Perception of quantity in Estonian (Part II).

Krull, D.

journal: Proceedings of Fonetik, TMH-QPSR
volume: 44
number: 1
year: 2002
pages: 057-060

http://www.speech.kth.se/qpsr
Perception of quantity in Estonian (Part II)

Diana Krull and Hartmut Traunmüller
Dept. of Linguistics, Stockholm University

Abstract

In the experiment reported here, the speech rate of a short preceding or following context was manipulated in addition to that of a V, C or VC-sequence that carried a quantity distinction. The results showed that the durations of the other segments within the same two-syllable rhythmic foot and the presence or absence of a third syllable contributed to quantity perception.

Introduction

This paper is a sequel to an earlier study (Krull & Traunmüller, 2000), where we investigated the effect of changes in segment durations on the perception of quantity in Estonian. We manipulated the speech rate of a short preceding or following context as well as that of the vowel or consonant carrying the quantity distinction. The present study is a complement to the earlier one in that we have included also words where the quantity distinction is carried by a vowel-consonant combination.

The three-way quantity opposition in Estonian, which shows itself in accented positions within a rhythmic foot, is characterized by certain duration ratios between the accented first and the second syllable (Lehiste 1960; Eek 1980a, b). These relations have been shown to remain stable also in spontaneous speech (Engstrand & Krull, 1994). The perceptual importance of the duration of the first, stressed syllable is generally accepted. However, the role of the second unstressed syllable, whose duration is inversely related with that of the first, has been questioned. According to one account, the duration ratio between the two syllables is considered important (e.g. Lehiste 1960, 1997; Eek 1980a, b). According to a different account (Hint 1998), it is the first syllable alone that determines quantity. The existence of a second syllable is necessary, though, because quantity distinctions are not possible in monosyllables.

The results of our earlier investigation (Krull & Traunmüller, 2000) showed that segments preceding and following the segment that carries the quantity distinction contribute to the perception of quantity, and that this effect is in contradiction with the hypotheses that ascribe decisive perceptual importance to the duration ratios between the first and second syllables of a rhythmic foot. A lengthening of [sa] in the first syllable of [sat-tu] had an effect similar to that of lengthening the [u] of the second syllable: it made [t] seem shorter. According to the syllabic ratio hypothesis, the effect of lengthening [sa] should be opposite to that of lengthening [u].

In the preceding study, word forms were used where either the consonant or the vowel was short, long or overlong: CVCV, CV:CV, CV::CV, CVC::V, CV:C::V. The investigation reported in part in this paper covered all the seven word forms distinguished by quantity that are possible in a word initial foot in Estonian. Thus, it included also word forms in which the quantity distinction is carried by a vowel-consonant combination: CV:C::V and CV::C::V.

Method

Stimuli

The stimuli were obtained by manipulating the durations of selected sections of recordings of the word saate [sa:t:e] ‘get’ (2nd p. pl.) produced by a professional female speaker, preceded by ja [ja] ‘and’, in isolation, and followed by ka [ka] ‘also’, with list reading intonation. This word had been chosen bearing in mind that most of the forms: [sate], [sa:te], [sa::te], [sat:e], [sat::e] and [sa::t::e] are also common words. The segment durations of the three original utterances are listed in Table 1.

A comparison of the segment durations in Table 1 reveals a substantial final lengthening effect and also a minor effect of initial lengthening. These effects can be analyzed as follows: Initial word lengthening: [sate] 5.4 %, [at:] 4.6 % - Initial consonant [s] lengthening: 9.7 % (4.9 % more than [at:]) - Final word lengthening: [sate] 40.9 %, [at:] 25.6 % - Final vowel [e] lengthening: 61.2 % (28.3 % more than [at:]). These values are not necessarily representative of Estonian in general.
Table 1. Segment durations in ms and in % of the duration of \([a:t:]\) in the original utterances.

<table>
<thead>
<tr>
<th></th>
<th>[ja]</th>
<th>[s]</th>
<th>[a:]</th>
<th>[t:]</th>
<th>[e]</th>
<th>[ka]</th>
<th>[a:t:]</th>
<th>[sa:t:e]</th>
</tr>
</thead>
<tbody>
<tr>
<td>duration (ms)</td>
<td>* 96</td>
<td>137</td>
<td>152</td>
<td>101</td>
<td>378</td>
<td>289</td>
<td>486</td>
<td></td>
</tr>
<tr>
<td>% of [a:t:]</td>
<td>* 33.2</td>
<td>47.4</td>
<td>52.6</td>
<td>34.9</td>
<td>130.8</td>
<td>100</td>
<td>168.2</td>
<td></td>
</tr>
<tr>
<td>duration (ms)</td>
<td>* 121</td>
<td>175</td>
<td>188</td>
<td>201</td>
<td>* 363</td>
<td>685</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of [a:t:]</td>
<td>* 33.3</td>
<td>48.2</td>
<td>51.8</td>
<td>55.4</td>
<td>* 100</td>
<td>188.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duration (ms)</td>
<td>160</td>
<td>105</td>
<td>176</td>
<td>171</td>
<td>198</td>
<td>* 347</td>
<td>650</td>
<td></td>
</tr>
<tr>
<td>% of [a:t:]</td>
<td>46.1</td>
<td>30.3</td>
<td>50.7</td>
<td>49.3</td>
<td>57.1</td>
<td>* 100</td>
<td>187.3</td>
<td></td>
</tr>
</tbody>
</table>

The durations of the [a:] and [t:] in [sa:t:e] (occlusion + burst) were modified in steps of, nominally, a factor \(2^n/8\) with \(-8 < n < 8\), and with 7 or 8 even or uneven values of \(n\) used in each series of stimuli. In order to obtain phase-clean joints for the [a:], a deviation within \(\pm 1/2\) pitch period from the nominal duration was tolerated and neglected in the evaluation. The durations of either those parts of the utterances that preceded or that followed the [a:t:]-sequence were modified similarly, allowing \(n\) to take the values \(-2, 0,\) and \(+2\). Quantity space was traversed in 7 ways: 3 with a constant duration of the [a], 3 with a constant duration of the [t] and 1 diagonal way in which [a] and [t] had been varied together (see Figure 1).

**Listeners and Procedure**

The stimuli, which had been recorded on a CD, were presented through headphones to three groups of Estonian university students, 5 male and 18 female in all. There were 1071 stimuli separated by 0.6 s of silence and arranged in 23 x 7 = 161 series with additional pauses of 1 s in between. Each stimulus was presented once in a block with successively increasing duration of the segment in focus. There were some additional series — not mentioned in the following — in which the stimuli were presented in reversed order.

**Results**

In order to obtain an overview of the results, the boundaries between perceived Q1 and Q2 (or Q3) and between Q2 (or Q1) and Q3 were calculated in each series based on linear interpolation between the identifications obtained for the stimuli that were closest to the boundary. In a few cases, extrapolation was applied, but this was only allowed where the number of votes for the other side reached at least a third. In a few series, the number of Q1 responses was too low to obtain a Q1/ boundary.

Figure 1 shows the boundaries obtained in all seven ways for stimuli based on *saate* (without *ja* and *ka*) in which the durations of the segments preceding and following the VC-sequence all were unchanged. The difference in slope between the Q1/ and /Q3 boundaries for [a] as well as for [t] can be understood as due to the non-use of V:C:: and V::C: in the system.

The boundary values obtained in some of the diagonal way series in which both V and C were increasing in duration are listed in Table 2. Figure 2 shows the variation of the quantity boundaries obtained in the diagonal series as a function of the duration of the vowel [e] of the second syllable.

The data presented in Table 3 reflect the result of a multidimensional regression analysis performed separately for each type of context (*ja s-e, s-e*, and *s-e ka*) with each of the two quantity boundaries as the dependent variables. The result of an overall linear regression analysis is shown in Table 4.
Table 2. Quantity boundaries between Q1 and Q>1 and between Q3 and Q<3 of [a] and [t] obtained for the series of utterances based on "saate" (no ja- or -ka) in which the duration of [a] and [t] had been varied together. Segment durations in ms. Upright: given durations; in italics: obtained boundary values. Rate structure: ">" increased, "=" unmodified, "<" decreased speech rate of the segments before or after the [at]-sequence.

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Rate structure</th>
<th>[ja]</th>
<th>[s]</th>
<th>[a:]</th>
<th>[t]</th>
<th>[t:]</th>
<th>[e]</th>
<th>[ka]</th>
</tr>
</thead>
<tbody>
<tr>
<td>saate</td>
<td>= x =</td>
<td>121</td>
<td>128</td>
<td>193</td>
<td>114</td>
<td>207</td>
<td>201</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>&gt; x &gt;</td>
<td>102</td>
<td>114</td>
<td>172</td>
<td>*</td>
<td>184</td>
<td>169</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>&lt; x &lt;</td>
<td>144</td>
<td>152</td>
<td>225</td>
<td>131</td>
<td>242</td>
<td>239</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>&gt; x =</td>
<td>102</td>
<td>130</td>
<td>195</td>
<td>110</td>
<td>209</td>
<td>201</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>&lt; x =</td>
<td>144</td>
<td>135</td>
<td>210</td>
<td>120</td>
<td>226</td>
<td>201</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>= x &gt;</td>
<td>121</td>
<td>115</td>
<td>175</td>
<td>*</td>
<td>188</td>
<td>169</td>
<td>*</td>
</tr>
<tr>
<td>saate</td>
<td>= x &lt;</td>
<td>121</td>
<td>146</td>
<td>220</td>
<td>131</td>
<td>236</td>
<td>239</td>
<td>*</td>
</tr>
</tbody>
</table>

Figure 2: Quantity boundaries obtained for [a] as a function of the duration of the [e] in the diagonal series affected by final lengthening of the [e] (based on [sa:t:e] and [ja sa:t:e]) and unaffected (based on [sa:t:e ka]).

Discussion
In Table 3 it can be seen that all significant effects of changes in the duration of any non-internal segments are opposite to that of the same change in the duration of the internal segment in focus. However, the vowel of the second syllable has in each case a stronger effect than the initial consonant of the first syllable. At the /Q3 boundary of [a], a given increase (in ms) has even a larger effect (108%) when it occurs on the vowel [e] of the second syllable. In calculating this percentage, it was assumed that the effect of a change in duration of the [e] alone would be equal to the effect actually observed with the change distributed uniformly across [e ka]. Since we can not assure this, the calculated value must be seen as an upper limit. Even so, the value is lower than required in order for the perceived quantity of the [a] to be given by the duration ratio [a]/[e]. This is due to the shorter length of [e] in [e ka] as compared with the /Q3 boundary length of the vowel of the preceding syllable. There is an analogous discrepancy also at the Q1/ boundary and in the perceived quantity of the [t]. This confirms the results obtained in our previous investigation.
Our observations are similar to those made by van Dommelen (1999) on the effects of the duration of a vowel in the post-stressed syllable on the perception of quantity in the stressed syllable of Norwegian words. While van Dommelen (1999) showed speech specific processes to be at work, it appears not to be required to invoke any language-specific processes for Estonian.

Acknowledgements

This research has been supported, in part, by grant F0558/1998 from HSFR. Our special thanks are due to Einar Meister for his help with the arrangement of the listening sessions at the Institute of Cybernetics, Tallinn Technical University.

References


