**0.0198***	**<0.0001*****	**<0.0001*****
V_2	0.0738	**0.0018****	0.9869	**0.0376***	0.4929	0.6631
word	**0.0006****	**<0.0001*****	**<0.0001*****	**<0.0001*****	**0.0006****	**<0.0001*****

*p<0.05; **p<0.005; ***p<0.0005

(V1), and consonant quantity words in the consonant on the syllable boundary (C2). In other words, emphatic lengthening is more significant in the segment that bears the quantity, as was found for pre-boundary lengthening.

In the words with overlong vowel quantity, only the lengthening of V1 is significant. In consonant quantity words, emphatic lengthening also affects the onset of the first syllable (C1), which does not play a role in the signalling of the quantity degree of the word. However, more data is needed in order to draw conclusions on other word quantity-dependant characteristics of emphatic lengthening besides the fact that it clearly affects most the segment bearing the quantity.

We hypothesised that, like pre-boundary lengthening, emphatic lengthening will not affect the characteristic duration ratios of the quantity degrees. Our measurements confirm indeed that quantity features remain stable under emphasis. Fig. 2 shows, by quantity and focus type, the two main parameters of the Estonian quantity distinction: the duration ratio of the rhymes of the stressed and unstressed syllable, and the position of the F0 peak in the rhyme of the stressed syllable. As can be seen, the values of these parameters vary only according to quantity degree and not according to focus condition. This is confirmed by the significance statistics, which are [F(1, 142) = 1.69, p = 0.1935] for the duration ratio and [F(1, 142) = 0.96, p = 0.3292] for peak alignment. This means that the quantity features are not significant in the focus conditions.

F0 range of the target words, and the difference between the F0 maximum of the emphatic accent and that of the preceding accent correlated more weakly with emphasis (see Fig. 1 boxplots B and C), the significance statistics being respectively [F(1, 142) = 7.75, p = 0.0162] and [F(1, 142) = 33.15, p =

0.0003]. In other words, the narrow-focus accent displayed a slightly larger F0 range and its declination with respect to the previous accent was somewhat smaller than was the case for the last accent in the corresponding broad-focus sentence.

Figure 2: Boxplots of target word quantity degree features.
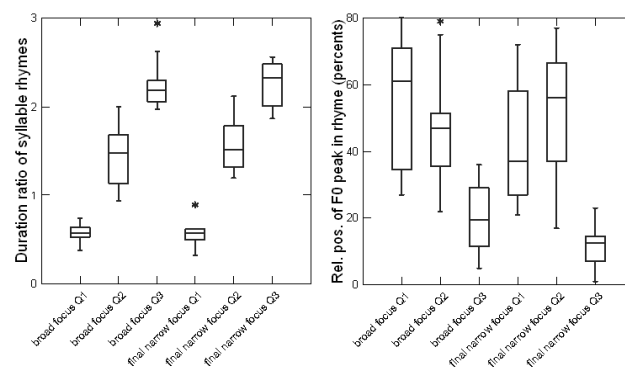
Consequently, the three F0-related parameters that were tested, i.e. peak alignment, F0 range of the target word, and the difference between F0 maxima in the sentence, did not differentiate significantly (in the case of peak alignment) or differentiated weakly between the focus conditions, confirming the earlier finding that F0 is not the primary correlate of emphasis in Estonian.

A feature that correlated relatively strongly with emphasis was the difference between the mean intensity of the target word and that of the previous accented word: [F(1, 142) = 40.72, p < 0.0002]. Both in broad-focus and narrow-focus sentences the mean intensity of the target word is inferior to that of the preceding accented word, but in narrow-focus sentences this difference is smaller. A weakly significant feature was target-word internal intensity difference ([F(1, 142) = 5.42, p = 0.0215]), cf. Fig. 1, boxplot D: under emphasis, the intensity

difference between the stressed and the unstressed syllable is larger.

Table 2: Logistic regression analysis of accent strength (significant values in bold).

Parameter	Est.	S.E.	Z	p-Value
Constant	-4.600	2.311	-1.991	**0.047**
Dur_rel_change	42.573	8.859	4.806	**<0.0005**
Dur_ratio_rhymes	0.605	0.551	1.098	0.272
F0_range	0.426	0.265	1.610	0.107
F0_peak_alignm	0.020	0.019	1.081	0.279
F0_peak_dif	-0.126	0.328	-0.384	0.701
INT_local	-0.400	0.181	-2.210	**0.027**
INT_global	0.449	0.193	2.328	**0.020**
EMPH_local	0.043	0.065	0.663	0.507
EMPH_global	-0.071	0.062	-1.154	0.248
Overall model fit				
Chi-Square		133.170		
p-Value		<0.0005		
Naglekerke's R-Square		0.805		

The word-internal and global spectral emphasis parameters were not significant with respect to emphasis (respectively [$F(1, 142) = 0.58$, $p = 0.4481$] and [$F(1, 142) = 0.31$, $p = 0.5841$]).

A binary logistic regression model was used to determine the significance and the interaction of the combination of the various parameters in describing the two focus conditions. The argument features used were the nine acoustic parameters described above (the eight candidate correlates and the duration ratio signalling the quantity degree). Table 2 presents the results of the regression analysis. No new correlations were revealed by potential hidden interactions of the parameters. In principle, the same primary features of emphasis emerged. Again, by far the most significant feature is the relative lengthening of the emphatic word. The intensity range of the sentence and the target-word internal intensity difference appear as weaker significant components of the regression equation. Somewhat surprisingly, none of the F0-related parameters turned out to be significant, including the F0 range of the target word and the difference of the F0

maxima in the sentence, which were significant separately.

The model is statistically significant ($p < 0.0005$) and describes more than 80% of the variation in the data.

4. CONCLUSIONS

The study confirmed that the strongest correlate of the emphasis signalling sentence-final narrow focus in Estonian is the lengthening of the focused word.

To a lesser degree, narrow focus is signalled by a higher intensity of the focal word in relation to the preceding accented word, and by a larger intensity difference between the stressed and the unstressed syllable of the focal word.

Interestingly, none of the three F0-related parameters that were tested for correlated clearly with emphasis. One of these parameters, alignment of the F0 peak with the stressed syllable, is also a secondary feature signalling the quantity degree; this parameter clearly correlated only with quantity. Consequently, while F0 seems to be cross-linguistically the strongest correlate of emphasis, this is not the case in Estonian.

Two other features that did not turn out to be significant were related to spectral emphasis.

The results further suggest that emphasis is signalled not only in the emphatic word, but also by the global features of the sentence: global intensity range, a slight deviation in the declination line of the F0 peaks, and a shortening of the previous accented word, shown by an earlier study [12].

One conclusion of the study is that duration plays a major role in Estonian prosody: in addition to being the primary correlate of quantity, which is a central feature of Estonian word prosody, and of prosodic boundaries, it also turns out to be the main correlate of emphasis. Given these multiple functions of duration, the study additionally examined the interaction of emphatic lengthening with the duration-related primary quantity feature, namely, the duration ratio of the rhyme of the stressed and unstressed syllable of the disyllabic foot. It was found that, like pre-boundary lengthening, emphatic lengthening does not affect the duration ratio of the foot, and is primarily realised in the segment bearing the quantity, i.e. the vowel of the stressed syllable in vowel quantity words and the consonant on the syllable boundary in consonant quantity words.

5. REFERENCES

[1] Asu, E. L. 2004. *The phonetics and phonology of Estonian intonation.* PhD dissertation, University of Cambridge.

[2] Asu, E. L., Lippus, P., Teras, P., Tuisk, T. 2009. The realization of Estonian quantity characteristics in spontaneous speech. In: Vainio, M., Aulanko, R., Aaltonen, O. (eds), *Proc. of Int. Conference: Nordic Prosody X* Helsinki. Frankfurt, Berlin, New York: Peter Lang Verlag, 49–56.

[3] Boersma, P., Weenink, D. 2012. Praat: Doing phonetics by computer. Version 5.3.14 [Computer program]. Retrieved on 13.04.2012 from <http://www.praat.org/>.

[4] Eek, A. 1983. Kvantiteet ja rõhk eesti keeles. *Keel ja Kirjandus* 26, 9, 481–489; 10, 549–559.

[5] Eek, A., Meister, E. 2003. Foneetilisi katseid ja arutlusi kvantiteedi alalt: Häälikukestusi muutvad kontekstid ja välde. *Keel ja Kirjandus* 46, 11, 815–837; 12, 904–918.

[6] Görs, K., Niebuhr, O. 2012. Hocus Focus - What the eliciation method tells us about types and exponents of contrastive focus. *Proc. of the 6th Int. Conference of Speech Prosody* Shanghai, 262–265.

[7] Krull, D. 1997. Prepausal lengthening in Estonian: Evidence from Conversational speech. In: Lehiste, I., Ross, J. (eds), *Estonian Prosody: Papers from a Symposium. Proc. of the Int. Symposium on Estonian Prosody* Tallinn. Tallinn: Institute of the Estonian Language,136–148.

[8] Ladd, D. R. 2008. *Intonational phonology.* Cambridge: Cambridge University Press.

[9] Lehiste, I. 1960. Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics I.* Bloomington: Indiana University, 21–82.

[10] Lehiste, I. 1970. *Suprasegmentals.* Cambridge (Mass.), London: The M.I.T. Press.

[11] Plüschke, M., Harrington, J. 2013. The domain of phrase-final lengthening in Estonian. In: Asu, E. L., Lippus, P. (eds), *Proc. of Int. Conference: Nordic Prosody XI* Tartu. Frankfurt am Main: Peter Lang Verlag, 293–302.

[12] Sahkai, H., Kalvik, M.-L., Mihkla, M. 2013a. Prosodic effects of Information Structure in Estonian. In: Asu, E. L., Lippus, P. (eds) *Proc. of Int. Conference: Nordic Prosody XI* Tartu. Frankfurt am Main: Peter Lang Verlag, 323–332.

[13] Sahkai, H., Kalvik, M.-L., Mihkla, M. 2013b. Prosody of contrastive focus in Estonian. *Proc. of INTERSPEECH-2013* Lyon, 315–319.

[14] Sahkai, H., Kalvik, M.-L., Mihkla, M. 2014. Perception of focus size and focus type in Estonian. In: Jähi, K., Taimi, L. (eds) *Proc. of the XXVIII Fonetiikan päivät* Turku, 72–79. <http://urn.fi/URN:ISBN:978-951-29-5980-8>

[15] Salveste, N., Reichel, U. 2015. *Investigation of Estonian narrow focus prosody with automatic elbow detection.* Presentation at the Workshop on Finno-Ugric Prosody, 9–10 April 2015, University of Tartu.