PAUSING IN TEXTS READ ALOUD

E. Strangert

Department of Phonetics, University of Umeå, Sweden

ABSTRACT

Perceived pauses in Swedish news texts read aloud were investigated. The pauses were analyzed to determine their distribution as well as their acoustic correlates and the perceptual relevance of these correlates. Most pauses occurred at syntactic boundaries, and the higher the rank of the boundary, the greater the probability of a pause. The acoustic correlates of pauses, in addition to silence, include prepausal lengthening, resetting of intensity and Fo, and voice quality irregularities. In general, the higher the rank of the boundary, the stronger and more varied were the acoustic correlates. Moreover, the data demonstrate that syntax plays a role not only in the production but also in the perception of pauses.

1. INTRODUCTION

This paper reports results from an ongoing project about pausing in Swedish. First, it concerns pausing in texts read aloud. Thus, the analysis only marginally includes hesitations and other phenomena that characterize ordinary speech situations. Secondly, the project combines a prosodic and a syntactic as well as a textual perspective on pauses. The purpose is to describe where pauses occur in relation to language structure, in particular to boundaries of different kinds. The purpose is, moreover, to learn about how these pauses are manifested acoustically, and finally, how the acoustic correlates contribute to the impression of a pause. Thirdly, "pause" in this study means "perceived pause". The focus is on those parts in the speech stream at which a pause is heard. By choosing this rather than an acoustic definition, pauses without a silent interval will not be excluded from analysis. The study includes normal, fast and slow renderings of the texts. A detailed account of the purpose and general outline of the project is given in [13]. Other studies with a similarly wide perspective on pausing include [10, 2, 15, 4].

2. MATERIAL AND ANALYSES

The material consisted of two news cables with a total of 810 words. Some of the original words had been exchanged for specific test words inserted in different syntactic positions to make it possible to study prepausal lengthening at different types of boundaries.The texts were read by ten male speakers. Each one read the material at his normal speed and at a faster as well as a slower speed. All the material was recorded on tape and registered on mingograms.

Prior to further analyses, two listeners identified the pauses from the recordings. Of the total number of pauses identified, the interrater reliability varied between 78 and 94% for the different speakers. These percentages may be compared to the 72% reliability in a Dutch study by de Rooij [10]. de Rooij had five persons listening to one speaker which reasonably should give a lower figure.

A syntactic analysis of the texts was also carried out with units such as paragraphs, sentences, clauses and phrases defined as in traditional grammar. The boundaries separating these units were marked as paragraph ($$), sentence ($), clause (//) and phrase (/) boundaries, respectively [13].

3. PAUSE DISTRIBUTION

The occurrence of pauses in relation to language structure has been investigated for different languages and conditions. Studies of speech read aloud have been based on German [2], English [15] and Dutch [10, 1].

In the present study some positions seemed to almost obligatorily attract pauses, while in other positions the occurrence of pauses varied for the different speakers. Positions where at least five of the speakers made a pause perceived by both listeners were termed "strong pause positions". All strong pause positions coincide with syntactic boundaries, and as might be expected, all paragraph and sentence boundaries are strong pause positions, independently of speech rate. For clause and phrase boundaries, speech rate is more important. The slower the speech, the more frequent the pauses. Figure 1 shows how strong pause positions are distributed over clause and phrase boundaries.

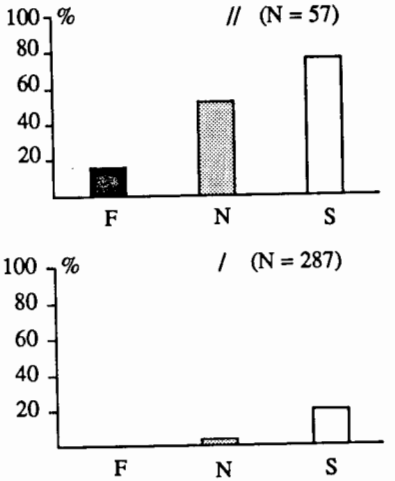

Figure 1. Relative frequency of strong pause positions at clause and phrase boundaries, in percent. Fast (F), normal (N), slow (S) rate. Based on 10 speakers.

A detailed analysis of the data indicates that different kinds of clauses did not attract pauses to the same degree. For example, complement clauses starting with *att* 'that' (as in *He said that* ...) were almost never preceded by a pause, not even at a slow speech rate. Temporal clauses, on the other hand, were generally preceded by a pause, as were conjoined clauses. Pauses also occurred very frequently before main clauses and some of the relative clauses. A similar pattern emerges from German data with comparable clause categories [2].

However, clauses of the same type were sometimes delimited by a pause, sometimes not. Length may be an important factor in these cases, as it seems that the probability of a pause between clauses is higher the longer and more complex the clause. Alternatively, it is not length but information load that is the important factor. Longer clauses and clause fragments contain more information than shorter ones. Thus, pausing may be a means for avoiding the clustering of too much information. That semantics plays a role for the insertion of pauses is supported also by the phrase data. The few phrase boundaries that were strong pause positions all delimited phrases with a high information load, viz., complex adverbial phrases and phrases expressing negation.

Thus, the present study suggests a multifactorial influence on pause distribution. (See [11, 13] for a more detailed account.) A similar complex basis for pausing is discussed by Umeda [15]. To isolate these determinants has the highest priority when it comes to predicting pausing. A number of studies have developed pausing algorithms as a means for revealing the "performance structures" of sentences [3, p 182-193; 6].

4. ACOUSTIC CORRELATES

So far, the normal rate data for six of the speakers have been analyzed. Measurements were made of silent intervals, test word durations (to estimate prepausal lengthening), as well as Fo before and after pauses. There was also an evaluation of voice quality irregularities before (and after) pauses. Figure 2 presents data for silent intervals.

It is apparent that even though the absolute durations vary widely between the speakers, they follow the same pattern: The duration of the silent interval matches the rank of the boundary. If the mean silent interval at paragraph boundaries is given a duration of 1 for each of the speakers, then at sentence boundaries the mean silent interval is about .6 and at clause boundaries about .2 of the reference duration. In general the mean silent interval at phrase boundaries is somewhat

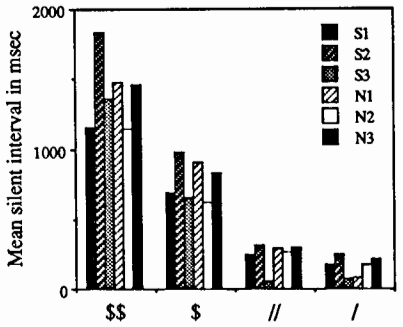

Figure 2. Mean silent intervals at paragraph, sentence, clause and phrase boundaries. Data for six speakers.

shorter than at clause boundaries, but the differences between these categories are very small [12]. Butcher [2, p 175-179] similarly measured silent intervals between sentences as well as between different types of clauses. As in the present study, the intervals were longer between sentences than between clauses. In addition, Butcher found significant differences of the silent intervals within the clause category.

There is a positive correlation between the acoustic signalling and the rank of the boundary for other pause correlates, too [11, 14]. Fo before a pause tends to drop to a lower value, and Fo after a pause tends to start at a higher value the higher the rank of the boundary. Thus, the resetting is greatest at paragraph boundaries. Irregularities of voicing, e. g. creaky voice, present a similar pattern. Most pauses with such irregularities occur at paragraph and sentence boundaries, and the higher the rank of the boundary, the stronger the irregularities. However, prepausal lengthening deviates from the general trend. There is no apparent positive correlation between the degree of lengthening and the boundary rank. This fits in with the observation that there is no obvious difference between the lengthening before a sentence and a paragraph boundary [8]. Several studies indicate complementarity between lengthening and the following silent interval [4, 5].

5. PERCEPTUAL ASPECTS

The pauses in this study were aurally identified whereupon acoustic data related to the pause positions were collected. This procedure does not permit conclusions as to the perceptual significance of the respective correlates or how they combine to the impression of a pause. (There may also be other relevant correlates than those which were chosen. In fact, it seems that resetting of intensity is such a correlate.) So far some preliminary observations have been made.

There is a high proportion of pauses without a silent interval. Over the six speakers the proportion ranges between 7 and 26%. In addition, there are many pauses with silent intervals 200 msec or shorter. Apparently there are other cues than silence to pause perception. Obvious candidates are Fo and intensity resetting, prepausal lengtening and voice quality irregularities. Several studies have shown that lengthening before a syntactic boundary may be a cue to boundary perception [7, 9]. Fo and intensity seem to be used as cues, too, but they are less effective than duration cues, including lengthening and silence [9, and references cited there]. A study of sentence and paragraph boundary perception points to a complex interaction of lengthening, voice quality irregularities (laryngealization), and silence [8].

Silence seems to be the more powerful cue. This may be inferred from the previous work cited above as well as from the present data. For example, listener agreement was 100% or close on pauses with silent intervals longer than 200 msec. It was the pauses with no or very short silent intervals (0-200 msec) that the listeners did not agree upon [12].

The silent interval, moreover, has to be adjusted to the specific boundary type. This conclusion may be drawn from a pilot experiment [16]. Three sections of the original recording of one speaker reading at his normal speed were stored digitally. The three sections each contained a boundary at which a pause had been perceived; one sentence boundary and two clause boundaries, one of which was longer than the other. In each section the boundary under study was preceded and followed, respectively, by a stretch of speech starting at the immediately preceding and following pause (boundary). A speech editing program made it possible for subjects to adjust the silent interval over a range from 0 to 1000 msec. The sections were tested one at a time and the subjects alternated between setting a duration and listening to the result until they decided they had found the optimal silent interval. Each of the sections were tested three times in this way. The results are presented in Figure 3, which contains the original durations produced by the speaker alongside with the adjusted durations averaged over the three trials and nine subjects. Though the adjusted intervals are generally shorter than those originally produced, the temporal relations between the three boundaries are more or less the same in production and perception. These data, moreover, demonstrate that syntactic structure plays a role in the production as well as the perception of pauses.

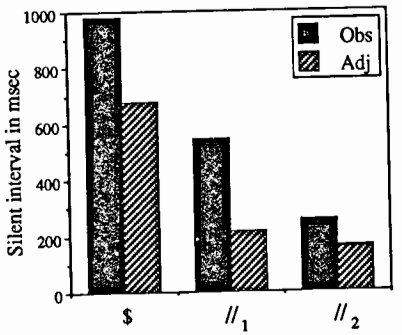

Figure 3. Originally produced silent intervals and adjusted intervals at one sentence and two clause boundaries. Averaged over 3 trials and 9 subjects.

6. REFERENCES

[1] BRINGMANN, E. (1990), "The distribution of Dutch reading pauses: A preliminary investigation on the influence of prosodic phrasing and punctuation on pause duration", Doctoral paper, Utrecht University.

[2] BUTCHER, A. (1981), "Aspects of the speech pause: Phonetic correlates and communicative functions", *Arbeitsberichte, 15,* Institut für Phonetik, Universität Kiel.

[3] COOPER, W. & PACCIA-COOPER, J. (1980), "*Syntax and speech*", Cambridge, Mass: Harvard University Press.

[4] FANT, G. & KRUCKENBERG, A. (1989), "Preliminaries to the study of Swedish prose reading and reading style", *STL-QPSR, 2/1989.*

[5] FARNETANI, E. (1989), "Acoustic correlates of linguistic boundaries in Italian: A study on duration and fundamental frequency", *Eurospeech '89*, vol 2, pp 332-335.

[6] GEE, J. P. & GROSJEAN, F. (1983), "Performance structures: A psycholinguistic and linguistic appraisal", *Cognitive psychology, 15*, 411-458.

[7] KLATT, D. H. (1976), "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence", *JASA, 59, 5*, 1208-1221.

[8] LEHISTE, I. (1979), "Perception of sentence and paragraph boundaries", In: Lindblom, B. & Öhman, S. E. G. (eds), *Frontiers of speech communication research*, New York: Academic press, pp 191-201.

[9] LEHISTE, I. (1980), "Phonetic manifestation of syntactic structure in English", *Annual bulletin, Research Institute of Logopedics and Phoniatrics, Tokyo, 14*, 1-27.

[10] ROOIJ DE, J. J. (1979), "*Speech punctuation. An acoustic and perceptual study of some aspects of speech prosody in Dutch*", Dissertation, Rijksuniversiteit Utrecht.

[11] STRANGERT, E. (1990a) "Pauses, syntax, and prosody", In: Wiik, K. & Raimo, I. (eds.), *Nordic prosody V*, Phonetics, University of Turku, pp 294-305.

[12] STRANGERT, E. (1990b) "Perceived pauses, silent intervals, and syntactic boundaries", *PHONUM, 1*, 35-38, Department of Phonetics, University of Umeå.

[13] STRANGERT, E. (1991), "Where do pauses occur in speech read aloud?", *Proceedings from The Twelfth Scandinavian Conference of Linguistics*, June 14-16, 1990, Reykjavík, forthcoming.

[14] STRANGERT, E. & ZHI, M. (1989), "Pause patterns in Swedish: A project presentation and some data", *STL-QPSR, 1/1989*, 27-31.

[15] UMEDA, N. (1982), "Boundary: Perceptual and acoustic properties and syntactic and statistical determinants", In: *Speech and language: Advances in basic research and practice*, vol 7, New York: Academic Press, pp 333-371.

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