Prosodic phrasing and articulation rate variation

Hansson, P.

journal: Proceedings of Fonetik, TMH-QPSR
volume: 44
number: 1
year: 2002
pages: 173-176

http://www.speech.kth.se/qpsr
Prosodic phrasing and articulation rate variation

Petra Hansson
Department of Linguistics and Phonetics
Lund University

Abstract
In this paper, results from two studies on articulation rate variation and prosodic phrasing are presented. Production data are presented that suggest a progressive articulation rate reduction over the course of the prosodic phrase in southern Swedish, and results from a perception experiment reveal the importance of articulation rate reduction for perceived boundary strength.

Introduction
Final lengthening is clearly visible in Swedish production data, and has also convincingly been shown to be an important cue in the perception of phrasing in standard Swedish (Bruce et al. 1992). However, the phonological and phonetic conventions of a standard variety are not always applicable to all varieties of a language, and there are several known differences in accent realization and accent distribution between the dialects of Sweden (e.g. between southern Swedish and the much more studied so-called standard variety of Swedish) that may have implications for the phrasing strategies used. The aim of the present paper is to shed some light on prosodic phrasing in southern Swedish.

Southern Swedish share many prosodic properties with Danish, a language that has been claimed to lack phrase-final lengthening in some dialects (Gronnum Thorsen 1988). Like Danish, southern Swedish has no (high) phrase accent, the high turning point that in standard Swedish follows the word accent fall in focal and phrase-final positions. Lyberg (1981) has suggested that the phrase-final lengthening phenomenon in Swedish is related to this characteristic phrase end contour of the fundamental frequency, the rise-fall gesture after the last word accent. In southern Swedish, there is no such gesture after the last word accent and there would therefore not be a need for slowing down the articulation rate phrase-finally. Secondly, like in Danish, there is no so-called default sentence accent (focal accent) at the end of the phrase in southern Swedish. The most prominent accent is not always found on the last word of the phrase. Even assuming that focal accentuation in itself (regardless of the tonal gesture associated with it) results in final lengthening, there is no reason to believe that final lengthening is obligatory in the south Swedish phrase.

The research questions addressed in the present paper concern the articulation rate variation within the prosodic phrase in southern Swedish. Can a reduction of the articulation rate be observed over the course of the prosodic phrase in southern Swedish, and if so, has the variation in articulation rate any effect on the perceived boundary strength of prosodic phrase boundaries in southern Swedish?

Methods and materials

The production data study
The speech material used is from the SweDia 2000 database (Bruce et al. 1999). The speech of four female speakers (two from the younger generation of speakers recorded, and two from the older generation) and two male speakers (from the older generation) from four of the recording locations in Skåne has been analyzed.

The first 100 prosodic phrases in the six speakers’ spontaneous recordings were initially chosen for the analysis. 447 of these were phrases without phrase internal pauses or fillers. Disfluent phrases were excluded since they may contain segmental lengthening associated with another domain than the prosodic phrase (see Dankovičová 1997).

For the analysis of the articulation rate variation within the prosodic phrase, we followed Dankovičová 1997 and measured and compared the articulation rate (in syllables per second) in each prosodic word in the prosodic phrase. Swedish is not a language with a fixed stress position, but the by far most common
stress pattern contains an initial stress. Therefore, the stressed syllables served as landmarks in the segmentation of the phrases into words. Only in phrase-initial position were unaccented syllables attached to the following accented syllable instead of the preceding.

In counting the number of syllables in each prosodic word, care was taken to count the actual number of syllables, rather than the numbers of syllables the word contains in its citation form.

**The perception experiment**

The purpose of the perception experiment was to relate three known cues for prosodic phrasing (pausing, F0 reset and reduced articulation rate in the final part of the phrase) to perceived boundary strength. For this purpose, another 50 short speech fragments were chosen from the SweDia 2000 recordings from Skåne (spontaneous speech from five female and five male speakers). All speech fragments (typically one or two utterances long) contain at least one prosodic phrase boundary. In the test, the listeners were presented with an orthographic transcription of the utterance(s) in which the prosodic boundary of interest was marked with a ‘/’. Their task was to indicate how strongly marked they perceived the indicated boundary to be in the recording.

They indicated the boundaries’ strength on a computer-based Visual Analogue Scale (VAS). The VAS has been used in the measurement of clinical phenomena (such as pain) since the 1920’s (Wewers and Lowe 1990), but it is suitable for measuring a variety of subjective non-clinical phenomena as well. The subjects respond by placing a mark at a position that represents their current perception of a given phenomenon (in this case a given boundary’s strength) between the labeled extremes of the 100 mm scale, see Figure 1.

The VAS is then scored by measuring the distance (in mm) from one end of the scale to the subject’s mark on the line.

So far, only ten subjects have completed the perception test, and therefore the results presented below have to be considered preliminary.

**Results**

**Articulation rate reduction in production**

There were 323 prosodic phrases containing more than one prosodic word in the material (134 2-word phrases, 118 3-word phrases, 50 4-word phrases, 18 5-word phrases and 3 6-word phrases). Due to the low number of 5- and 6-word phrases in the material, mainly those phrases containing 2 to 4 words have been analyzed.

As a first step in the analysis of the data, we chose to rank order the words in the phrases according to their articulation rate. As shown in Figure 2, 113 (84%) of the 2-word phrases show an AB pattern, i.e. a reduction of the articulation rate where the first word (A) is articulated with a higher articulation rate than the second and final word (B). All speakers but one use the AB pattern more frequently than the BA pattern. Speaker 3 uses the BA pattern as often as the AB pattern.

As shown in Figure 3, 59 (44%) of the 3-word phrases demonstrate an ABC pattern, i.e. a progressive slowing down. Another 36 phrases (27%) show a reduction of the articulation rate that is observable only in the comparison of the articulation rates in the second and third word, i.e. a BAC or CAB pattern. No phrase demonstrates a CBA pattern, i.e. a pattern in which the articulation rate is gradually increasing.
Figure 3. Ordinal patterns in 3-word phrases.

The most common pattern in the 4-word phrases is the ABCD pattern (13 of 50), i.e. a gradual decrease in articulation rate over the course of the phrase. The ACBD, BACD and CBAD patterns are also relatively common (16 of 50). In the BACD the reduction of the articulation rate is visible in the last three words, and in the ACBD and CBAD patterns in the two last words. It is also clear that in the majority of the 4-word phrases, it is either the first or the second word in the phrase that is pronounced the fastest (in 38 cases of 50), and that the phrase-final word is either the word with the lowest or second lowest articulation rate in the phrase (in 41 cases of 50).

The reduction of the articulation rate over the course of the phrase is also observable in the mean articulation rates in the material, as shown in Tables 1 and 2.

Table 1. Mean articulation rate (in syllables/second) and standard deviations for the words in 2- and 3-word phrases (all speakers).

<table>
<thead>
<tr>
<th></th>
<th>2-word phrases</th>
<th>3-word phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st word</td>
<td>Mean 6.8</td>
<td>Mean 6.8</td>
</tr>
<tr>
<td></td>
<td>SD 2.1</td>
<td>SD 1.8</td>
</tr>
<tr>
<td>2nd word</td>
<td>Mean 4.4</td>
<td>Mean 6.0</td>
</tr>
<tr>
<td></td>
<td>SD 2.0</td>
<td>SD 2.2</td>
</tr>
<tr>
<td>3rd word</td>
<td>Mean 4.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD 1.4</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Mean articulation rate (in syllables/second) and standard deviations for the words in 4- and 5-word phrases (all speakers).

<table>
<thead>
<tr>
<th></th>
<th>4-word phrases</th>
<th>5-word phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st word</td>
<td>Mean 7.0</td>
<td>Mean 7.2</td>
</tr>
<tr>
<td></td>
<td>SD 2.4</td>
<td>SD 1.7</td>
</tr>
<tr>
<td>2nd word</td>
<td>Mean 6.1</td>
<td>Mean 6.3</td>
</tr>
<tr>
<td></td>
<td>SD 1.8</td>
<td>SD 1.7</td>
</tr>
<tr>
<td>3rd word</td>
<td>Mean 6.0</td>
<td>Mean 6.2</td>
</tr>
<tr>
<td></td>
<td>SD 3.0</td>
<td>SD 1.8</td>
</tr>
<tr>
<td>4th word</td>
<td>Mean 3.7</td>
<td>Mean 5.5</td>
</tr>
<tr>
<td></td>
<td>SD 1.1</td>
<td>SD 1.5</td>
</tr>
<tr>
<td>5th word</td>
<td>Mean 4.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD 1.7</td>
<td></td>
</tr>
</tbody>
</table>

Typically, the articulation rate in the phrase-initial words is about 7 syllables per second and in the final word about 4 to 5 syllables per second.

In order to test if a word’s position in the phrase has an effect on its articulation rate, a GLM (general linear modeling) procedure was used. The dependent variable was articulation rate, and the two factors were position (word’s position in phrase) and speaker. Each phrase type (the 2-, 3- and 4-word phrase) was analyzed separately.

The factor of position was significant for all phrases: 2-word phrases (F(1, 256)=87.6, p<.001), 3-word phrases (F(2, 336)=67.0, p<.001) and 4-word phrases (F(3, 176)=22.4, p<.001). A posthoc Tukey revealed that the mean articulation rate of all three words were significantly different from each other at the .01 level in the 3-word phrases. In the 4-word phrases, on the other hand, significant differences in mean articulation rate were found only in the comparisons of the mean articulation rates in words 1 and 4, 2 and 4, and 3 and 4 (p < .01). In other words, the only significant difference in articulation rate between two successive words was found in the comparison between the final and penultimate word.

Typically, the articulation rate in the phrase-initial words is about 7 syllables per second and in the final word about 4 to 5 syllables per second.

Articulation rate reduction in perception

Two thirds of the prosodic phrase boundaries in the perception experiment (33 of 50) were associated with pauses (a silent interval) ranging in duration from 130 ms to 2860 ms. The F0 resets across the phrase boundaries ranged from a 190 Hz large reset of F0 to a lowering of F0 across the phrase boundary of 51 Hz. If expressed in percentages of the speakers’ F0 ranges (measured as the difference in F0 between the highest and the lowest F0 value in the recording), the F0 resets ranged from a reset of 94% to a lowering of 76%. The normalized difference in articulation rate (the difference was divided by the mean articulation rate in the two words being compared) ranged from 1.3 (corresponding to a change in articulation rate
from 7.7 syllables/second to 1.8 syllables/second) to –0.3 (corresponding to a change from 4.2 syllables/second to 5.4 syllables/second).

None of the three cues investigated are related to each other. There is no statistically significant correlation between the duration of the pause, the size of the F0 reset and/or the change in articulation rate. Apparently speakers do not maximize the use of several cues simultaneously to increase the strength of a boundary. This finding is partly consistent with the trading relation between pause duration and final lengthening reported on in Horne, Strangert and Heldner 1995.

Statistical analyses revealed that the listeners in the perception experiment agree very well in their perceptual judgments. The Pearson correlation coefficient of each pairwise combination of listeners was significant at the .001 level. Therefore, we have chosen to follow Sanderman (1996) and pool the scores of the listeners for each boundary and calculate a mean to obtain an estimate of each boundary’s perceived strength (in mm). Possible relationships between the perceived boundary strength and the cues for prosodic phrasing can then be investigated.

The correlation between pause duration and perceived boundary strength proved very strong \( r = .90, r^2 = .81 \). As shown in Figure 4, the longer the pause is, the stronger the listeners perceived the boundary. The only clear exceptions are pauses longer than 2 seconds, which were not perceived stronger than pauses of 1.5 to 2 seconds.

![Figure 4. Scatterplot demonstrating the relationship between pause duration and perceived boundary strength (PBS).](image)

Attempts made to relate the pauses’ duration to the speakers’ speaking rate did not strengthen the correlation further. Finally, no statistically significant correlation was found between change in articulation rate and perceived boundary strength, or between F0 reset and perceived boundary strength.

**Conclusions**

The results presented in this paper show that the articulation rate in south Swedish phrases is significantly lower in phrase-final words than in preceding words, i.e. that some sort of phrase-final lengthening exists in southern Swedish. The results also indicate that the reduction in articulation rate between successive words in the prosodic phrase is not restricted to the final part of the phrase. However, preliminary results from a perception experiment suggest that the amount of articulation rate reduction has no effect on the perceived strength of prosodic phrase boundaries.

**References**


