INTENSITY AND SPECTRAL PARAMETERS AS CORRELATES OF PHRASAL STRESS AND WORD QUANTITY IN ESTONIAN

Heete Sahkai and Meelis Mihkla

Institute of the Estonian Language
Heete.Sahkai@eki.ee, Meelis.Mihkla@eki.ee

ABSTRACT

The study examines overall intensity, spectral emphasis and spectral balance measures in test words that contrast minimally with respect to the Estonian three-level word quantity distinction and represent three accentuation conditions: no intonational pitch accent, nuclear pitch accent signalling broad focus, and emphatic nuclear pitch accent signalling narrow/contrastive focus. The study asks 1) whether the presence of an intonational pitch accent in a word also correlates with intensity and spectral parameters as cues to metrical stress, and 2) whether quantity as a word-level prosodic property affects the realisation of stress at phrase level. The study finds that overall intensity correlates both with the accentuation and quantity conditions, whereas among the spectral measures, only spectral emphasis weakly correlates with quantity. As for interactions, the intensity differences found between the quantity degrees in the deaccented condition are partly neutralised in the accented conditions, but the effect of accentuation/phrase stress is identical in all three quantity degrees.

Keywords: phrasal stress, Estonian quantity, overall intensity, spectral emphasis, spectral balance.

1. INTRODUCTION

The study[1] examines overall intensity, spectral emphasis and spectral balance measures in test words that contrast minimally with respect to the Estonian three-level word quantity distinction and represent three accentuation conditions: no intonational pitch accent, nuclear pitch accent signalling broad focus, and emphatic nuclear pitch accent signalling narrow/contrastive focus.

The first aim of the study is to examine whether the presence of an intonational pitch accent in a word also correlates with intensity and spectral parameters, evidencing that in addition to the intonation pattern, phrasal units also have a metrical stress pattern signalled by dynamic cues (cf. [4] p. 54). Previous studies on the acoustic correlates of phrase level prominence in Estonian have shown that words with both non-emphatic and emphatic nuclear accents are distinguished by increased F0 values as well as by an increase in duration, i.e. both by tonal and dynamic correlates ([8], [13]). The present study examines whether they also correlate with overall intensity and spectral measures as further cues to phrasal stress. Accented words have previously been found to correlate with overall intensity and spectral measures for instance in Swedish: [2] found that increases both in overall intensity and, especially, spectral emphasis were reliable acoustic correlates of focal accents in Swedish. More studies have examined and compared these parameters as correlates of stressed *vs.* unstressed syllables in words, but with conflicting results. For overall intensity, several studies have found that it varies together with F0 rather than word stress (e.g. [10], [14]); on the other hand, [6 p. 144] concluded that despite various dependence relationships, F0 and intensity can be considered at least to some extent independent. Spectral measures in turn have been found to correlate with word stress in some studies (e.g. [14]), but with vowel quality rather than word stress in others (e.g. [11]). Intensity and spectral measures have also been studied as correlates of word stress in Estonian by [1], but without controlling for accentuation; they found that only spectral balance distinguished the primary stressed syllable, while mean overall intensity and spectral emphasis distinguished the entire primary stressed foot from all the other syllables.

The second question addressed in the study is whether quantity as a word-level prosodic property affects the realisation of stress at phrase level. The characteristic property of Estonian word prosody, the three-way quantity distinction (for an overview see e.g. [7]), is primarily manifested in the duration ratio of the stressed and unstressed syllable of the disyllabic foot. Word quantity has been found to affect the realisation of several phrase-level duration-related phenomena: pre-boundary lengthening [e.g. [3], [12]), phrasal stress [5], emphatic lengthening [13]. Word quantity also determines the alignment of the intonational H*+L pitch accent.

[1] The study was supported by the Centre of Excellence in Estonian Studies (European Regional Development Fund) and the project IUT35-1 (Estonian Research Council).

2. DATA AND METHOD

The data consists of 18 disyllabic words that constitute three minimal vowel-based quantity triplets and three minimal consonant-based quantity triplets, for instance *lina* 'sheet' (Q1), Liina [proper name] (Q2), *vii:na* 'vodka.PAR' (Q3); *linu* 'sheet.PAR.PL' (Q1), *linnu* 'bird.GEN' (Q2), *lin:nu* 'town.PAR.PL' (Q3). The test words were embedded in four-word sentences where they occurred in the third position. The sentences were read aloud by nine speakers (five women and four men) in answer to three questions eliciting three information structures and accentuation conditions: 1) broad focus (the test word received the nuclear accent signalling broad focus), 2) narrow focus on the test word (the test word received an emphatic nuclear accent signalling narrow focus), and 3) narrow focus on the first word of the sentence (the test word was deaccented). The condenser microphone Neumann TLM 102 was used to record the speech material. The data set consisted of a total of 486 sentences.

The overall intensity measures that were examined were the mean intensity of the stressed syllable vowel, and the intensity range of the test word, calculated as the difference between the mean intensity of the stressed syllable vowel and the mean intensity of the unstressed syllable vowel.

Since overall intensity may depend on F0, we compared the intensity measures with the relative F0 maximum of the stressed syllable vowel and the F0 range of the test words. To calculate the relative F0 maximum, we first calculated for each speaker the average F0 maximum of the stressed syllable vowel of each test word across the three accentuation conditions, and then calculated the relative F0 maximum for each test word in each condition as the difference from the average value. The F0 range was calculated as the difference between the F0 maximum of the stressed syllable vowel and the F0 minimum of the unstressed syllable vowel.

Spectral emphasis was calculated in the vowel of the stressed syllable as the difference between the energy of the entire spectrum (i.e. up to the Nyquist frequency 0–24 kHz) and the energy of the frequency band 0–1.43*F0, following [15].

Spectral balance was characterised in two ways. First, spectral tilt was measured as the ratio of the spectral energy in the frequency bands 0–1 kHz and 1–5 kHz, following [9]. Secondly, following [14], spectral energy was measured as the change of intensity in four successive frequency bands: 0–0.5 (B1), 0.5–1 (B2), 1–2 (B3), and 2–4 kHz (B4).

We used classical discriminant analysis and multinomial logistic regression to evaluate the classification power and the contribution of the intensity and spectral measures to the prediction of the accentuation and quantity conditions.

3. RESULTS

Spectral balance and spectral emphasis. The intensity distribution over four frequency bands did not differ significantly in the three accentuation conditions. As shown in Figure 1 for the vowels /ɑ/ and /e/, the intensity level in all frequency bands tends to be highest in the emphatic accent condition and lowest in the deaccented condition, but these differences are not significant. The ratio of the spectral energy in the frequency bands 0–1 kHz and 1–5 kHz also did not show a significant difference between the accentuation conditions. The two measures also did not distinguish between the quantity degrees.

Figure 1: The mean intensity in different frequency bands of the stressed syllable vowels (/ɑ/ and /e/) in the three accentuation conditions.

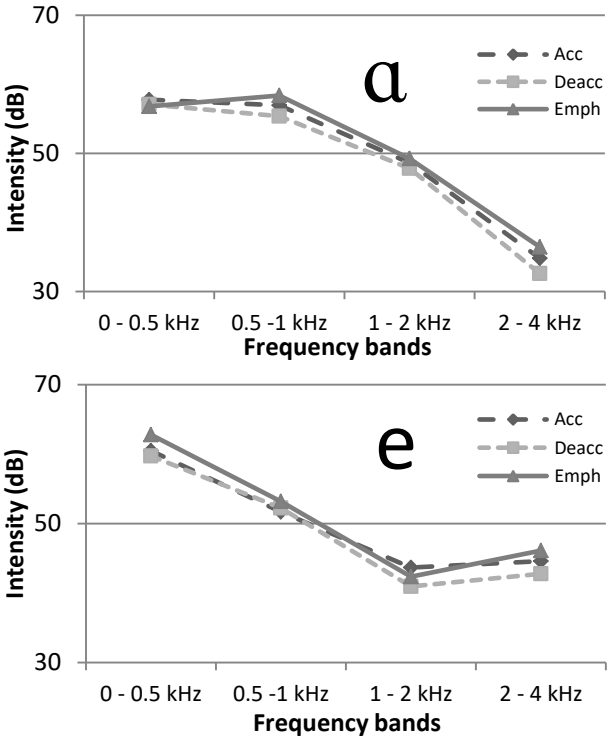
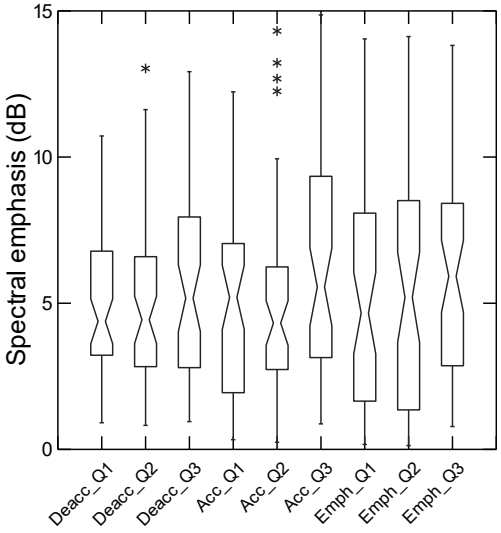


Figure 2 shows the distribution of the spectral emphasis values as a function of the accentuation and quantity conditions. The accentuation conditions did not have a significant effect on spectral emphasis ($F[2, 483]=1.65$; $p=0.192$), but the effect of quantity was moderately significant ($F[2,483]=3.29$; $p=0.038$). The clearest difference was obtained between the overlong quantity on the one hand and the short and the long quantity on the other hand,

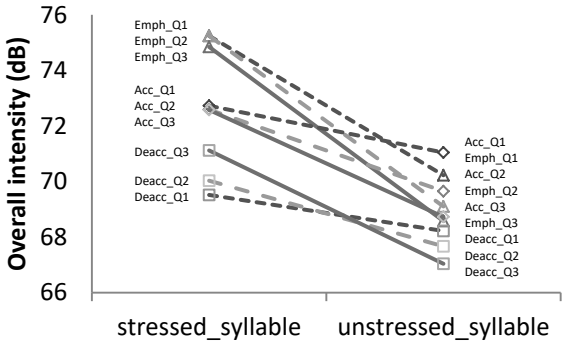
respectively (F[1,322]=4.75; p=0.029) and (F[1,322]=4.81; p=0.028).

Figure 2: Spectral emphasis of the stressed syllable vowel of the test words in the three accentuation and three quantity conditions.



Overall intensity. The results of the measurements of the overall intensity as a function of the accentuation and quantity conditions are summarised in Figure 3.

Figure 3: The average mean intensity of the stressed syllable and the intensity range of the word in the three accentuation and three quantity conditions (the unstressed syllable values were obtained by subtracting the intensity range values from the mean intensity values of the stressed syllable).



In the **deaccented condition**, both the mean intensity of the stressed syllable vowel and the intensity range of the test word distinguish between the three quantity degrees. The **non-emphatic accent condition** is distinguished from the deaccented condition by an increase of the intensity level in the stressed syllable, whereas the intensity range still distinguishes between the three quantity degrees in the same way as in the deaccented condition. The **emphatic accent condition** in turn is distinguished from the non-emphatic accent condition by increased intensity in the stressed syllable vowel, as well as by an increased intensity range of the word. Additionally, the intensity range also distinguishes between the three quantity degrees, as in the deaccented and non-emphatic accent conditions.

Since differences in overall intensity may result from differences in F0, we compared the intensity measures with F0 measures. Figure 4 shows that like the mean intensity of the stressed syllable vowel (F[2,483]=88.4; p<.001), the relative F0 maximum of the stressed syllable vowel distinguished between the three accentuation conditions (F[2,483]=281.7; p<.001). However, unlike the intensity level, it did not distinguish between the quantity degrees in the deaccented condition (F[2,483]=1.87; p=0.158).

Figure 4: The relative F0 maximum of the stressed syllable vowel in the three accentuation and three quantity conditions.

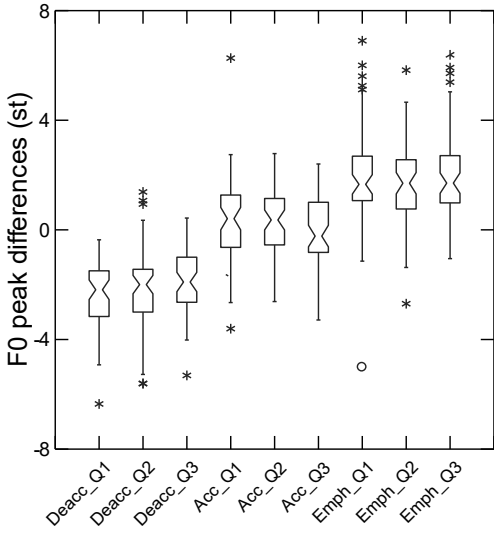
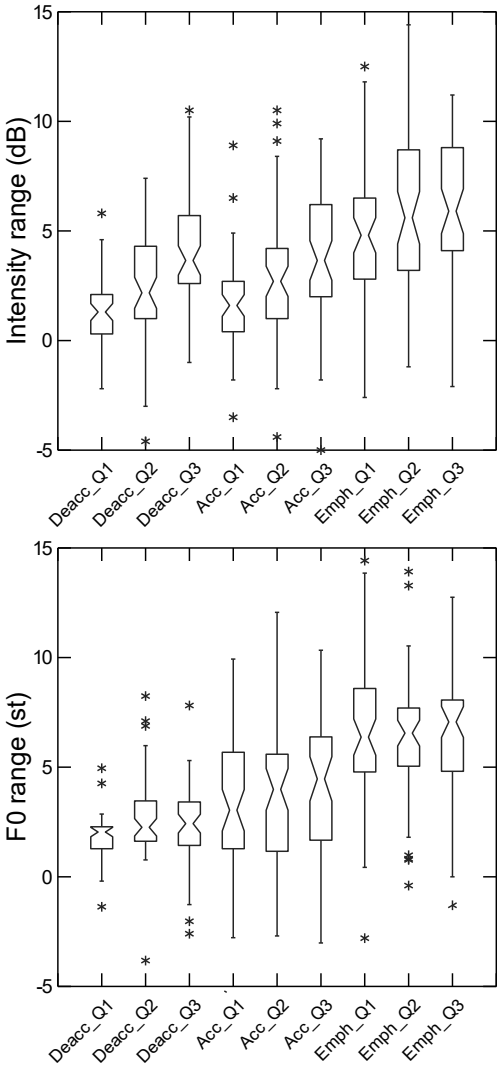


Figure 5 shows that the F0 range of the test words distinguished between all three accentuation conditions (F[2,483]=100.5; p<.001), unlike the intensity range which only distinguished the emphatic accent condition. Also unlike the intensity range, which consistently distinguished between all three quantity degrees in all the accentuation conditions (F[2,483]=15.5; p<.001), the F0 range in general did not differentiate the quantity degrees (F[2,483]=0.51; p=0.607), although it weakly distinguished between them in the non-emphatic accent condition (F[2,483]=3.41; p=0.034).

Figure 5: The intensity range (upper figure) and the F0 range (lower figure) of the test words in the three accentuation and three quantity conditions.



Modelling. Table 1 presents the significant features in the binary predictive models for accentuation and quantity obtained by multinomial linear regression analysis. The reference categories were respectively the non-emphatic accent condition (Acc) and the long quantity degree (Q2). The overall intensity level of the stressed syllable vowel dominates as the significant feature in the accentuation prediction models, while the intensity range dominates in the quantity prediction models. The signficance of the other parameters is arbitrary.

A classical discriminant analysis showed that the accentuation models M1 and M2 have a moderate classification power (respectively 67% and 70%) and the quantity models M3 and M4 only a marginal classification power (61% for both models) on the basis of the intensity and spectral parameters.

Table 1: Multinomial logistic regression models predicting the accentuation and quantity categories (+ significant feature, – insignificant feature).

	Int	Int Range	Spec Emph	Spec Tilt	Int B1	Int B2	Int B3	Int B4
M1: Acc vs Deacc	+	–	–	+	+	–	+	+
M2: Acc vs Emph	+	+	–	–	–	–	–	–
M3: Q2 vs Q1	–	+	–	–	+	–	–	–
M4: Q2 vs Q3	–	+	+	–	–	–	–	–

4. DISCUSSION AND CONCLUSIONS

The study asked whether the presence of an intonational pitch accent in a word also correlates with overall intensity, spectral emphasis and spectral balance as cues to phrasal stress, and whether this correlation is affected by the quantity degree of the word. The study found that the accentuation conditions only correlated with overall intensity: both non-emphatic and emphatic accentuation was accompanied by an increased intensity in the stressed syllable, and emphasis also correlated with an increased intensity range in the word; these effects were identical in all three quantity degrees. However, these intensity differences can be seen to reflect F0 differences, which also distinguished between the accentuation conditions: both the non-emphatic and the emphatic accent involved an increase in the relative F0 maximum of the stressed syllable, as well as in the F0 range. It can thus be concluded that phrasal stress in Estonian does not correlate directly with overall intensity, spectral emphasis or spectral balance.

Interestingly, the study found that overall intensity and spectral emphasis correlated with word quantity. The intensity range distinguished the quantity degrees in all the accentuation conditions, and the mean intensity of the stressed syllable vowel in the deaccented condition. A comparison of the intensity and F0 measures suggested that the distinctive intensity features of the three quantity degrees did not mirror F0 differences, as was the case for the accentuation conditions. In addition to overall intensity, spectral emphasis weakly distinguished the overlong quantity. In conclusion, although the intensity and spectral features played a relatively small role in the quantity prediction models, it is remarkable that the intensity range showed a consistent discrimination power in all the accentuation conditions. Intensity range is thus a potentially relevant correlate of word quantity in Estonian and requires further study.

5. REFERENCES

[1] Asu, E. L., Lippus, P. 2018. Acoustic correlates of secondary stress in Estonian. *Proceedings of Speech Prosody 2018*. Poznań: Adam Mickiewicz University, 602−606.

[2] Heldner, M. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics* 31, 39–62.

[3] Krull, D. 1997. Prepausal lengthening in Estonian: Evidence from Conversational speech. 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.

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

[5] Lehiste, I. 1968. Vowel quantity in word and utterance in Estonian. *Congressus Secundus Internationalis Fenno-Ugristarum*. Helsinki: Societas Finno-Ugrica, 293−303.

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

[7] Lippus, P., Asu, E. L., Teras, P., Tuisk, T. 2013. Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics* 41, 17−28.

[8] Mihkla, M., Sahkai, H. 2017. Lauserõhu akustilised korrelaadid eesti keeles [The acoustic correlates of phrasal stress in Estonian]. *Mäetagused* 68, 161−176.

[9] Murphy, P. J., McGuigan, K. G., Walsh. M., Colreavy, M. 2008. Investigation of a glottal related harmonics-to-noise ratio and spectral tilt as indicators of glottal noise in synthesized and human voice signals. *J. Acoust. Soc. Am.* 123, 1642–1652.

[10] Ortega-Llebaria, M., Prieto, P. 2007. Disentangling stress from accent in Spanish: Production patterns of the stress contrast in de-accented syllables. Prieto, P., Mascaró, J., Solé, M.-J. (Eds.) *Segmental and prosodic issues in Romance Phonology*. Amsterdam /Philadelphia: John Benjamins, 155−179.

[11] Ortega-Llebaria, M., Prieto, P. 2010. Acoustic Correlates of Stress in Central Catalan and Castilian Spanish. *Language and Speech* 54, 73−97.

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

[13] Sahkai, H., Mihkla, M., Kalvik, M.-L. 2015. Emfaas ja fookus eesti keeles [Emphasis and focus in Estonian]. *Journal of Estonian and Finno-Ugric Linguistics* 6, 97−122.

[14] Sluijter, A. M. C., Heuven, V. J. van 1996. Spectral balance as an acoustic correlate of linguistic stress. *J. Acoust. Soc. Am.* 100, 2471–2485.

[15] Traunmüller, H., Eriksson, A. 2000. Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451.