9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Acoustic correlates of secondary stress in Estonian

Eva Liina Asu, Pärtel Lippus

Institute of Estonian and General Linguistics, University of Tartu, Estonia
eva-liina.asu@ut.ee, partel.lippus@ut.ee

Abstract

The present study aims to add to the growing body of recent work addressing the acoustic correlates of secondary stress. Here the focus is on Estonian, a quantity language where the primary stress is fixed on the first syllable of the word and the placement of secondary stresses is determined by morphological constraints but typically coincides with odd-numbered syllables.

Words consisting of five and six CV syllables were analysed with respect to various acoustic measures relating to duration, pitch, and spectral characteristics. The results show that in Estonian secondary stress does not acoustically differ from unstress, calling into question the usefulness of the concept. This finding supports earlier results for several other languages (e.g. Hungarian, Brazilian Portuguese) where phonological secondary stress has been postulated but is not realised phonetically. It also underlines the crucial role of the primary stressed foot in the prosodic system of Estonian.

Index Terms: word stress, secondary stress, acoustic measures of stress, Estonian

1. Introduction

Traditionally, a phonological distinction is made between three levels of lexical stress: primary, secondary, and unstress (e.g. [1]). A vast majority of phonetic research on word stress has, however, focussed on the comparison of only two levels: stressed and unstressed. Acoustic analyses of stress in many languages show that the main correlates of primary stress are duration, pitch, intensity, vowel quality, and spectral characteristics (e.g. [2]). Considerably fewer studies have investigated the correlates of secondary stress or the acoustic distinction between primary and secondary stress, see [3], [4].

The present paper aims to address the issue of stress by focussing on the acoustic characteristics of secondary stress in Estonian, a quantity language where primary stress in most native words is fixed on the first syllable. The primary stressed syllable and the following unstressed syllable form a trochaic foot, which constitutes the core of the Estonian prosodic system, being also the domain of the three-way quantity distinction [5]–[8]. A foot in Estonian may be maximally trisyllabic. Disyllabic or trisyllabic words, thus, typically consist of one foot, whereas a tetrasyllabic word is normally made up of two disyllabic feet; words with five syllables are divided into a tetrasyllabic and a disyllabic foot, and words with six syllables contain either two trisyllabic or three disyllabic feet. Secondary stresses in Estonian fall on odd-numbered syllables (e.g. [9]). In longer words, the placement of secondary stresses is determined by the syllable structure and derivational structure of the word, e.g. certain suffixes such as, for instance, *-ik* or *-nna* always attract secondary stress. To a certain extent, the position of secondary stress is variable in Estonian. As to durational properties secondary stressed feet are supposed to behave like primary stressed feet [6].

Phonologically, secondary stress can be defined as a type of stress that is weaker than primary stress but stronger than unstress [10, p. 96], but phonetically secondary stress has proven to be rather elusive in many languages where it has been postulated, e.g. Brazilian Portuguese [11] and Hungarian [12], see [4] for an overview. A study on the lexical stress in Pitjantjatjara, a dialect of the Western Desert language of Australia, [13] found no acoustic evidence for secondary stress contrary to earlier phonological descriptions. Likewise, the lack of robust acoustic correlates for lower levels of stress has been observed in Polish pointing to the existence of only one stress per word [4]. A comparison of the cues of primary and secondary stress in accented and unaccented complex words in English [3] concludes that there is not much evidence for assuming a phonetic distinction between primary and secondary stress; primary and secondary stressed syllables differ only when the word is accented, suggesting that stress distinction in English should take into account accent-conditions.

An acoustic study on Estonian word stress [14] compared three levels of stress in different speaking styles considering four acoustic measures: duration, F0 mean, F0 standard deviation, and spectral emphasis. It was shown that the strongest correlate of stress was vowel duration, and that the secondary stressed vowels were always shorter than the primary stressed ones. The study did not, however, take into account the length of the stressed vowel or the structure of the stressed syllable, and therefore may not tell the whole story. In the same study, no difference was found between the F0 mean of secondary and unstressed vowels, while the results about the standard deviation of F0 were inconclusive. Spectral emphasis was shown to be the weakest measure of the four.

Durations of the secondary stressed feet were shown to be shorter in [15] whereas the duration ratios within the secondary stressed feet remained the same as in the primary stressed feet. In perception, it has been shown, though, that the Estonian listeners prefer F0 as a cue for stress, as opposed to Russian listeners who rely on duration [16]. Intensity has been shown not be a correlate of stress in Estonian [7], [17].

As our knowledge of the correlates of secondary stress in Estonian is somewhat patchy, this study aims to clarify it by testing the relevance of several acoustic measures normally used for stress. Based on previous work we can hypothesise that secondary stress behaves in durational terms like primary stress [6], i.e. the distinction between stressed and unstressed syllables is similar, but stronger in primary and weaker in secondary stressed feet [15]. Acoustically, we would expect secondary stressed syllables to have a shorter duration and a lower F0 than the primary stressed ones, and show no difference in spectral emphasis or intensity [14], [17].

10.21437/SpeechProsody.2018-122

2. Materials and methods

The materials, used previously in [15], consisted of 40 test words: 20 five-syllable words and 20 six-syllable words. All test words were made up of CV syllables where the vowel was always short (in Quantity 1), thus eliminating the factor of the three quantity degrees from the analysis. The segmental make-up of the words was not controlled for. All test words were placed in an accented position utterance-medially in different meaningful carrier sentences: e.g. *Külastas* ***sagedamini*** *sõpra* '(S/he) visited **more frequently** a friend'. *Lootis* ***tavalisemale*** *teole* '(S/he) was hoping for **a more ordinary** deed'.

The lists, where each test word occurred just once, were read by six informants (3 male and 3 female, 20-24 years old) who were all university students at the time of the recording, and spoke Standard Estonian. The recordings were conducted in a quiet setting using a Sony TCD-D 100 DAT recorder and an AKG D40S microphone.

Segmental boundaries of all target words were annotated with PRAAT [18]. As an initial step in the analysis, two experienced phoneticians independently marked the location of all primary and secondary stressed syllables on a separate tier based on their native-speaker auditory impression. Subsequently, all cases of divergence were re-examined and a consensus was reached. Two test words that were pronounced regularly as compounds with two primary stresses by all the speakers were excluded from the analysis as well as all instances of disfluencies. The final dataset comprised 216 test words of which 107 consisted of five syllables and 109 of six syllables. All five-syllable words contained two metrical feet: a disyllabic foot followed by a trisyllabic foot (83 words) or a trisyllabic foot followed by a disyllabic foot (24 words). Six-syllable words contained either three disyllabic (87 words) or two trisyllabic feet (22 words). In total 1189 CV syllables were analysed.

The analysis used various acoustic parameters that were measured in the vowel of each syllable: durational correlates (duration of the syllable onset consonant and duration of the vowel), pitch (average F0, standard deviation of F0, and F0 slope in the vowel calculated as $(F0_{max}-F0_{min})/(t_{max}-t_{min})$), mean intensity, spectral balance and spectral emphasis (as described in [19]).

As in Estonian disyllabic Quantity 1 feet the unstressed syllable is longer than the primary stressed one (c.f. [5], [7], [8]), all unstressed syllables in the data were divided into three different categories. Thus, for the purposes of the analysis the syllables were classified as follows: primary stressed (*Pri*), secondary stressed (*Sec*), unstressed syllable following primary stressed syllable (*Upri*), unstressed syllable following secondary stressed syllable (*Usec*), and unstressed syllables following another unstressed syllable (*Uuns*).

Statistical analysis was carried out in R [20] with the packages lme4 [21] and multcomp [22]. Each acoustic measure was tested with a linear mixed effect model for the effect of stress (5 levels described above), including random intercepts for the speaker and the test word. For post-hoc testing Tukey HSD was used.

3. Results

3.1. Durational properties

As seen in Figure 1, there was a significant effect of stress on the duration of syllable onset consonant ($F(4, 1153.3)=37.3$, $p<0.001$). Post-hoc testing showed that the duration of C in *Pri* (75 ms) is significantly longer than in all other syllable types ($p<0.001$), where it ranges between 55 and 59 ms. There was a significant difference between *Upri* vs. *Sec* and *Usec* ($p>0.01$) while the difference between *Sec*, *Usec* and *Uuns* is not significant.

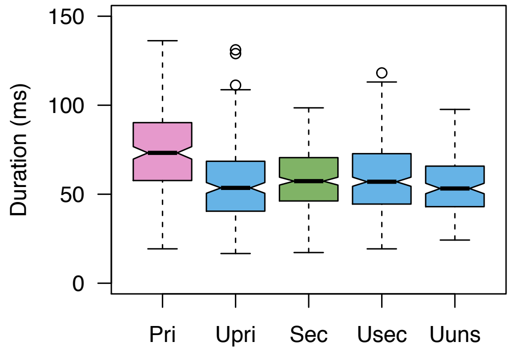

Figure 1: *Duration of the syllable onset consonant.*

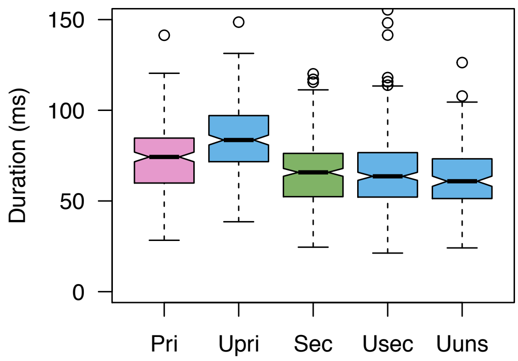

Figure 2: *Duration of the vowel.*

There was a significant effect of stress on vowel duration ($F(4, 1149.6)=75.2$, $p<0.001$). Post-hoc testing showed that the vowel duration in *Pri* (73 ms) is significantly shorter ($p<0.001$) than in the following unstressed syllable (85 ms), and significantly longer ($p>0.001$) than in all other syllables. The mean duration of *Sec*, *Usec* and *Uuns* ranges from 63 to 66 ms and the differences are not significant (see Figure 2).

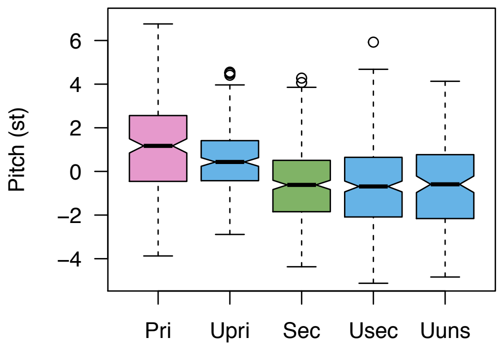

Figure 3: *Average F0 in the vowel.*

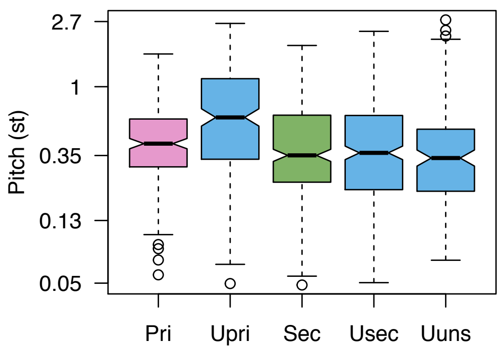

Figure 4: *Standard deviation of F0 in the vowel.*

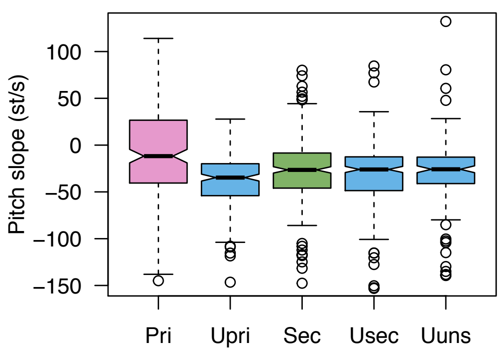

Figure 5: *F0 slope in the vowel.*

3.2. Pitch properties

There was a significant effect of stress on the average F0 in the vowel (*F*(4, 1145.7)=58.05, *p*<0.001). As seen in Figure 3 the mean F0 was normalised to the semitone scale with reference to the speaker's overall mean F0 (introducing negative values for syllables below the speaker's average pitch). Post-hoc testing showed that the mean F0 in *Pri* (1.2 st), and *Upri* (0.5 st) were significantly different from each other and that of the other syllables (-0.7 st; *p*<0.001). There were no significant differences between *Sec*, *Usec* and *Uuns*.

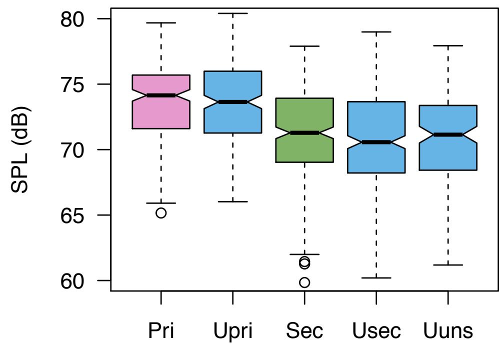

Figure 6: *Mean intensity of the vowel.*

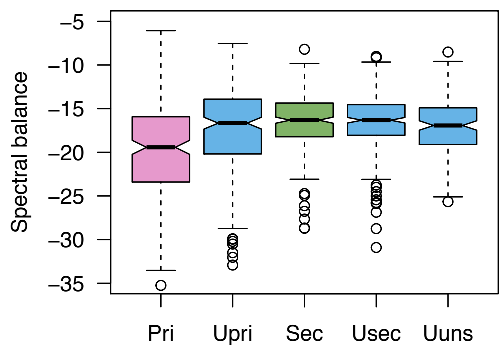

Figure 7: *Spectral balance in the vowel.*

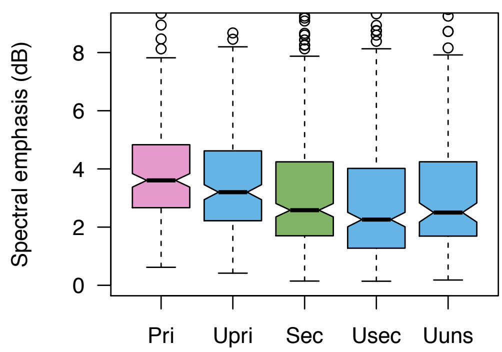

Figure 8*: Spectral emphasis in the vowel.*

There was a significant effect of stress on the standard deviation of F0 in the vowel (*F*(4, 1147.4)=16.2, *p*<0.001), which is the highest (0.8 st) in *Upri* (*p*<0.001). In all other syllables the standard deviation of F0 is 0.5 st and there are no significant differences (see Figure 4).

As seen in Figure 5, there was a significant effect of stress on F0 slope (*F*(4, 1129.1)=5.6, *p*<0.001). There is considerable variability in the values of pitch slope (in particular in *Pri*): on average it is -16 st/s in *Pri*, -37 st/s in *Upri*, -26 st/s in *Sec*, -28 st/s in *Usec* and -30 st/s in *Uuns*. Post-hoc testing showed that the difference is only significant between *Pri* and *Upri* (*p*<0.001), but not significant between all other combinations.

3.3. Intensity

As seen in Figure 6, there was a significant effect of stress on the mean intensity of the vowel ($F(4, 1136.8)=115.6$, $p<0.001$). The mean intensity in *Pri* and *Upri* (73 dB) is significantly higher ($p<0.001$) than in *Sec*, *Usec* and *Uuns* (71 dB). Post-hoc testing also showed a minute (0.5 dB) yet significant ($p<0.01$) difference between *Sec* and *Usec*.

3.4. Spectral properties

There was a significant effect of stress on spectral balance ($F(4, 1140)=39.4$, $p<0.001$). It can be seen in Figure 7 that the average spectral balance is significantly smaller in *Pri* (-20) compared to all other syllables ($p<0.001$). In *Upri* it was -18, which is significantly smaller ($p<0.01$) than in *Sec* and *Usec* (-16), but not significantly different from *Uuns* (-17). There are no significant differences between *Sec*, *Usec* and *Uuns*.

There was also a significant effect of stress on spectral emphasis ($F(4, 1146.3)=20.9$, $p<0.001$) as presented in Figure 8. Post-hoc testing showed that the difference between *Pri* (3.9 dB) and *Upri* (3.5 dB) is not significant, but there is a significant difference between *Pri* and the remaining syllables ($p<0.001$). *Upri* has a significantly higher spectral emphasis than that of *Sec* (3.3 dB) and *Usec* (2.9 dB) ($p<0.05$), but not significantly different from *Uuns* (3.2 dB). Finally, the spectral emphasis in *Usec* is significantly lower than in *Sec* and *Uuns* ($p<0.05$), but there is no difference between *Sec* and *Uuns*.

4. Discussion

The aim of the present study was to investigate the acoustic correlates of secondary stress in Estonian using five-syllable and six-syllable words. Based on earlier findings, we expected the secondary stressed syllables to be shorter and have a lower F0 than the primary stressed ones, but display no difference in intensity or spectral emphasis.

The results showed that the secondary stressed syllables were indeed significantly different from the primary stressed syllables with respect to all the studied acoustic measures, including also intensity and spectral properties, but did not at the same time differ significantly from the unstressed syllables, except the unstressed syllable following the primary stressed syllable (*Upri*) that has been shown to behave differently [5], [8].

Earlier results [7], [17] have showed that intensity is not a reliable correlate of stress in Estonian, but these studies have focussed on disyllabic feet. In this respect, the findings of the present study support the earlier work, as there was no significant difference in the intensity in the primary stressed (*Pri*) and the following unstressed syllable (*Upri*) in the present material either. However, as this study analysed longer words, it emerges that the intensity in the primary stressed foot is significantly higher than in all the other syllables. There being a significant difference in the intensities of the secondary stressed syllable (*Sec*) and the following unstressed syllable (*Usec*) implies that the secondary stressed foot is different from the primary stressed foot.

The results for the F0 reflect the accentuation pattern of the test words which all appeared in the accented position. Most prominent were the primary stressed syllables that had the highest average F0 (as well as F0 slope) while the following unstressed syllables were lower followed by all the other syllables that showed no significant difference. The standard deviation of F0 was the largest in the unstressed syllable following the primary stressed syllable which implies that the main F0 fall coincides with the second syllable of the word and that the F0 is most variable within this syllable.

Contrary to what was hypothesised, the present findings imply that the secondary stressed feet do not behave like the primary stressed feet. Importantly, the findings underline the crucial role of the primary stressed foot in the prosodic system of Estonian, where a disproportionally large number of options of the phonological system has been granted to the primary stressed syllable, and consequently, all other syllables are much poorer [23, p. 149]. For instance, of the nine vowels of the primary stressed syllable only five can occur in the non-primary syllables, or of the 36 diphthongs just three occur in the non-primary syllables (only starting from the third syllable).

The fact that the secondary stressed syllables did not differ from the unstressed syllables leads us to ask whether it is useful to distinguish secondary stress at all, at least in phonetic terms. This is in line with recent work on several other languages [3], [4], [13] that seems to suggest that there is no phonetic evidence for the previously postulated phonological secondary stress. Why do we perceive it then? Tabain et al. [13] claim that a native speaker's stress perception is guided by a higher level grammatical knowledge. Likewise Chomsky and Halle [2, pp. 149–150] argue that if one is familiar with the language system one is able to hear the predicted phonetic patterns although there is no acoustic or physical reality that would support this. It might be possible to clarify this issue with further work on the perception of secondary stress.

Finally, a note should be made about creaky voice as a stress correlate. Interestingly it was found in [24] that it was the secondary stressed syllables (not unstressed) that were most likely to be creaky, although both unstressed and secondary stressed syllables were significantly more creaky than the primary stressed syllables. An abundance of creaky voice in non-initial syllables was also noticed in the data of the word stress study [14] which prompted the suggestion that creaky voice should be included as a measure in the stress model. The present data contained only five syllables labelled as creaky, one of them being a primary stressed and four word-final syllables.

5. Conclusions

The results of the present study show that the secondary stressed syllables in Estonian are not acoustically different from the unstressed syllables with respect to various durational, pitch-related and spectral measures normally used to investigate correlates of stress. The only exception is the unstressed syllable following the primary stressed syllable that behaves differently from other unstressed syllables. This finding also underlines the important role of the primary stressed foot in the prosodic system of the Estonian language, as the primary stressed foot seems to be the domain for durational as well as intensity differences, among others.

6. Acknowledgements

The study was supported by Estonian Research Council grant IUT2-37. The paper is dedicated to the 100th anniversary of the Republic of Estonia.

7. References

[1] B. Hayes, *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press, 1995.

[2] I. Lehiste, *Suprasegmentals*. Cambridge Mass.: M.I.T. Press, 1970.

[3] I. Plag, G. Kunter, and M. Schramm, 'Acoustic correlates of primary and secondary stress in North American English', *Journal of Phonetics*, vol. 39, no. 3, pp. 362–374, Jul. 2011.

[4] L. Newlin-Łukowicz, 'Polish stress: looking for phonetic evidence of a bidirectional system', *Phonology*, vol. 29, no. 02, pp. 271–329, Aug. 2012.

[5] I. Lehiste, 'Segmental and syllabic quantity in Estonian', in *American Studies in Uralic Linguistics*, T. A. Sebeok, Ed. Bloomington: Indiana University Publications, 1960, pp. 21–82.

[6] A. Eek, 'Units of temporal organisation and word accents in Estonian', *LU*, vol. 26, no. 4, pp. 251–263, 1990.

[7] A. Eek and E. Meister, 'Simple perception experiments on Estonian word prosody', in *Estonian prosody: Papers from a symposium*, I. Lehiste and J. Ross, Eds. Tallinn: Institute of Estonian Language, 1997, pp. 71–99.

[8] P. Lippus, E. L. Asu, P. Teras, and T. Tuisk, 'Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian', *Journal of Phonetics*, vol. 41, no. 1, pp. 17–28, Jan. 2013.

[9] T.-R. Viitso, 'Phonology, morphology and word formation', in *Estonian language*, M. Erelt, Ed. Tallinn: Estonian Academy Publishers, 2003, pp. 9–92.

[10] P. Roach, *English phonetics and phonology: A practical course*. Cambridge: Cambridge University Press, 2000.

[11] P. Arantes and P. A. Barbosa, 'F1 and spectral correlates of secondary stress in Brazilian Portuguese', in *Proceedings of the Speech Prosody 2008 Conference*, Campinas, 2008, pp. 559–562.

[12] S. Blaho and D. Szeredi, 'Secondary stress in Hungarian: (morpho)- syntactic, not metrical', in *Proceedings of the 28th West Coast Conference on Formal Linguistics*, Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, Eds. Somerville, MA, USA: Cascadilla Proceedings Project, 2011, pp. 51–59.

[13] M. Tabain, J. Fletcher, and A. Butcher, 'Lexical stress in Pitjantjatjara', *Journal of Phonetics*, vol. 42, pp. 52–66, Jan. 2014.

[14] P. Lippus, E. L. Asu, and M.-L. Kalvik, 'An acoustic study of Estonian word stress', in *Social and Linguistic Speech Prosody. Proceedings of the 7th international conference on Speech Prosody*, N. Campbell, D. Gibbon, and D. Hirst, Eds. Dublin, 2014, pp. 232–235.

[15] P. Lippus, K. Pajusalu, and P. Teras, 'The Temporal Structure of Penta- and Hexasyllabic Words in Estonian', in *Proceedings of the 3rd International Conference Speech Prosody*, R. Hoffmann and H. Mixdorff, Eds. Dresden: TUDpress, 2006, pp. 759–762.

[16] A. Eek, 'The perception of word stress: A comparison of Estonian and Russian', in *Honor of Ilse Lehiste. Ilse Lehiste pühendusteos*, R. Channon and L. Shockey, Eds. Dordrecht, Providence: Foris Publications, 1987, pp. 19–32.

[17] M.-L. Kalvik and M. Mihkla, 'Modelling the Temporal Structure of Estonian Speech', in *Human Language Technologies – The Baltic Perspective - Proceedings of the Fourth International Conference Baltic HLT 2010*, Amsterdam, The Netherlands, The Netherlands, 2010, pp. 53–60.

[18] P. Boersma and D. Weenink, *Praat: doing phonetics by computer*. 2017.

[19] H. Traunmüller and A. Eriksson, 'Acoustic effects of variation in vocal effort by men, women, and children', *The Journal of the Acoustical Society of America*, vol. 107, no. 6, pp. 3438–3451, 2000.

[20] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017.

[21] D. Bates, M. Maechler, B. Bolker, and S. Walker, *lme4: Linear mixed-effects models using Eigen and S4*. 2014.

[22] T. Hothorn, F. Bretz, and P. Westfall, 'Simultaneous Inference in General Parametric Models', *Biometrical Journal*, vol. 50, no. 3, pp. 346–363, 2008.

[23] M. Hint, *Häälikutest sõnadeni*, 2. Tallinn, 1998.

[24] K. Aare, P. Lippus, and J. Šimko, 'Creak as a Feature of Lexical Stress in Estonian', in *Proceedings of Interspeech 2017*, 2017, pp. 1029–1033.