Articulation Rate Variation in South Swedish Phrases

Petra Hansson

Department of Linguistics and Phonetics
Lund University, Sweden
petra.hansson@ling.lu.se

Abstract

In this paper, a study on articulation rate variation within the prosodic phrase is presented. The dialect under investigation is south Swedish. South Swedish shares many prosodic properties with Danish, a language claimed to lack phrase-final lengthening. The results of the present study nevertheless indicate that the articulation rate in south Swedish phrases is significantly lower in phrase-final words than in preceding words, and that the reduction in articulation rate between successive words in the prosodic phrase is not restricted to the final part of the phrase.

1. Introduction

1.1. Cues for prosodic phrasing in Swedish

Why is it interesting to examine whether changes in articulation rate within the prosodic phrase, such as final lengthening (a slowing down of the articulation rate in the last syllable of the phrase), is a cue for phrasing in southern Swedish? Final lengthening is clearly visible in Swedish production data [1], and has also convincingly been shown to be an important cue in the perception of phrasing in Swedish [2]. However, in Swedish prosody research, although a great deal of attention has been paid to dialectal variation in word accent realization (see e.g. [3]), possible differences as regards phrasing between the dialects have been almost completely overlooked (see [4] however). As e.g. shown in [5], the phonological and phonetic conventions of a standard variety are not always applicable to all varieties of a language, and there are several differences in the accent realization and accent distribution between the dialects that may have implications for the phrasing strategies used.

1.2. Final lengthening and F0 dependence

There are several known differences between southern Swedish and the much more studied so-called standard variety of Swedish that may have implications for the cues used for prosodic phrasing in the two dialects. For example, southern Swedish share many prosodic properties with Danish, a language that has been claimed to lack phrase-final lengthening in some dialects [6].

Like Danish, southern Swedish also lacks a (high) phrase accent, the high turning point that in standard Swedish follows the word accent fall in focal and phrase-final positions. Lyberg [7] 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 fall. In southern Swedish, however, there is no such gesture after the last word accent and there would therefore not be a need for prolonging the phrase-final durations or slowing down the articulation rate phrase-finally. Furthermore, like in Danish, there is no so-called default sentence accent (focal accent) at the end of the phrase in southern Swedish [6]. In other words, 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.

Despite the fact that we do not expect to find F0-dependent final lengthening in southern Swedish, it is interesting to look for non-tonal cues for prosodic phrasing in this dialect. Since the prosodic phrase in southern Swedish is not ended with a prominent rise-fall gesture (after the last word accent), but with an accent that does not deviate from preceding ones neither regarding direction nor necessarily as to the extent of its F0 movement (much like Copenhagen Danish), non-tonal cues may prove to be very important.

A method for measuring final lengthening is developed and discussed in [8]. Since the inherent duration of a phone is known to be the largest source of variation in segmental duration, in measuring final lengthening one would like to measure the difference between the duration of a given segment and the mean duration of that specific phone type. Wrightman et al. [8] therefore introduce *normalized duration*, which measures the duration of a segment as the number of standard deviations from the mean duration of the phone contained in the segment. In order to obtain the means and standard deviations of a speaker's phones, a fairly large amount of speech data needs to be segmented and labeled. This can be done automatically, as it has been for standard Swedish (see [9]), but since we have no speech recognizer available to us for segmenting and labeling our south Swedish data, we will choose another method in the present investigation.

The method used in the present study will not allow us to pinpoint the exact domain of possible final lengthening. Nevertheless, it will allow us to answer the question of whether or not there is a lower articulation rate in phrase-final words, and if so, whether the reduction in articulation rate begins before the last prosodic word of the phrase or not.

1.3. Articulation rate variation

In an experiment conducted to determine if articulation rate variation has a specific domain in spontaneous speech (Czech), Dankovičová [10] found a regular pattern within the intonation phrase. By measuring and comparing the articulation rate in each phonological word in the phrase, she was able to show that the articulation rate was slowing down over the course of the phrase. The first or second word was demonstrated to have the highest articulation rate and the last word to have the slowest. The so-called interpause stretch (stretch of speech bounded by consecutive pauses) was also found to be a domain of phrase-final lengthening.

1.4. Research question

The research question to be answered in the present study concerns the existence of articulation rate variation within the prosodic phrase in southern Swedish. Is there variation in articulation rate within the phrase, and if so, what is the nature of this variation? Two hypotheses will be tested: 1) that the position of a word has an effect on the articulation rate, and 2) that the articulation pattern demonstrates a progressive slowing down.

2. Method

2.1. Speech material

The speech material used in the present study has been extracted from the SweDia 2000 database [11]. Within SweDia 2000 (Phonetics and Phonology of the Swedish Dialects around Year 2000) speech samples from over 100 Swedish dialects have been collected. Each dialect is represented by at least 12 speakers, and both spontaneous speech and words and phrases elicited with a number of specific research questions in mind have been recorded. In the present study, speech extracted from the spontaneous part of the database has been used. 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 five recording locations in Scania (the southernmost region of Sweden) has so far been analyzed.

The recordings were made with a sampling rate of 44.1 kHz and 16 bit resolution, and transferred digitally to a Sun workstation and stored as ESPS/Waves+™ files. The sampling frequency has subsequently been converted to 16 kHz.

2.2. Measurements and criteria for segmentation

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 [10], the interpause stretch). The positions of prosodic phrase boundaries were marked after listening and visually inspecting F0 traces of the speech material in the speech analysis program ESPS/Waves+™.

For the analysis of the articulation rate variation within the prosodic phrase, the articulation rate (measured in syllables per second) was calculated for each prosodic word, the defining feature of the prosodic word in Swedish being that it contains one accent. 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 order to decide the exact positions of the prosodic word boundaries (the onsets of the stressed syllables), the Maximal Onset Principle (which puts consonants preferentially into onsets rather than codas) was applied. Figure 1 shows an example of how the prosodic phrases where segmented. The phrase is *och haft hela tiden* 'and have had all along' (male speaker).

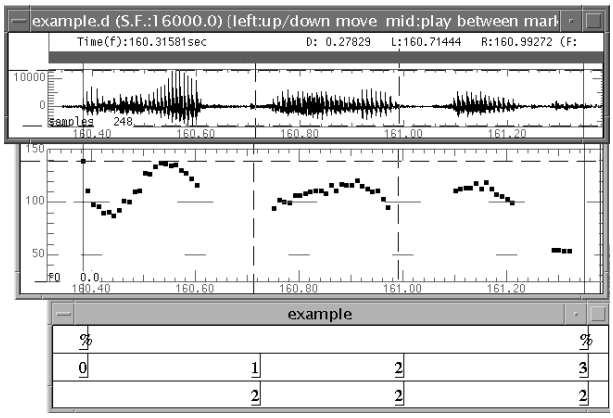

Figure 1: Waveform and F0-contour of a prosodic phrase. The three label tiers are, from top to bottom: 1) a phrase tier with the beginning and the end of the prosodic phrase marked with %, 2) a word tier with labels for the beginning (0) of the initial word as well as the end of the initial (1) and possible following words, and 3) a syllable tier with labels for the number of syllables within the words in the word tier.

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.

3. Results

There were 323 prosodic phrases containing more than one prosodic word in the material. The distribution of 2-, 3-, 4-, 5- and 6-word phrases for each speaker separately and all speakers pooled together is shown in Table 1. Due to the low number of 5- and 6-word phrases in the material, only those phrases containing 2 to 4 words have been analyzed.

Table 1: Number of prosodic phrases containing two to six prosodic words

Speaker	2-word phrases	3-word phrases	4-word phrases	5-word phrases	6-word phrases
1	20	21	8	5	1
2	17	21	8	2	1
3	16	13	10	5	0
4	18	20	9	5	0
5	39	20	7	1	0
6	24	23	8	0	1
Total:	134	118	50	18	3

The idea proposed in [10], is that the articulation rate variation within the prosodic phrase follows a certain pattern, a slowing down of the articulation rate. In order to test whether this is also the case in southern Swedish, we will begin by rank ordering the words in the phrases according to their articulation rate.

As shown in Table 2, 113 (84%) of all 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.

Table 2: Ordinal patterns in 2-word phrases

Speaker	AB	BA
1	19	1
2	14	3
3	8	8
4	16	2
5	32	7
6	24	0
Total:	113	21

As shown in Table 3, 59 (44%) of the 3-word phrases demonstrate an ABC pattern, i.e. a progressive slowing down of the articulation rate where the first word (A) is pronounced faster than the second word (B), which in its turn is pronounced faster than the third and final word (C). 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.

Table 3: Ordinal patterns in 3-word phrases

Speaker	ABC	ACB	BAC	BCA	CAB	CBA
1	10	5	6	0	0	0
2	10	6	5	0	0	0
3	5	4	3	1	0	0
4	13	2	4	0	1	0
5	10	2	7	0	1	0
6	11	2	7	1	2	0
Total:	59	21	32	2	4	0

As shown in Table 4, the most common pattern in the 4-word phrases is the ABCD pattern, i.e. a gradual decrease in articulation rate over the course of the phrase. The ACBD, BACD and CBAD patterns are also relatively common. 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).

Table 4: Ordinal patterns in 4-word phrases (only patterns existing in the data included)

Pattern	Number of phrases
ABCD	13
ABDC	2
ACBD	6
ACDB	1
ADBC	2
ADCB	1
BACD	5
BADC	2
BCAD	4
BCDA	1
BDCA	1
CABD	2
CADB	4
CBAD	5
CDBA	1

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 5 and 6. 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. The largest differences in mean articulation rate between successive words are found in the comparisons between the articulation rate in the phrase-final and penultimate words.

Table 5: Mean articulation rate (in syllables/second) and standard deviations for the words in 2- and 3-word phrases (all speakers).

	2-word phrases		3-word phrases	
	Mean	SD	Mean	SD
1st word	6.8	2.1	6.8	1.8
2nd word	4.4	2.0	6.0	2.2
3rd word			4.2	1.4

Table 6: Mean articulation rate (in syllables/second) and standard deviations for the words in 4- and 5-word phrases (all speakers).

	4-word phrases		5-word phrases	
	Mean	SD	Mean	SD
1st word	7.0	2.4	7.2	1.7
2nd word	6.1	1.8	6.3	1.7
3rd word	6.0	3.0	6.2	1.8
4th word	3.7	1.1	5.5	1.5
5th word			4.6	1.7

In order to test the hypothesis that 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 ($F(3, 176) = 22.4$, $p < .001$). No analyses were made of the 5- and 6-word phrases because of the low number of occurrences. In order to analyze the pattern of articulation rate variation, i.e. if a pattern of progressive slowing down exists, a posthoc Tukey test was done on the means. In the 3-word phrases the mean articulation rate of all three words were significantly different from each other at the .01 level. 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.

The factor of speaker was significant at the .01 level in the 3-word phrases ($F(5, 336) = 5.5$, $p < .001$), although the only speaker that demonstrated a significantly different mean articulation rate was speaker 3 (significantly different from all other speakers'). The two-way interaction position by speaker was significant at the .01 level only in the 2-word phrases

($F(5, 256) = 6.1$, $p < .001$), indicating that the speakers were not, in general, different in their patterns of articulation rate. This finding is consistent with the results of the rank ordering of the words according to their articulation rate.

4. Concluding remarks and future work

Despite the fact that we did not expect to find any F0-dependent final lengthening in southern Swedish, we have argued that non-tonal cues for prosodic phrasing (such as articulation rate variation) nevertheless are likely to be important in this particular dialect. We believe this to be so because the prosodic phrase in southern Swedish is not ended with a prominent extra rise-fall gesture as it is in standard Swedish, but with an accent that does not deviate from preceding ones neither as regards direction nor necessarily as to the extent of its F0 movement.

The analysis of speech from six south Swedish speakers has revealed a significant effect of word position on articulation rate. Like Dankovičová [10], we have found some evidence to suggest a progressive reduction of the articulation rate over the phrase. However, regardless of the length of the phrase (as expressed in the number of words it contains), the articulation rate in the first and the last word is roughly the same. The difference in mean articulation rate between the words in the phrase is therefore small in the 4- and 5-word phrases, and the reduction in articulation rate between all successive words in the phrase is only significant in the 2- and 3-word phrases.

The largest differences in mean articulation rate were found in the comparison of the articulation rate in phrase-final and penultimate words, indicating that some sort of final lengthening exists in southern Swedish. Due to the method used to measure the articulation rate variation, no conclusions can be drawn with regard to its domain (whether or not it is restricted to the last syllable of the phrase-final word).

Future work involves extending the study to include a total of ten speakers (two speakers from each recording location in Scania, five female and five male speakers) and to further investigate individual differences regarding the patterns of articulation rate. A possible factor to include in the analysis of the articulation rate variation, is the factor of word size (expressed in number of syllables). Although it was shown in [10] that size is not a reliable determinant of a word's articulation rate, the fact that the phrase-initial words tend to contain more syllables than non-initial words (as a consequence of the segmentation criteria discussed in section 2.2), it is important to control whether this has had any effect on their measured articulation rates. Finally, the perceptual relevance of the decreasing articulation rate needs to be tested. Is the observed reduction in articulation rate over the phrase used by listeners in their perception of prosodic phrasing?

5. References

[1] Bruce; G.; Granström, B.; Gustafson, K.; House, D. 1991. Prosodic Phrasing in Swedish. *Working Papers* 38. Department of Linguistics and Phonetics, Lund University, 5-17.

[2] Bruce; G.; Granström, B.; Gustafson, K.; House, D. 1992. Interaction of F0 and duration in the perception of prosodic phrasing in Swedish. In *Nordic Prosody VI. Papers from a Symposium, Stockholm, August 12-14 1992*, 7-22.

[3] Bruce, G. 2001. Secondary stress and pitch accent synchronization in Swedish. In *Nordic Prosody. Proceedings of the VIIIth Conference, Trondheim 2000*, Van Dommelen, W. A.; Fretheim, T. (eds). Peter Lang, 33-44.

[4] Gårding, E. 1974. Den efterhängsna prosodin. In *Språket i bruk*, U. Teleman; T. G. Hultman (eds.). Lund: Gleerups, 50-71.

[5] Fletcher, J.; Grabe, E.; Warren, P. To appear. Intonational variation in four dialects of English: the high rising tune. In *Prosodic typology and transcription - a unified approach*, Sun-Ah Jun (ed.). OUP.

[6] Grønnum Thorsen, N. 1988. Intonation on Bornholm – Between Danish and Swedish. *ANNUAL REPORT of the Institute of Phonetics, University of Copenhagen*, 25-138.

[7] Lyberg, B. 1981. Some consequences of a model for segment duration based on F0-dependence. *Journal of Phonetics* 9, 97-103.

[8] Wightman, C. W.; Shattuck-Hufnagel, S.; Ostendorf, M.; Price, P.J. 1992. Segmental Durations in the Vicinity of Prosodic Phrase Boundaries. *Journal of the Acoustical Society of America* 91(3), 1707-1717.

[9] Lindström, A.; Bretan, I.; Ljungqvist, M. 1996. Prosody Generation in Text-to-Speech Conversion Using Dependency Graphs. *ICSLP 96. The Fourth International Conference on Spoken Language Processing*. October 3-6, Philadelphia, USA. Volume 3, 1341-1344.

[10] Dankovičová, J. 1997. The domain of articulation rate variation in Czech. *Journal of Phonetics* 25, 287-312.

[11] Bruce, G.; Elert, C.-C.; Engstrand, O.; Eriksson, A.; Wretling, P. 1999. Database tools for a prosodic analysis of the Swedish dialects. *Proceedings Fonetik 99. The Swedish Phonetics Conference.* June 2-4 1999. Gothenburg Papers in Theoretical Linguistics 81, 37-40.