The loss of intervocalic short /h/ in Estonian

Pire Teras

Institute of Estonian and General Linguistics, University of Tartu, Estonia
pire.teras@ut.ee

ABSTRACT

Estonian is a Finno-Ugric language that is characterised by the voicing or loss of both word-initial and intervocalic short /h/. The present paper aims to determine the acoustic characteristics of Estonian words where a short intervocalic /h/ has been lost. Disyllabic CVhV-words where the loss of /h/ has occurred are analysed. Two tendencies have been found: the loss of /h/ results either in a long monophthong or diphthong (CVV) or in a boundary between two vowels (CV.V). In most cases, the duration of V2 is longer than V1 but there are also variants where the loss of /h/ has lengthened the duration of the first vowel. Variants with a boundary are characterised by a slight dip in intensity between two vowels. Both variants' pitch contours resemble those of Estonian quantity 3 words that are characterised by an early F0 turning point and falling pitch in accented words.

Keywords: phonetics, variation, glottal fricative, sound loss, Estonian.

1. INTRODUCTION

Previous studies have shown that the voicing of short intervocalic unvoiced obstruents occurs quite frequently in Estonian [1]. Ermus [2, p. 36] pointed out that the reduction of short intervocalic plosives can result in this loss: for velar /k/ in 6% and alveolar /t/ in 4% of all analysed tokens.

Estonian /h/ is a voiceless glottal fricative. Teras [1] has suggested that the main variant of short intervocalic /h/ is its voiced variant: this occurred in 70% of analysed tokens. However, it appears that the loss of intervocalic short /h/ is also quite frequent: the reduction of /h/ resulted in a complete loss in 21% of analysed tokens and, additionally, in 3% of cases in complete loss and vowel shortening. This loss occurs first of all in informal situations, it is more frequent in the speech of male speakers and in unaccented words [1].

Voicing and loss or elision of consonants have both been connected with the term *lenition*, which means that the constriction degree or length of a consonant reduces as the result of minimising more effort-demanding articulatory gestures [3, pp. 313–314]. This phenomenon is well-known in the world's languages and the most typical context for it is intervocalic [3, p. 314]. Garellek et al. have analysed the voicing of glottal consonants and found that "both [h] and [ɦ] are likely to be as voiced as breathy vowels" [4, p. 23].

Melero García [5, p. 11] has shown that complete loss of intervocalic /d/ in Madrid Spanish leads to changes in intensity curves. Intervocalic /d/ was completely lost in 14% of analysed words, and in most cases, there was a slight dip in intensity between vowels (intensity ratio between 0.95 and 1). A study on the elision of /h/ in the Shirazi dialect of Persian shows that this loss occurs in weak positions (intervocalic, coda, clusters) and lengthens the preceding vowel [6]. Davidson [7] has shown that in Hawaiian in some cases an intervocalic glottal stop is realised – among other variants (e.g., creaky voice, full glottal closure) – as an intensity dip. The percentage of such cases was slightly greater when the flanking vowels were identical.

Recent studies on the South Estonian Leivu dialect have shown that in Leivu the loss of intervocalic short /h/ has led to the development of broken tone [8], [9]. Compared to plain tone words, broken tone words in Leivu are characterised by longer vowel duration (the first vowel is often lengthened), an intensity dip between two vowels, a more consistent early F0 turning point, and in some cases by a laryngealisation period during the second vowel. Loss of intervocalic short /h/ has also been considered to be one reason for the development of broken tone in Livonian [9], [10] and the South Estonian Lutsi dialect [11].

Teras [1] has made observations about tendencies in the pronunciation of words where /h/ has been lost: originally disyllabic quantity 1 (Q1) CVhV-words (e.g., *kohe* [koe̯ːe̯] 'immediately') are pronounced either as monosyllabic quantity 3 (Q3) CVV-words, where the syllable nucleus is a long monophthong or diphthong (e.g., *kohe* [koeː]), or they are pronounced with a boundary so that the first syllable vowel lengthens (e.g., *kohe* [koː.e]). However, no closer acoustic analysis of such variants has yet been done and it is not known how frequently such variants occur and whether some other variants are possible.

The present study aims to determine what acoustically characterises the variants of CVhV-words where the loss of intervocalic short /h/ occurs. The research questions are as follows:

(1) What variants result from the loss of /h/ and what is their percentage?
(2) What characterises vowel duration and duration ratios in these variants?
(3) What can be said about the F0 contours of the variants?
(4) How often does the dip in intensity occur between two vowels?

2. MATERIALS AND METHODS

The Phonetic Corpus of Estonian Spontaneous Speech [12] was used for data retrieval. The corpus contained 77 hours of speech from 111 speakers at the time of data retrieval. The corpus is manually segmented both on the word level (annotated using Estonian orthography) and segmental level (annotated using SAMPA transcription). Disyllabic CVhV-words were searched for at the word level. The same material as in Teras [1] was used for acoustic analysis, but variants with consonant loss were reinspected and two misanalysed variants were left out.

In Estonian, there are 9 vowels (close [i y u], mid [e ø ɤ o], open [æ ɑ]) that occur in the primary stressed first syllable, and 5 of these [i u e o ɑ] occur in successive syllables. The second syllable /e/ is often pronounced lower: [ę] ~ [æ]. [13] The current material did not contain [ø] in the first syllable or [o] in the second syllable.

There are three quantity degrees in Estonian: words with short stressed syllables are Q1 and those with long stressed syllables are Q2 or Q3. All monosyllabic words are considered to be Q3, e.g., *maa* [mɑː] ‘land, earth’, *kae* [kɑeː] ‘cataract’. The three quantity degrees are differentiated by the duration ratios of syllables (2/3, 3/2, and 3/1, respectively), and Q2 and Q3 are additionally differentiated by the location of the F0 turning point (late or early). See, e.g., [13], [14]. CVhV-words are Q1 where both vowels are phonologically short. However, Q1 words are characterised by the lengthening of the second syllable vowel (V1/V2 ratio 0.5–0.8), e.g., *raha* [rɑhɑ̆·] ‘money’.

Of all CVhV-words (1173), there were 250 tokens where a loss of /h/ occurred (see Table 1).

Table 1. The number of analysed tokens.

Variant	Number of tokens		
	No intensity dip	Intensity dip	Total
CVV	66	22	88
CV1V2	129	33	162
Total	195	55	250

These tokens were analysed acoustically using Praat [15]. The duration of all segments was measured. In the case of two different vowels, the boundary was set dividing the transition between the two vowels in half. When /h/ between two identical vowels was lost, a dip in intensity was used to mark the boundary.

Fundamental frequency and intensity were measured from the beginning of the first vowel and the end of the second vowel, and also from the turning point (TP) where the F0 or intensity started to fall. The location of the TP was calculated as a percent of the total duration of the vowels. If there was a dip in intensity during the vowels, then an intensity maximum of the first and the second vowel was measured as well as an intensity minimum during the dip. See an example of segmentation and annotation in Figure 1.

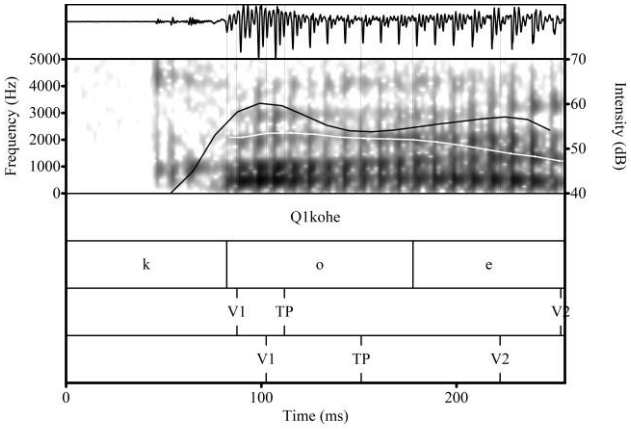

Figure 1. An example of segmentation and annotation of the word *kohe* [koˑ.e] ‘immediately’ (a male speaker). A sound wave, spectrogram, intensity contour (solid black line), F0 contour (dotted white line, scale 60–180 Hz). The annotation shows four tiers: word, sounds, F0 (V1, V2 – the beginning and the end, TP – the turning point), and intensity (V1, V2 – maximum, TP - minimum) measurement points.

Statistical analysis was carried out in R [16]. Descriptive statistics included the mean and median values. Average F0 values were calculated separately for female and male speakers. The chi-square test was used to determine significant associations between variants and linguistic factors: accentuation (levels: accented, unaccented), V1 and V2 (levels: close, mid, open), utterance position (levels: initial, medial, final), and adjacent vowels (levels: different, identical).

3. RESULTS AND DISCUSSION

3.1. Variants and duration ratios

Intervocalic /h/ can be lost between two identical or two different vowels. It appears that the loss of intervocalic /h/ results in four variants. The average duration of segments in these variants is shown in

Figure 2 and the duration ratios of vowels in the three variants are given in Figure 3.

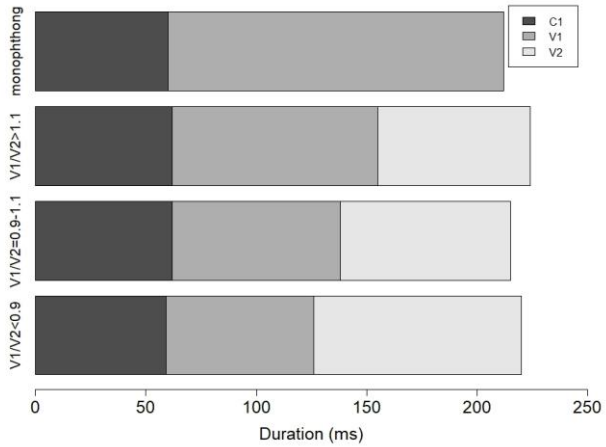

Figure 2. Sound durations (C1, V1, V2) for four variants of CVhV words where intervocalic /h/ has been lost.

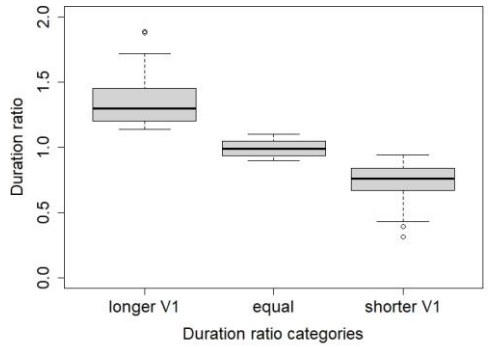

Figure 3. The duration ratio of V1/V2 for three variants: longer V1 than V2 (35 tokens), similar duration for V1 and V2 (48 tokens), shorter V1 than V2 (101 tokens).

The variants are as follows:

(1) two identical vowels are pronounced as a long monophthong (VV 152 ms);
(2) V1 is longer than V2 (V1 93 ms, V2 69 ms, V1/V2 = 1.35);
(3) two vowels have similar duration (V1 76 ms, V2 77 ms, V1/V2 = 0.99);
V1 is shorter than V2 (V1 67 ms, V2 94 ms, V1/V2 = 0.73).

Two identical vowels were pronounced in most cases as a long monophthong (66 tokens, 26% of all tokens). However, there were an additional 22 tokens with a dip in intensity between two identical vowels. These tokens were grouped into three variants.

The most frequent variant is the variant where V1 is shorter than V2 (41% of all tokens). This vowel duration ratio (0.73) is characteristic of Estonian Q1 words [14] but also of the vowel duration ratio for diphthongs in Q3 words where V2 tends to be longer than V1 [17], e.g., *lai* [lɑiː] 'broad'. It can be hypothesised that listeners would perceive these tokens as Q3 words.

In 14% of all tokens, the duration of V1 was longer than the duration of V2 (ratio: 1.35). Lengthening of the preceding vowel resulting from the loss of intervocalic /h/ has also been found in the South Estonian Leivu dialect [8] and in the Shirazi dialect of Persian [6]. In 19% of all tokens, V1 and V2 had quite similar durations (ratio: 0.99).

When the percentage of accented to unaccented tokens is similar (44% vs. 56%) for tokens with shorter V1 than V2, then there is a smaller percentage of accented than unaccented words (40% vs. 60%) for tokens with longer V1 than V2 and with a similar duration for V1 and V2. However, the chi-square test shows no significant association between the variants and accentuation or between other linguistic factors.

3.2. Pitch contours

Previous research has shown that Estonian accented Q1 and Q2 words are characterised by a late F0 TP in the stressed syllable followed by an F0 fall in the second syllable, whereas accented Q3 words have an early F0 TP followed by a fall already in the stressed syllable (e.g., [14]). Unaccented words do not have F0 TP and their contours are flat [14].

The current material contained 109 accented and 141 unaccented words. All accented words had a falling pitch contour. In unaccented words, the pitch contours were rather flat: on average there was only a slight fall from beginning to end (12–15 Hz). In 13 unaccented tokens, the pitch rose slightly (8–19 Hz).

The pitch contours of accented words will be analysed in more detail next. In Figure 4, the location of F0 TP can be seen. In accented words, in almost all cases (96 tokens, 88% of accented tokens) an early F0 TP occurred at 25% of the total vowel duration. In only 13 tokens (12% of accented tokens), the TP occurred late at 61% of the total vowel duration.

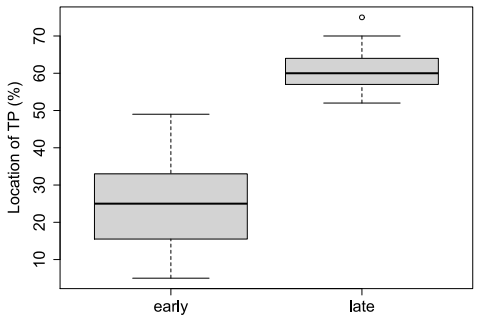

Figure 4. The location of the F0 turning point in accented words.

Figure 5 shows the average F0 contours of tokens where there was an early F0 TP (average values were calculated separately for female and male speakers). The F0 range from the beginning to the end is 32–36

Hz. It can be said that the F0 contours of these words are similar to Estonian Q3 words that are also characterised by an early F0 turning point [14].

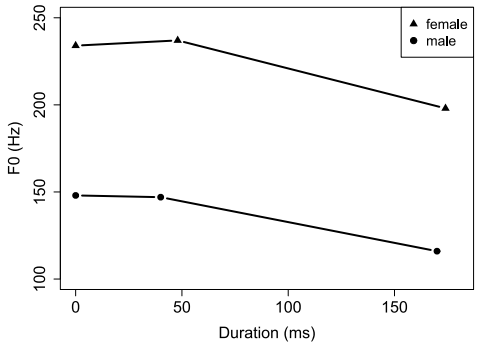

Figure 5. Average F0 contours of accented tokens with an early F0 turning point (points: the beginning of V1, TP, and end of V2).

3.3. Intensity

In the case when an intervocalic /h/ is present, there is a dip in intensity between adjacent vowels. When an intervocalic /h/ is lost, in most cases (78%, see table 1) there is no dip in intensity. In these tokens, there is in most cases (71%) an early intensity TP (occurring at 27%) and in other cases a late intensity TP (at 69%). Average intensity contours are presented in Figure 6.

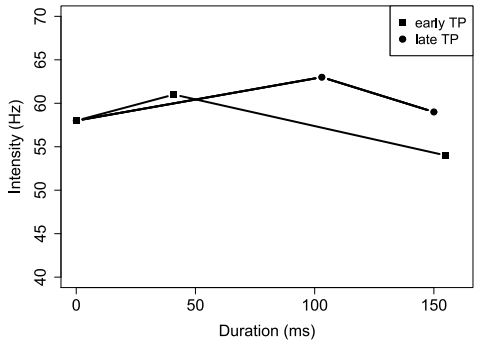

Figure 6. Average intensity contours of accented tokens with an early and late intensity turning point (points: the beginning of V1, TP, and end of V2).

In 22% of cases, there was a slight dip in intensity during two vowels. When /h/ is lost between two identical vowels, this percentage is greater (25%) than when it is lost between two different vowels (20%). The ratio of the intensity minimum (61 dB) and the quotient of V1 and the V2 intensity maximum (both 63 dB) was calculated for these tokens. The ratio varied between 0.92 and 0.99, the average ratio is 0.97. These values are similar to those found in Madrid Spanish where intervocalic /d/ was reduced [5]. Such a dip in intensity has also been found in Hawaiian when an intervocalic glottal stop is lenited, c.f. [7].

The chi-square test shows a significant association between the dip and accentuation with a small effect size ($\chi^2 = 4.67$, df = 1, $p < 0.05$, Cramer's V = 0.137), and between the dip and preceding vowel with a small effect size ($\chi^2 = 11.47$, df = 2, $p < 0.05$, Cramer's V = 0.214). Pearson's residuals suggest that the dip occurs more than expected in accented words and when preceded by an open vowel. No other significant associations were found.

4. CONCLUSION

Lenition of intervocalic obstruents is well-known in the world's languages. Its final stage is complete elision or loss. The loss of intervocalic short plosives and /h/ occurs in Estonian. When an intervocalic /h/ between two identical vowels is lost, in most cases the two vowels are pronounced as a long monophthong but in some cases, there is a dip in intensity between the vowels, and, in most cases, V2 is longer than V1. The latter is most characteristic also of variants where an intervocalic /h/ is lost between two different vowels. The ratio of V1/V2 in such variants is 0.73. Such a ratio also characterises diphthongs in Q3 words. However, the loss of /h/ can also lengthen the first syllable vowel. The two vowels can also have quite similar durations (V1/V2 ratio 0.99).

Accented words, where an intervocalic /h/ has been lost, are characterised by an early F0 TP, which is also characteristic of Estonian Q3 words. The intensity turning point is also early in most cases, but in 22% of analysed tokens, there was a slight dip in intensity between the two vowels (the ratio of the intensity minimum and two intensity maximums of vowels was 0.97). The dip in intensity occurs more in accented than in unaccented words and after an open vowel.

It can be concluded that loss of intervocalic /h/ can result in a prosodic change: originally disyllabic Q1 words are pronounced like monosyllabic Q3 words with open long syllables where the second vowel lengthens and F0 is falling. However, there are also variants where the loss of /h/ lengthens the first vowel, resulting in a different pattern of duration ratios for the two adjacent vowels and most likely indicating that there can still be a syllable boundary between the two vowels despite the loss of /h/. The dip in intensity that was found in some tokens also seems to refer to such a boundary.

5. ACKNOWLEDGEMENTS

This research has been supported by the University of Tartu, grant PHVEE19903. The author would like to thank the two anonymous reviewers for their comments on and suggestions for this article, and Uldis Balodis for his English-language editing.

6. REFERENCES

[1] P. Teras, ‘The phonetic variation of short intervocalic /h/ in Estonian’, in *Speech Prosody 2018*, ISCA, 2018, pp. 883–887. doi: 10.21437/SpeechProsody.2018-178.

[2] L. Ermus, ‘Eesti keele lühikeste klusiilide häälduse variatsioon ja seda mõjutavad tegurid’, *Mäetagused*, vol. 69, pp. 27–52, 2017, doi: 10.7592/MT2017.68.ermus.

[3] R. Kirchner, ‘Consonant lenition’, in *Phonetically Based Phonology*, B. Hayes, D. Steriade, and R. Kirchner, Eds., Cambridge: Cambridge University Press, 2004, pp. 313–345. doi: 10.1017/CBO9780511486401.010.

[4] M. Garellek, Y. Chai, Y. Huang, and M. Van Doren, ‘Voicing of glottal consonants and non-modal vowels’, *J. Int. Phon. Assoc.*, pp. 1–28, 2021, doi: 10.1017/S0025100321000116.

[5] F. Melero García, ‘Variación sociofonética: comparación de análisis graduales y categóricos’, *Cuad. Lingüíst. Hispánica*, no. 37, 2021, doi: 10.19053/0121053x.n37.2021.12184.

[6] N. Rashidi and M. Shokrollahi, ‘Elision of /h, ?/ in the Shirazi Dialect of Persian (SHDP): An Optimality Theory Based Analysis’, *Raspr. Instituta Za Hrvat. Jez. Jezikoslovlje*, vol. 36, no. 1, pp. 135–162, 2010.

[7] L. Davidson, ‘Effects of word position and flanking vowel on the implementation of glottal stop: Evidence from Hawaiian’, *J. Phon.*, vol. 88, p. 101075, Sep. 2021, doi: 10.1016/j.wocn.2021.101075.

[8] P. Teras, ‘Broken tone in Leivu CV’V-words’, *J. Est. Finno-Ugric Linguist.*, vol. 12, no. 2, pp. 169–190, 2021, doi: 10.12697/jeful.2021.12.2.07.

[9] P. Teras and T. Tuisk, ‘Broken tone in Livonian and Leivu’, in *1st International Conference on Tone and Intonation (TAI)*, ISCA, 2021, pp. 26–30. doi: 10.21437/TAI.2021-6.

[10] T.-R. Viitso, ‘Livonian and Leivu: Shared innovations and problems’, *Linguist. Ural.*, vol. 45, no. 4, pp. 269–282, Dec. 2009, doi: 10.3176/lu.2009.4.03.

[11] U. Balodis, K. Pajusalu, and P. Teras, ‘Broken tone in South Estonian dialects in Latvia’, vol. 24, pp. 98–114, 2016.

[12] P. Lippus, K. Aare, A. Malmi, T. Tuisk, and P. Teras, ‘Phonetic Corpus of Estonian Spontaneous Speech v1.2’, 2021, doi: 10.23673/RE-293.

[13] E. L. Asu and P. Teras, ‘Estonian’, *J. Int. Phon. Assoc.*, vol. 39, no. 3, pp. 367–372, 2009, doi: 10.1017/S002510030999017X.

[14] P. Lippus, E. L. Asu, P. Teras, and T. Tuisk, ‘Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian’, *J. Phon.*, vol. 41, no. 1, pp. 17–28, 2013, doi: 10.1016/j.wocn.2012.09.005.

[15] P. Boersma and D. Weenink, ‘Praat: doing phonetics by computer [Computer program].’ 2022. Accessed: Dec. 17, 2022. [Online]. Available: http://www.praat.org/

[16] R Core Team, ‘R: A Language and Environment for Statistical Computing’. Vienna, Austria, 2022. [R Foundation for Statistical Computing]. Available: https://www.R-project.org/

[17] P. Lippus and E. L. Asu, ‘The timing and quality of diphthong components in spontaneous Estonian’, in *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds., Canberra, Australia: Australasian Speech Science and Technology Association Inc, 2019, pp. 1139–1143.