PROSODIC BOUNDARY STRENGTH IN SWEDISH: FINAL LENGTHENING AND SILENT INTERVAL DURATION

Merle Horne*, Eva Strangert** and Mattias Heldner**
*Dept. of Linguistics & Phonetics, Lund Univ. **Dept. of Phonetics, Umeå Univ.

ABSTRACT

Production data are presented that support the assumption of Final Lengthening and Silent Interval duration as parameters of prosodic boundary strength in Swedish.

INTRODUCTION

Within the hierarchical model of prosodic constituents assumed for Swedish [1], four boundary strengths (0-3) have been assumed: the 0-boundary is thus associated with the end of a word *within* a Prosodic Word (PW), boundary strength 1 with the end of a PW, boundary strength 2 with the end of a Prosodic Phrase (PPh), and boundary strength 3 with the end of a Prosodic Utterance (PU). Although we have a good idea as to how tonal parameters are associated with the various prosodic categories, it is not clear how other parameters, in particular Final Lengthening (FL) and Silent Interval (SI) duration are associated with the various prosodic constituents.

Prosodic constituents have often been assumed to be associated with specific degrees of FL. The domain of lengthening is generally believed to be the rhyme of the final syllable [2-3]. Recent studies have also shown that the lengthening is progressive, i.e. the final consonant is lengthened to a greater extent than the preceding vowel [4]. It is also known that FL is influenced by the presence of prominent accents/boundary tones [3]. Researchers on Swedish [5-6] have provided evidence from production experiments that shows that segment lengthening in final position is indeed influenced by the presence of a focal accent on the last word in an utterance. Lyberg & Ekholm, basing themselves on measurements made on the stressed vowel rather than the final rhyme consonant in their test words [7], do not find any consistent evidence for the appearance of FL as an independent marker of the end of a phrase. Fant and colleagues [8-9], in studies on read prose, show that there is a negative correlation between SI duration and FL in their analysis of stress foot structure in Swedish. These results prompted us to make an investigation in order to tease out the relation between accenting, FL and SI duration in order to relate the findings mentioned above to boundaries assumed in the prosodic constituent hierarchy.

DATABASE STUDY

We made a preliminary data base study to determine 1) whether boundaries between the four types of prosodic constituents we have assumed for Swedish are associated with different degrees of perceived boundary strength and if so, to determine 2) if and how these boundary strengths correlate with segment lengthening and silent intervals for the radio commentator style we are modelling.

17 broadcasts from Radio Sweden on Stock-Market rates were studied. The boundary strength after the Accent 1 word *procent* 'percent' [pru'sent] was chosen for analysis since this word was uttered on the average of 5 times during the broadcasts comprising the database. Also the 5 different occurrences of *procent* always had the same respective syntactic position in each text. The material was presented to 2 native listeners, who scored the strength of the perceived boundary after *procent* on a 4 point scale, where 0 corresponds to no boundary (0-boundary), and 3 corresponds to the strongest boundary (PU-boundary). An example of one of the texts follows (subscripts after the word *procent*: refer to the predicted boundary strength):

Vid 13-tiden noterades Stockholms fondbörs generalindex till 1026,1. Det är en uppgång med 0,1 procent(1) jämfört med gårdagens slutindex. 16-i-topp-index hade då gått upp med 0,4 procent(3).

Marknadsräntorna vid middagstid: den 4-åriga standardobligationen låg då stilla på gårdagens slutränta på 10,12 procent(2), 12-månaders statsskuldväxlar hade gått tillbaka 1 räntepunkt till 10,58 procent(2), medan sexmånadersväxlar gått upp 5 punkter till 10,50 procent(3).

Preliminary results

The boundaries were given scores of 1, 2, or 3 in accordance with the predictions. (The text contained no test word followed by a predicted 0-boundary and there there were no 0-scores either.) A MANOVA analysis demonstrated significant differences in the rhyme of the stressed syllable for the scored boundaries. The greatest differences were found in the [t] segment which differed significantly between all three perceived boundary strengths and was furthermore positively correlated with the strength of the boundary. The SI durations were also positively correlated with the strength of the boundary, and significant differences were found between all three boundary strengths.

Although the results indicated a clear correlation between type of boundary and the phonetic correlates, there were certain gaps in the database: the test word was non-focal in 4 cases and focal in one case. Furthermore, as there were no examples of the test word followed by a 0-boundary, we decided to conduct a more structured lab study in order to include all boundary types.

LAB STUDY

The study was thus undertaken in order to include all boundaries after both [+focus] and [-focus] words to determine to what extent 1) the boundary-type and 2) the focal/non-focal status of a word interacts with FL and SI duration.

One text from the database study was modified so that in new versions *procent* was followed by all 4 types of boundary, 0-boundary, PW, PPh and PU-boundaries, respectively. Moreover, 2 subcategories of boundary within the PPh and PU were distinguished: clause-final/sentence-final position for PPh and paragraph-final/textfinal position for PU. Thus there were 6 boundary categories: 0, PW, PPh/C, PPh/S, PU/P and PU/T. All 6 categories occurred after both [+focus] and [-focus] *percent*. In this way 6 new texts were created which were read 10 times by the same speaker as in the database study. Altogether the material contained 120 occurrences of the testword (10 readings x 6 boundaries x +/- focus).

Results

Figure 1 shows the SI duration for the 5 types of boundary, and Figure 2, the duration of the entire word *procent* before each boundary. The data are given separately for +/-focus (abbreviated +/-fa), thus giving an overview of the general trends. Concerning SI duration, there is a gradual increase as the rank of the boundary becomes higher. The SI varies between 0 (0-boundary) to more than 900 msec (PU/T). However, only three levels of boundary differ significantly from each other on the basis of the measured intervals: 0 and PW versus PPh/C and PPh/S versus PU/P ($p<.05$). The focus distinction, moreover, has no significant effect on the SI duration. The duration of *procent*, on the other hand, differs considerably between the +/-focus condition ($p<.001$). The [+focus] increase in duration varies over the different boundaries between 40 and 100 msec. Figure 2 also demonstrates complex effects of boundary type on word duration, an increase in the higher-rank end of the curve (PPh/C - PU/T and a decrease in the lower-rank end (0 - PPh/C). Data disentangling the complex information in Figure 2 are presented in Figure 3 showing in separate parts the duration of the segments [s], [ɛ], [n], [t].

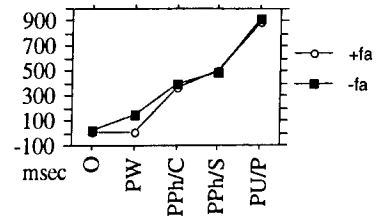

Figure 1. Silent interval duration following the test word with +/- focus accent.

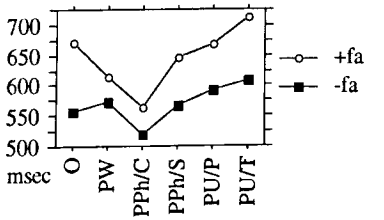

Figure 2. Duration of the test word with +/- focus accent before each boundary.

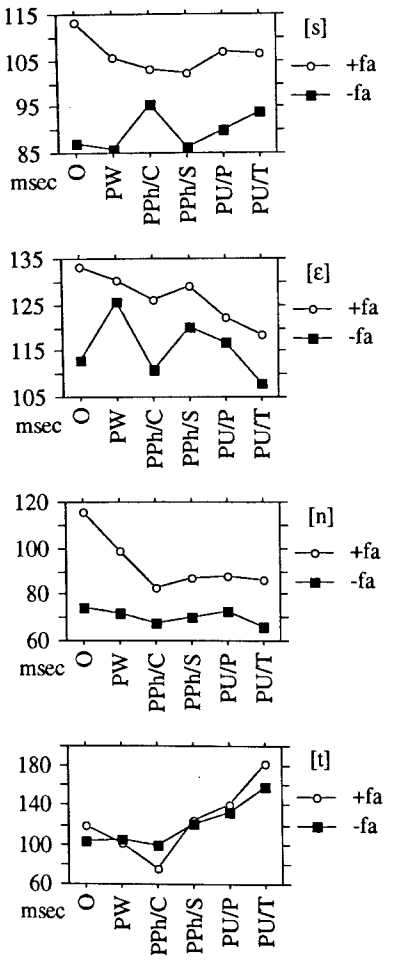

Figure 3. Stressed final syllable segment durations in +/- focus accented test word before each boundary.

First, whether the test word is focussed or not has significant effects on all segments except the final [t] ($p<.001$). (The first syllable of the test word is affected similarly.) Secondly, the final [t] as well as the preceding [ε] and [n], are significantly ($p<.001$) affected by boundary type (as is also the initial [pr]-segment). However, the three segments are affected in very different ways. For [ε] there seems to be a negative correlation between segment duration and the rank of the boundary. For [n] there is also a negative correlation, but only for the boundaries ranked lowest, 0, PW and PPh/C. The higher ranked boundaries appear to be unaffected. For both [ε] and [n] these adjustments primarily affect [+focus] words. For [t] on the other hand, there is a positive correlation between SI duration and the boundaries with higher ranking, PPh/C, PPh/S, PU/P and PU/T. The interaction between boundary type and +/-focus makes it difficult to state the effects of boundary type for 0, PW and PPh/C. Thus, the segment duration data demonstrate where the effects found in Figure 2 come from. The increase of duration in the higher rank end of the curve stems primarily from the [t], while the decrease in the lower rank end is a combined effect of adjustments made in [ε], [n] and, to some extent, [t].

CONCLUSIONS

We may first conclude that the database study indicates that the boundaries associated with the four environments we have analysed can be perceptually distinguished. The subcategorizations of the PPh and PU categories in the present study have not been tested perceptually yet. Thus, we cannot be certain as to how many distinguishable categories there are.

A major adjustment affecting the rhyme segment durations is associated with the focus accent, with [+focus] associated with significantly longer durations than [-focus). This is to be expected, as focus accent has temporal correlates in additon to the primary F0-correlates [6]. However, [t] as well as the SI following the test word do not conform to the general trend, both being unaffected by the +/-focus distinction.

Concerning the temporal adjustments associated with the boundary types investigated we found a more or less gradual increase in the duration of the SI upon an increase in the rank of the boundary observed (see also [8,10] for Swedish). Concerning segment durations, the increase in [t] duration associated with the higher rank end of the boundary scale and the decrease in [ε], [n] and [t] duration in the lower end, together, as we have seen, sum up to a v-shaped curve with PPh/C forming the lowest point.

The increase in [t] duration between the higher ranks in the boundary hierarchy (starting with the PPh/C) is evidence for FL existing independent of focus accent in Swedish (cf. also results by Edwards & Beckman [11] who claim that FL does not exist below the PPh). In contrast, Lyberg and Ekholm [7] measuring only the stressed rhyme vowel, could not find any evidence of FL as an independent marker of the end of a phrase. (Neither could we find any independent lengthening when measuring the stressed [ε] in the present study (see also [4]). However, the observations on Swedish made by Lyberg and Ekholm may be given an alternative interpretation in the light of the present study. The claim they made that FL is a consequence of focus position on the last content word - they, like us, observed lengthening of the [+focus] stressed *vowel* - may therefore be a consequence of the segment they chose for analysis.

We also have reported a decrease of duration in the rhyme segments at the lower ranks in the boundary hierarchy. In our test word with a stressed syllable containing a short vowel [ε] followed by two consonants [n] and [t], it is particularly the consonants which are affected, though only in the [+focal] condition. Duration is greatest at the 0 boundary, less at the end of a PW and least at the end of the PPh/C word. Combined with the silent interval data we present, these results corroborate previous observations of a trading relation between FL and SI duration [8-9]. In contrast to the findings of the other researchers, however, negative correlations are not obtained for our data above the PPh/C level.

The trading effect may be looked upon as a means to optimize boundary signalling, maximizing segment duration cues when SI duration is at its minimum. However, why is this pattern more or less restricted to the [+focus] conditions? And how does one explain the fact that the significant +/- focus differences we have reported are much more pronounced, especially in the [n], at the lower ranked boundaries; the difference is greatest at the lowest ranked boundary, 0, and least at the PPh/C boundary. These questions need to be answered in future research.

In summary, what one can conclude from this study is that the phenomenon of Final Lengthening does exist in Swedish. Its domain would appear to be PPh and PU. It affects the final segment of the rhyme. Silent Intervals, moreover, are intimately tied to the higher-ranked boundaries, PPh and PU. Further, there appears to be a trading relation such that at the lower-ranked boundaries, segment and Silent Interval duration are negatively correlated.

ACKNOWLEDGEMENTS

We would like to thank Suzanne Haage-Palm for reading the material, Antonio de Serpa-Leitao for assistance with the recordings and Björn Granström for generously placing the sound studio at KTH at our disposal. This research was supported by grants from the Swedish HSFR/NUTEK Language Technology Programme.

REFERENCES

[1] Horne, M. & Filipsson, M. (1995), "Computational modelling and generation of prosodic structure in Swedish", *Proceedings, XIII ICPhS 95*, Stockholm.

[2] Crystal, T.H. & House, A.S. (1990), "Articulation rate and the duration of syllables and stress groups in connected speech", *J. Acoust. Soc. Am.*, vol. 88, pp. 101-112.

[3] C.W. Wightman, S. Shattuck-Hufnagel, M. Ostendorf & P. Price (1992), "Segmental durations in the vicinity of prosodic phrase boundaries", *J. Acoust. Soc.Am.*, vol. 91, pp. 1707-1717.

[4] Berkovits, R. (1994), "Durational effects in final lengthening, gapping, and contrastive stress",. *Language and Speech* , vol 37, pp. 237-250.

[5] Lyberg, B. (1981), *Temporal properties of spoken Swedish*. MILUS, vol. 6, Stockholm U.

[6] Bruce, G. (1981), "Tonal and temporal interplay", *Working Papers* (Dept. of Linguistics, U. of Lund), vol. 21, pp. 49-60.

[7] Lyberg, B. & Ekholm, B. (1994), "The final lengthening phenomenon in Swedish – a consequence of default sentence accent?" *Proceedings of ICSLP* 94, 135-138.

[8] Fant, G. & Kruckenberg, A. (1989),. Preliminaries to the study of Swedish prose reading and reading style. *STL-QPSR* 2.

[9] Fant, G., Kruckenberg, A., & Nord, L. (1991),. "Prosodic and segmental speaker variations", *Speech Communication* , vol. 10, pp. 521-531.

[10] Strangert, E., (1990), "Perceived pauses, silent intervals and syntactic boundaries", PHONUM (Dept Phon., Umeå Univ.), vol. 1, pp. 35-38.

[11] Edwards, J. and Beckman, M. (1988), Articulatory timing and prosodic interpretation of syllabic duration. *Phonetica* 45, pp. 156-174.