

ON THE AMOUNT AND DOMAIN OF FOCAL LENGTHENING IN SWEDISH

Eva Strangert and Mattias Heldner

Department of Phonetics, Umeå University, S-901 87 Umeå, Sweden

ABSTRACT

Temporal effects of focus in Swedish were studied in short sentences with systematic variation of the length and prosodic pattern of target words in different syntactic positions.

Generally, focus caused an average increase in word duration of about 25%. Variations of word length, stress, or word accent pattern did not produce any systematic effects on the amount of lengthening of the target word, but the lengthening varied considerably between speakers and different positions in the sentence. The most extensive lengthening occurred in combination with the insertion of a boundary after the word in focus in those cases when the word preceded a strong syntactic boundary.

Within words, stressed syllables were lengthened most, and lengthening of a primary and following secondary stressed syllable was equal to that of a single primary stressed syllable. As unstressed syllables were also affected, the domain of focal lengthening is assumed to be the word.

INTRODUCTION

Over the years, a good deal of attention has been paid to acoustic correlates of accents and focus. It is generally agreed that the primary correlate of focus in languages such as English, Dutch and Swedish is a pitch accent. It is also well known that the pitch accent is usually accompanied by other correlates such as prolonged durations within the focused word. However, the observed amount of focal lengthening varies greatly between different studies. For example, in data based on Dutch from one speaker, an average focal lengthening of about 25% was reported (Eefting, 1991). In a study of American English using six speakers, non-final focused words were 37% to 43% longer than non-focused, and the amount of focal lengthening was dependent on the word's position in the sentence (Cooper, Eady, and Mueller, 1985).

There is also the issue of the domain of accentual (or focal) lengthening, that is, what is lengthened in the focused word. Studies of Dutch (Eefting, 1991; Sluijter and van Heuven, 1995) suggest that the domain of lengthening is the word. All syllables and segments in Dutch words are longer when accented than when they are unaccented. Accenting a particular word, moreover, resulted in an almost linear time expansion of the entire word. Regarding the domain of lengthening in English, there is some disagreement in the literature. Sluijter and van Heuven (1996), on the one hand, report similar results for American English as for Dutch, that is, lengthening extending over the word as a whole. Turk and Sawusch (1997) and Turk and White (1997), on the other hand, basing their studies on

American and Scottish English respectively, suggest that the domain of accentual lengthening begins with the pitch-accented syllable and includes at least one following unstressed syllable. As the duration of unstressed syllables preceding the stressed syllable is unaffected by focus in their data, pre-stress syllables are not included in their lengthening domain. Studies of Swedish have shown that the domain of focal lengthening is at least the stressed syllable (Fant, Kruckenberg, and Nord, 1991), but most probably the entire word (Bruce, 1981).

The study to be reported here relates to a project on acoustic correlates to focus in Swedish and their perceptual relevance. Having established that focus may be signalled by other cues than F0 (Heldner and Strangert, 1997), we here examine the temporal effects of focus in short Swedish sentences. We want to (1) quantify the amount of focal lengthening and (2) investigate whether and to what extent the lengthening is influenced by the position in the syntactic structure and (3) specify the domain of lengthening.

METHOD

The material consists of three sets of meaningful three-word sentences, all with the same [_S NP [_{VP} V NP]] structure. Variations of these sentences were produced as answers to questions designed to elicit narrow focus on any one of the three words while keeping the syntactic and segmental structure constant.

The material in Set 1 includes two base sentences. One of them has acute accent words in all sentence positions (*Mannen tömmer dammen* 'The man is draining the pond') and the other only grave accented words (*Kvinnan dammar kannan* 'The woman is dusting the pitcher'). All the words are two-syllable non-compounds and stress is always on the first syllable. While the variations (of word accent pattern) in Set 1 occurred in all three sentence positions, Sets 2 and 3 consisted of a fixed sentence frame, where either the words in the first position (Set 2) or those in medial position (Set 3) were varied. The target words (simplex words and compounds) varied in length (2-5 syllables) and stress pattern (initially and non-initially stressed words). The simplex words had one (primary) and the compounds two (primary + secondary) stressed syllables. The material included acute as well as grave accented words, but there was no systematic variation of word accent. The syntactic, prosodic and segmental structure of the three sets of sentences is shown in Table 1.

The material in Set 1 was read by four native Swedish speakers, one male and three females, all used to reading aloud. Each base sentence occurred in six versions as answers to different questions (2 different questions to elicit focus in each position).

All versions were repeated twelve times by each speaker, thus yielding a total of 576 productions. Two speakers, identical to speaker 1 and 2 from Set 1, read the material in Sets 2 and 3. The two speakers read each version of the base sentences with appropriate focus on either the target word or the words in the sentence frame (depending on the preceding question) six times, yielding a total of 108 and 216 productions respectively in Sets 2 and 3.

Table 1: Syntactic, segmental and prosodic structure of the speech material. * = nonsense word

Set 1	[_S NP	'Mannen 'The man'
		'Kvinnan 'The woman'
	[_{VP} V	'tömmer 'is draining'
	NP]]	'dammar 'dusting'
Set 2	[_S NP	'Sëssan 'The princess' (abbr.)
		Prin'sëssan 'The princess'
		Kronprin'sëssan 'The crown princess'
	[_{VP} V	'säljer 'sells'
Set 3	[_S NP	'Mannen 'The man'
	[_{VP} V	'ställer 'parks'
		beställer 'reserves'
		'B-,ställer* 're-serves'
		'ävb,eställer 'cancels'
		'föravbeställer* 'pre-cancels'
		spe'cialbeställer 'special orders'
	NP]]	'bilen 'the car'

The recordings took place in a sound-treated room. The speakers were seated in front of a microphone and a computer screen. They were presented with one question-answer pair at a time on the screen and were instructed to read the answers appropriate to the questions. The question-answer pairs were presented in random order. A computer program (VoicePro™) handled the presentation of text on the screen as well as the recording. The answers were recorded to hard disk at 48 kHz sample rate, 16 bit amplitude resolution. In the subsequent analysis of the material from Set 1, all the words/sentence positions were included, while the analysis of the material in Sets 2 and 3 was restricted to the target words. Boundaries for all segments in the answers (Set 1) and for all segments in the target words (Sets 2 and 3) were determined and labelled using ESPS/waves+™. Syllable- and word-durations were calculated using the labelled segment boundaries.

RESULTS AND DISCUSSION

Since, according to the procedure described above, all three words in the answer sentences occurred in focus, there are two non-focused versions for each focused version of a word. As the differences between the two non-focused versions are usually very small, the average of the non-focused durations will be reported.

The data (based on Set 1) in Table 2 present duration means for focused and non-focused words and stressed syllables-to-word lengthening ratios (L-ratios) for each speaker and each syntactic position separately. In general, the mean word durations are

considerably longer when focused. The grand mean across all positions reveals that focused words are about 25% longer than non-focused (cf. Eefting, 1991) and there are no apparent differences between the acute and grave words. However, a closer inspection of the data reveals considerable variation in the amount of lengthening.

Table 2: Word (W) duration means for focal (+F) and non-focal (-F) words, differences between $\pm F$ in absolute values (ΔW) and as a percentage (%W). The last column (L-ratio) shows the ratio between the lengthening in the stressed syllable and in the word as a percentage (S/W). Data based on Set 1.

Word		Word duration				L-ratio S/W
		+F	-F	ΔW	%W	
Sp1	Mannen	733	504	229	45%	55%
Sp1	Kvinnan	767	527	240	46%	53%
Sp2	Mannen	560	443	117	26%	89%
Sp2	Kvinnan	600	469	131	28%	95%
Sp3	Mannen	455	331	124	37%	71%
Sp3	Kvinnan	490	355	135	38%	78%
Sp4	Mannen	453	435	18	4%	89%
Sp4	Kvinnan	499	433	66	15%	70%
[_S NP mean					30%	75%
Sp1	tömmer	567	445	122	27%	87%
Sp1	dammar	553	488	65	13%	77%
Sp2	tömmer	507	419	88	21%	91%
Sp2	dammar	475	383	92	24%	88%
Sp3	tömmer	360	286	74	26%	85%
Sp3	dammar	397	311	86	28%	87%
Sp4	tömmer	387	331	56	17%	100%
Sp4	dammar	415	347	68	20%	97%
[_{VP} V mean					22%	89%
Sp1	dammen	604	497	107	22%	64%
Sp1	kannan	629	514	115	22%	74%
Sp2	dammen	659	588	71	12%	61%
Sp2	kannan	739	594	145	24%	69%
Sp3	dammen	466	376	90	24%	82%
Sp3	kannan	551	408	143	35%	59%
Sp4	dammen	571	518	53	10%	55%
Sp4	kannan	663	546	117	21%	39%
NP]] mean					21%	63%

This variation seems to have two sources. One such source is pure speaker differences. Some speakers apparently lengthen words in focus more than others do; for example, 'Mannen' is lengthened by 4% by Speaker 4 and by 45% by Speaker 1. Comparing different positions in the sentence, we observe that 'tömmer' produced by the same Speaker 1 is lengthened by 27%. Thus position appears to be another source of variation (cf. Cooper, et al., 1985), and moreover, speaker and position seem to interact. Speaker 1 treats the first NP differently than the other speakers. The first NP is substantially lengthened, by 45%-46%, but only part of it, 53%-55%, comes from the stressed syllable. We believe that these data reflect a strategy used by speaker 1 to make the focused word more prominent by inserting a prosodic boundary after the word. One means of realizing this boundary is final lengthening in the unstressed syllable, and this would explain the relatively low values in the L-ratio column. We also believe that this strategy is determined by the syntactic structure, as it occurs in the first NP but not in the verb. It is possible to insert a boundary after the first NP since this is also the NP-VP

boundary. There is no boundary inserted after the verb since this would split the VP into two parts. There is always a boundary after the second NP and there is probably some final lengthening even in the non-focal versions of the second NP. We believe that this is why the L-ratios are somewhat higher in the second NP compared to the first NP but still lower than in the verb.

A similar pattern appears from the data given in Tables 3 and 4 (based on Sets 2 and 3). On average, the words in Sets 2 and 3 respectively are lengthened by 249 and 152 ms by speaker 1 and by 101 and 78 ms by speaker 2. There are no systematic influences of either word length or stress pattern on the amount of lengthening. On the basis of these data we assume that there is a speaker-specific focal lengthening component that is used irrespective of word length and stress pattern. We further assume that the means found in Set 3 may be good estimates of this component for the two speakers in this study. In addition to the focus component, we assume an optional final lengthening component supported by the data from speaker 1 in Set 1. The final lengthening, the extra duration given to the unstressed final syllable, is obviously restricted to Set 2, the words occurring immediately before the NP-VP boundary. Through the final lengthening component we can explain the differences between the data in Sets 2 (words before a strong syntactic boundary) and 3 (words occurring within a syntactic constituent) for speaker 1. As reported above, the words in Sets 2 and 3 are lengthened by 249 and 152 ms by speaker 1 and by 101 and 78 ms by speaker 2. Expressed as terms in a suggested two-factor model, speaker 1 then has a focus component of about 120-180 ms and a final lengthening component of about 100 ms. For speaker 2, who did not use final lengthening in Set 2, the focus component is about 80-100 ms. These results thus support our previous assumption that the variation found has two sources: speaker-specific differences and differences due to position in the syntactic structure.

The data reported here also reveal that all syllables contribute to the lengthening of the word, even though most of the durational increase (as shown in Tables 2, 3 and 4) is found in the stressed syllable (cf. Eefting, 1991, Fant, et al., 1991, Sluijter and van Heuven, 1995, Turk and Sawusch, 1997, Turk and White, 1997). There are no tendencies in our data to a linear time expansion of the entire word similar to that reported for Dutch in Sluijter and van Heuven (1995). Moreover, the part of the lengthening of the word that comes from the stressed syllable (as expressed by the L-ratios) also seems to vary with speaker and position in the syntactic structure, e.g. Speaker 1 has lower values for the first NP than the other speakers whereas the values for the verb and the second NP are similar to the others' (cf. Table 2).

From Tables 3 and 4 it appears that simplex and compound words are lengthened to a similar extent, but they differ in the distribution of lengthening. In a simplex word (with one primary stressed syllable) most of the lengthening (50%-100%) is observed to occur in the stressed syllable. In a compound word (with one primary and one secondary stress), there is less lengthening of the primary stressed syllable. However, the sum of the lengthening of the primary and secondary stressed syllables is about as great as the lengthening found in simplex words.

Table 3: Word (W) duration means for focal (+F) and non-focal (-F) words, differences between $\pm F$ in absolute values (ΔW) and as percentages (%W). The last column (L-ratio) shows the ratio between the lengthening in the stressed syllable and in the word as a percentage (S/W). Data based on Set 2.

	Word duration				L-ratio
	+F	-F	ΔW	%W	S/W
Sp1					
'Sëssan	823	590	233	39%	65%
Prin'sëssan	925	666	259	39%	57%
Kronprin'sëssan	1209	953	256	27%	50%
<i>mean</i>			249	35%	57%
Sp2					
'Sëssan	596	520	76	14%	100%
Prin'sëssan	754	630	124	20%	79%
Kronprin'sëssan	942	839	103	12%	65%
<i>mean</i>			101	15%	81%

Table 4: Word (W) duration means for focal (+F) and non-focal (-F) words, differences between $\pm F$ in absolute values (ΔW) and as percentages (%W). The last columns (L-ratios) show the ratio between the lengthening in the stressed syllable and in the word (S/W), and the ratio between the lengthening in the primary plus the secondary stressed syllables and the word (S+S/W) as percentages. Data based on Set 3.

	Word duration				L-ratios	
	+F	-F	ΔW	%W	S/W	S+S/W
Sp1						
'ställer	629	489	140	29%	88%	88%
be'ställer	734	564	170	30%	81%	81%
'B-ställer	868	718	150	21%	47%	84%
'ävbe'ställer	953	770	183	24%	23%	47%
'föravbe'ställer	1240	1094	146	13%	53%	72%
spe'ciälbe'ställer	1308	1184	124	10%	70%	85%
<i>mean</i>			152	21%	60%	76%
Sp2						
'ställer	480	402	78	19%	87%	87%
be'ställer	550	476	74	16%	84%	84%
'B-ställer	628	582	46	8%	61%	100%
'ävbe'ställer	788	653	135	21%	27%	53%
'föravbe'ställer	1008	924	84	9%	80%	96%
spe'ciälbe'ställer	1083	1035	48	5%	85%	100%
<i>mean</i>			78	13%	71%	87%

Apart from the lengthening occurring in stressed syllables as well as in the unstressed final syllable in the first NP - the final lengthening assumed to be an optional focus marker - unstressed syllables in other positions are to some extent lengthened also. From these data we conclude that the domain of lengthening in Swedish includes at least the primary stressed syllable and, if the word contains a secondary stress, the domain extends from the beginning of the primary stressed to the end of the secondary stressed syllable including the intervening syllables. However, we might as well equate the domain with the word, as all syllables in a word, including unstressed syllables in pre-stress position, seem to have the potential for being affected by focal lengthening. Thus our results conform to those reported for Dutch by, for example, (Eefting, 1991) and (Sluijter and van Heuven, 1995).

CONCLUSIONS

To conclude, this study shows that lengthening is a common but perhaps optional correlate of focus. Focused words are about 25% longer than non-focused in general in our data, but there are great differences between speakers and between positions in the syntactic structure.

We find no systematic effects of word length, stress, or word accent pattern on the amount of lengthening. Regarding position effects, we believe that the insertion of a prosodic boundary after the focused word in certain syntactic positions (strong syntactic boundaries) is a way of making the word more prominent and hence part of the signalling of focus. As this is not an obligatory pattern, we assume focal lengthening to have two components, a focal and an optional final lengthening component.

Within the focused word, most of the lengthening occurs in the primary stressed syllable, but lengthening of a primary and following secondary stressed syllable (in compounds) equals that of a single primary stressed syllable. As the unstressed syllables also may contribute to the lengthening of the word it seems reasonable to assume the domain of lengthening to be the word.

ACKNOWLEDGEMENTS

We are grateful to Anna Aasa who carried out part of the acoustic analyses. The work was supported by a grant from the Swedish Council for Research in the Humanities and Social Sciences.

REFERENCES

1. Bruce, G., "Tonal and temporal interplay", In T. Fretheim (Ed.), *Nordic prosody II: Papers from a symposium* Tapir, Trondheim, 1981, p 63-74.
2. Cooper, W. E., Eady, S. J., and Mueller, P. R., "Acoustical aspects of contrastive stress in question-answer contexts", *Journal of the Acoustical Society of America*, Vol. 77(6), 1985, p 2142-2156.
3. Eefting, W., "The effect of "information value" and "accentuation" on the duration of Dutch words, syllables, and segments", *Journal of the Acoustical Society of America*, Vol. 89(1), 1991, p 412-424.
4. Fant, G., Kruckenberg, A., and Nord, L., "Durational correlates of stress in Swedish, French and English", *Journal of Phonetics*, Vol. 19, 1991, p 351-365.
5. Heldner, M., "To what extent is perceived focus determined by F0-cues?", In *Eurospeech '97 Proceedings: ESCA 5th European conference on speech communication and technology* ESCA, Rhodes, Greece, 1997, p 875-877.
6. Sluijter, A., and van Heuven, V. "Acoustic correlates of linguistic stress and accent in Dutch and American English", In *Proceedings from ICSLP'96*, University of Delaware and Alfred I. duPont Institute, Philadelphia, US, 1996, p 630-633.
7. Sluijter, A. M. C., and van Heuven, V. J., "Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch", *Phonetica*, Vol. 52, 1995, p 71-89.
8. Turk, A. E., and Sawusch, J. R., "The domain of accentual lengthening in American English", *Journal of Phonetics*, Vol. 25(1), 1997, p 25-41.
9. Turk, A. E., and White, L., "The domain of accentual lengthening in Scottish English", In *Eurospeech '97 Proceedings: ESCA 5th European conference on speech communication and technology* ESCA, Rhodes, Greece, 1997, p 795-798.