2nd International Conference on
Spoken Language Processing (ICSLP 92)
Banff, Alberta, Canada
October 12-16, 1992

PROSODIC CUES TO THE PERCEPTION OF SYNTACTIC BOUNDARIES

Eva Strangert

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

ABSTRACT

The reported experiments were designed to study the contribution of prosodic cues to syntactic boundary categorization. The results indicate (1) that it is possible to differentiate between syntactic boundaries on the basis of prosodic cues alone (2) that the best results are obtained when pre- and post-boundary information is combined with information about the boundary itself (the silent interval) (3) that a fairly good categorization may be based exclusively on pre-boundary cues and (4) that the silent interval appears to be the stronger cue when in conflict with other prosodic cues.

INTRODUCTION

The experiments reported below extend previous work on pausing and syntactic boundaries in Swedish.

Strangert [1] investigated prosodic characteristics of perceived pauses at sentence, clause and phrase boundaries in text reading. Silent intervals were measured as well as pre-boundary lengthening and Fo characteristics before and after the boundary. The results indicated that the different types of boundaries were signalled acoustically in different ways. Generally, the higher the rank of the boundary, the stronger were its acoustic correlates.

Some preliminary perceptual experimentation [2] points to the length of the silent interval as a powerful indicator of syntactic boundary type. Subjects asked to adjust the silent interval at different boundaries, using material from the text-reading study, closely replicated the temporal relations found in the production data. It seems, then, that this interval has to be adjusted to the specific boundary, if speech is to sound acceptable. However, listeners frequently perceive a pause, even though no silent interval exists [3]. Pauses apparently may be signalled by other acoustic cues than silence.

The experiments to be reported aim at exploring whether it is possible to differentiate between syntactic boundaries on the basis of prosodic cues alone. Such cues may include the length of the silent interval, pre-boundary lengthening, as well as Fo and intensity characteristics. In addition, the respective contributions of the silent interval and other prosodic information as indicators of syntactic boundary type are explored.

Recordings were made of a speaker reading three different sentence sequences in Swedish. Each sequence occurred four times. All the material (3 x 4 sequences) was read in random order. The sequences were all identical in length and segmental composition, but differed in their syntactic structure. After *9.15-tåget* there was either a sentence boundary (A), a clause boundary (B) or a phrase boundary (C):

A) *Han blev tvungen att ta 9.15-tåget. Sedan Dessy hade sålt bilen i januari, blev det lugnare.*
'He had to take the 9.15 train. After Dessy had sold the car in January, things became calmer.'

B) *Han blev tvungen att ta 9.15-tåget, sedan Dessy hade sålt bilen. I januari blev det lugnare.*
'He had to take the 9.15 train after Dessy had sold the car. In January things became calmer.'

C) *Han blev tvungen att ta 9.15-tåget sedan dess. Y hade sålt bilen. I januari blev det lugnare.*
'He had to take the 9.15 train since then. Y had sold the car. In January things became calmer.'

After recording them, the sequences were cut so as to include only the initial part, containing a sentence, clause and phrase boundary, respectively. The portions used for experimentation have been underlined above. Some measurements made from mingograms of these test sequences, including the four renderings of each, are presented in Table 1.

Table 1. Fo and duration values measured from mingograms of the test sequences.

	Fo fin* in Hz	Dur (t)åget in msec	Silent interv in msec	Dur (s)edan in msec	Fo init** in Hz
A1	105	425	840	255	115
A2	100	420	735	250	115
A3	95	435	1020	270	110
A4	100	410	1135	215	125
$\bar{x}$	*100*	*423*	*933*	*243*	*116*
B1	110	445	420	255	125
B2	100	460	265	250	125
B3	110	360	0	250	125
B4	110	300	0	215	125
$\bar{x}$	*108*	*453/330*	*343/0*	*243*	*125*
C1	115	310	0	260	125
C2	115	320	0	245	125
C3	125	290	0	245	135
C4	125	320	0	265	135
$\bar{x}$	*120*	*310*	*0*	*254*	*130*

*measured at the lowest point in *tåget*
** measured at the first stable voice period in *sedan*

50 university students participated in the experiments, 10 in each of the 5 separate experiments run (for details, see below). Listening to the test sequences through headphones, the subjects were required to judge what kind of boundary occurred after *9.15-tåget* by indicating whether there was a period (sentence boundary), a comma (clause boundary) or nothing (phrase boundary) following that word. The test sequences occurred in random order in blocks. Each specific test sequence occurred 10 times. Thus, there were 100 judgements (10 occurrences x 10 subjects) for each sequence.

EXPERIMENT 1

The test sequences occurring in this experiment were:

A) *Han blev tvungen att ta 9.15-tåget. Sedan ...*
B) *Han blev tvungen att ta 9.15-tåget, sedan ...*
C) *Han blev tvungen att ta 9.15-tåget sedan ...*

preserving pre- and post-boundary information as well as the boundary itself (the silent interval occurring at a sentence, A, clause, B, and phrase, C, boundary).

The categorization of each of the four renderings of the test sequences is shown in Table 2. The confusion appears mainly as sentence>clause, clause>sentence, clause>phrase and phrase>clause. The closer in rank the boundaries are, the more they are confused. (Lehiste, [4] has reported a similar confusion between paragraph and sentence boundaries.) Moreover, the categorization differs somewhat for the different renderings of the test sequences. This is a reasonable consequence of their different acoustic realizations (see Table 1). Acoustic differences, the length of the silent interval in particular, seem to be the reason for the very striking difference in the categorization of B1 and B2 (mainly clause) on the one hand, and B3 and B4 (mainly phrase) on the other. The clause>phrase confusion (test sequences B3 and B4) is considerable. However, B3 and B4 are the only exceptions to the general tendency of small confusion in these data.

Table 2. Categorization of the boundary in 12 test stimuli (4 renderings of each of 3 test sequences) including pre- and post-boundary information and a silent interval co-occurring with a sentence (A), clause (B) or phrase (C) boundary. Percentages refer to 10 judgements of each test stimulus by 10 listeners.

% judged as	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4
sentence	89	85	94	93	10	2	1					
clause	11	15	5	3	85	89	19	14	8	21	13	18
phrase			1	4	5	9	80	86	92	79	87	82
	sentence				clause				phrase			

EXPERIMENT 2

In this case there was only pre-boundary information. The test sequences were:

A) *Han blev tvungen att ta 9.15-tåget. ...*
B) *Han blev tvungen att ta 9.15-tåget, ...*
C) *Han blev tvungen att ta 9.15-tåget ...*

The results are given in Table 3. Generally, though the confusion pattern is qualitatively the same as in Experiment 1, the successful categorizations are fewer. This holds especially for phrase and clause boundaries. Sentence boundaries, on the other hand, have high identification scores, though not as high as in Experiment 1. As in Experiment 1, the different renderings of the test sequences get different scores. However, an interesting difference between the two experiments is the higher scores in this experiment for B3 and B4. Those are the renderings with a clause boundary which were categorized almost exclusively as "phrase" in Experiment 1. Apparently, when there is no longer any information indicating a phrase boundary (a zero silent interval) the number of correct categorizations increase for B3 and B4. These cases as well as the general pattern emerging from a comparison of Experiments 1 and 2 (see Figure 1) point to the length of the silent interval as an important indicator of syntactic boundary type.

Table 3. Categorization of the boundary in 12 test stimuli (4 renderings of each of 3 test sequences) including preboundary information of a sentence (A), clause (B) or phrase (C) boundary. Percentages refer to 10 judgements of each test stimulus by 10 listeners.

% judged as	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4
sentence	76	95	92	73	27	14	2		10	1	1	
clause	23	5	7	22	57	65	41	29	50	39	34	41
phrase	1		1	5	16	21	57	71	40	60	65	59
	sentence				clause				phrase			

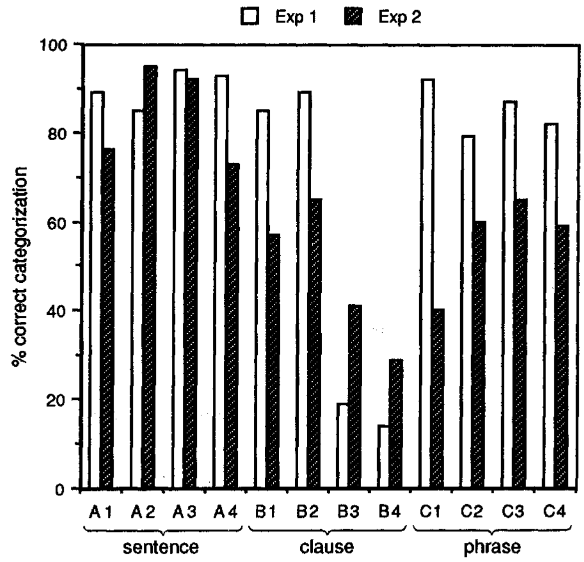

Figure 1. Comparison of judgements in Experiments 1 and 2.

EXPERIMENTS 3-5

These experiments were designed to study the respective perceptual contributions of pre- and post-boundary information as well as the silent interval. The test material consisted of modified versions of three test sequences from Experiment 1 with a sentence, clause and phrase boundary, respectively. In these experiments the pre-boundary portions, the silent intervals, and the post-boundary portions of the three test sequences occurred in all possible combinations (9 in each of the experiments), according to the following scheme:

Experiment 3: pre- + post-boundary same; silent interval same or different: AAA, ABA, CBC etc.

Experiment 4: pre-boundary + silent interval same; post-boundary same or different: AAA, AAB, CCB etc.

Experiment 5: silent interval + post-boundary same; pre-boundary same or different: AAA, ABB, CBB etc.

In these experiments then, the silent interval and the other cues to boundary categorization were sometimes in conflict. When this happened, a complex response pattern emerged. Table 4 shows the results from Experiment 3. Apparently, both the silent interval <u>and</u> the pre-and post-boundary information contribute to the categorization. However, the silent interval seems to be the stronger cue. This conclusion gets even stronger support from Experiments 4 and 5. Though this is the general tendency, individual subjects vary to some extent. The majority, of course, were influenced by interval length in combination with the other cues. However, some subjects seemed to respond on the basis of interval length alone, while for a few the silent interval was the weaker cue.

Table 4. Categorization of the boundary in 9 test stimuli (3 versions of each of 3 test sequences) combining pre- and post-boundary information, and silent interval duration according to the scheme in Experiment 3 (see text). Percentages refer to 10 judgements of each test stimulus by 10 listeners.

% judged as	AAA	BAB	CAC	ABA	BBB	CBC	ACA	BCB	CCC
sentence	95	65	57	43	11	4	16		
clause	4	32	30	48	78	71	25	24	5
phrase	1	3	13	9	11	25	59	76	95

CONCLUSIONS

Summarizing, the following conclusions may be drawn from the reported experiments:

- It is possible to differentiate between different syntactic boundaries on the basis of prosodic cues alone.
- The best scores (76% correct on average) are obtained when information from the silent interval is combined with pre- and post-boundary information.
- The boundary itself (the silent interval) is not neccessary to differentiate between different boundaries. There is enough pre-boundary information to attain a fairly good categorization (63% correct on average).
- When there is conflicting information in the silent interval and the pre- and post-boundary portions, the silent interval is the main contributor to boundary categorization for the majority of the subjects.

ACKNOWLEDGEMENT

The present work has been carried out within the Language Technology Programme under a grant from the Swedish Council for Research in the Humanities and Social Sciences.

REFERENCES

[1] E. Strangert. "Pauses, syntax and prosody." In *Nordic Prosody*, ed. by K. Wiik and I. Raimo University of Turku, Phonetics, pp. 294-305. 1990 a.

[2] Unpublished work in collaboration with Rolf Carlson and Björn Granström, KTH, Stockholm.

[3] E. Strangert. "Perceived pauses, silent intervals and syntactic boundaries." *PHONUM 1*, Reports from the Department of Phonetics, University of Umeå, pp. 35-38. 1990 b.

[4] I. Lehiste. "Perception of sentence and paragraph boundaries." In *Frontiers of Speech Communication Research*, ed. by B. Lindblom and S. Öhman. London: Academic Press, pp. 191-201. 1979.