Quantity and duration in early speech: preliminary observations on three Swedish children

Olle Engstrand and Lina Bonsdroff
Department of Linguistics, Stockholm University

Abstract
The paper presents preliminary data relevant to Swedish children’s acquisition of the quantity contrast. Since Swedish children are heavily exposed to lexically conditioned quantity patterns, and since the contrast involves some fairly robust phonetic correlates, distinct ambient language influences on early vocalizations can be expected. To date, however, such effects have been observed only informally. In the present study, durational measurements were made on the relevant vowels and consonants in disyllabic words produced by three Swedish children aged 30 (two children) and 24 months (one child), respectively. The data suggested adult-like durational patterns in all three children irrespective of age. So far, then, the experimental evidence has encouraged the hypothesis that the quantity contrast may be in place by 24 months of age or even earlier.

1. Introduction
Quantity appears in many languages with medium-sized on large vowel inventories (Maddieson 1984, p. 129). The domain of quantity differs between languages, the primary target being single vowels, consonants or whole rhymes. In many languages, e.g., Finno-Ugric languages such as Finnish, quantity is contrastive in both vowels and consonants and appears independently of lexical stress.

In Standard Swedish, the contrast displays a complementary pattern in the rhyme of stressed syllables with either V(C) or VC(C); cf. the spectrograms in figure 1 which represent the words bada [‘bad[a] ‘to bathe’ and badda [‘bad[a] ‘to sponge’ as produced by an adult male Swedish speaker. The contrast is mostly lexically conditioned even though phonologically or morphologically conditioned regularities can also be found. For example, /j/ and /y/ are always preceded by short vowels (e.g., haj [‘haj] ‘shark’ and lång [‘lån] ‘long’) whereas vowels preceding /v/ are normally long (e.g., liv [‘liv] ‘life’). Also, morpheme-final clusters are usually preceded by short vowels (e.g., burk [‘børk] ‘can’). Thus, forms such as *[ha:v], *[lo:n], *[li:v] and *[bur:k] are counterintuitive to native Swedish speakers. Even though duration is a prime quantity correlate, these words also exemplify that qualitative differences between long and short vowels may be quite salient, more so in some vowel pairs than in others (e.g., Hadding-Koch & Abramson 1964); for example, the difference between (long) [a] and (short) [a] qualities is visible in the first two formants in the spectrograms in figure 1. However, the relationship between durational and qualitative quantity correlates is a matter of considerable regional variation (Strangert & Wretling 2003).

Since every stressed syllable carries quantity, and since the contrast involves fairly robust, multiple correlates, a heavy ambient language impact on Swedish children’s early vocalizations can be expected in this regard. According to recent findings, on the other hand, an equally pervasive prosodic feature, the tonal word accent contrast (Bruce 1977, Engstrand 1995, 1997) does not seem to be under productive control even by most 18-monthers (Engstrand et al. 1991, Engstrand et al. 2003). It can be speculated that this reflects a difference in relation to lexicon and grammar; whereas the quantity contrast is predominantly lexically conditioned, as noted above, the word accent distinction is largely predictable from morphological structure. In
consequence, quantity carries a greater semantic load than do the word accents; this, in turn, might necessitate a focus on quantity at early stages of word production.

Informal observations suggest adult-like durational patterns in 30- and 24-month-olds’, perhaps even in 18-month-olds’ disyllabic word productions. However, this possible ambient language effect has not been empirically verified to date. The present paper offers some preliminary durational evidence from two ages: 30 and 24 months. More specifically, we will present evidence to suggest that, at those ages, children produce words that are consistent with the adult norm in terms of durational implementation of quantity.

2. Methods

The subjects for this study were drawn from a larger database described in Engstrand et al. (2003). In summary, audio and video recordings were made with children at 6, 12, 18, 24 and 30 months of age, with approximately 8 Swedish and 8 American English children in each age group. The children were recorded individually at two separate, 30-45 minutes long sessions separated by an interval of approximately one week; the material used in this study comes from the first of these recording sessions. Each child was accompanied by a parent, and age-appropriate objects and pictures were used to assist parents in eliciting word types of interest. The parents were representative of the regional standard spoken in the Stockholm and Seattle areas, respectively.

For the purpose of this preliminary study, disyllabic words produced by three Swedish children, two 30-monthers and one 24-monther, were sampled (16 kHz sampling rate) and analyzed with respect to the durational aspect of the quantity contrast, i.e., the duration of vowels and consonants pertaining to the rhyme of the word’s stressed syllable. Selected utterances were to be non-cry, non-screaming, non-whispered, and free from environmental noise. Apart from these conditions, utterances were selected semi-randomly such that a relatively even distribution of consonant types was aimed at. In some cases, the video sequence or the parent’s response to the child’s vocalizations were used to interpret word meanings.

Segmentation of vowels and consonants were made according to conventional acoustic landmarks as exemplified in figure 2, and V:C ratios were calculated based on the measured durations. Note that the durational pattern in figure 2 is similar to that of the adult production of the same word, shown to the left in figure 1 above.

![Figure 2](image-url)

Figure 2. Spectrogram of the word bada [ˈbud̥a] ‘to bathe’ as produced by the male Swedish 24-month-old observed in this study. The durations of the vowel and stop closure segments are approximately 225 and 100 ms, respectively.

3. Results

The vowel-to-consonant durational ratios pertaining to the respective children are summarized in table 1. The V:C data are shown under the heading ‘long vowels’ in the left part of the table, and the VC: data (‘short vowels’) are shown to the right. It can be seen, for example, that the 30 months old boy displays V:C and VC: ratios of 2.90 and 0.83, respectively, that the respective standard deviations are 1.67 and 0.38, and that these figures are based on 12 and 20 word tokens, respectively. This means that, on average, this subject’s long vowels were almost three times as long as the accompanying consonants, and that the duration of his short vowels were approximately 80 percent of that pertaining to the accompanying consonants. The remaining two subjects display similar durational relationships; it can be noted, in particular, that, in spite of a six months age difference, the 24-monther follows the same general pattern as the two older children. Considering the relatively modest standard deviations, these effects are clearly reliable. It can be concluded, then, that the durational ratios evidenced by these children are basically similar to those
encountered in adults’ phonetic implementation of the quantity contrast.

Table 1. Mean vowel-to-consonant ratios for words produced by three Swedish children (boy, 30 months, girl, 30 months and boy 24 months).

<table>
<thead>
<tr>
<th></th>
<th>Long vowels</th>
<th></th>
<th>Short vowels</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ratio N</td>
<td>S.D.</td>
<td>Ratio N</td>
<td>S.D.</td>
</tr>
<tr>
<td>Boy 30</td>
<td>2.90 12</td>
<td>1.67</td>
<td>0.83 20</td>
<td>0.38</td>
</tr>
<tr>
<td>Girl, 30</td>
<td>1.83 17</td>
<td>1.17</td>
<td>0.73 16</td>
<td>0.17</td>
</tr>
<tr>
<td>Boy, 24</td>
<td>2.11 9</td>
<td>0.59</td>
<td>0.88 22</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Figure 3. Consonant by vowel durations for the three children. Filled circles: long vowels, unfilled circles: short vowels.

The ratios shown in table 1 are, however, uninformative as to the precise location of the main durational effect. Figure 3 depicts consonant against vowel durations for the three children. The filled circles refer to the V:C structures, and the unfilled circles refer to the VC: structures. In spite of a certain amount of overlap, the respective data sets are relatively well separated. In general, long vowels tend to entail short consonants and vice versa; thus, there is a clear tendency to the expected durational complementarity. Also, however, vowel and consonant variation tend to go together suggesting a tendency to linear increase of syllable duration as a whole. The 30 months old girl shows a notable exception to this pattern; the almost horizontal regression line indicates a variation in long vowel durations that leaves the corresponding consonant durations nearly intact.

4. Conclusions

In summary, these children have evidenced durational ratios that are basically similar to those encountered in adults’ phonetic implementation of the quantity contrast. Also, there are clear traces of the expected durational complementarity. On the other hand, the data contain a considerable amount of variability as expected from vocalizations produced by young children. It should be noted, however, that phonetic variability is a hallmark of ‘spontaneous’ or ‘unscripted’ speech in general, not only of that produced by young children.

Do Swedish children in the 24-30 months age range control the quantity contrast according to the adult norm? Clearly, the three subjects studied here are too few to provide the conclusive evidence. For another thing, the Swedish children’s vocalizations will have to be compared to those produced by non-Swedish controls. As far as the data go, however, it is reasonable to maintain the hypothesis that quantity is mastered by children in the 24-30 months age interval and perhaps earlier. Further work should aim to analyze vocalizations produced by additional children in this age range and, subsequently, to subject younger children’s vocalizations to similar
analyses. In particular, a long term goal should be to integrate durational data relevant to the (lexically conditioned) quantity contrast with F0 data associated with the (morphologically conditioned) tonal word accent distinction (cf. Engstrand & Kadin 2004). This will hopefully help to clarify the role of linguistic constraints on ambient phonology acquisition.

Acknowledgment

This work was supported in part by a grant from Vetenskapsrådet (the Swedish Research Council), VR 421-2003-1757.

References


