

PAUSING IN SWEDISH SPONTANEOUS SPEECH

Petra Hansson

Department of Linguistics and Phonetics
Lund University
Heigonabacken 12, S-223 62 Lund, Sweden

ABSTRACT

Pauses in spontaneous speech have a less restricted distribution than pauses in read discourse; however, they are not distributed in a haphazard way. The majority of the perceived pauses in the examined Swedish spontaneous speech material, 73%, occurred in one of the following positions: between sentences, after discourse markers and conjunctions, and before accented content words.

There is a range of acoustic correlates of perceived pauses in spontaneous speech, such as silent intervals, hesitation sounds, prepausal lengthening, glottalization and specific F0 patterns. The acoustic manifestation of a pause, e.g. the duration of the pause and the F0 pattern associated with the pause, is to some extent dependent on the pause's position and function.

1. INTRODUCTION

Prosody plays an important role in the structuring of spoken language. Structure is assigned to the flow of speech by marking boundaries between sequences of words that belong together on syntactic, semantic and pragmatic grounds and by marking coherence within those sequences of words [1, 2]. One way of signalling a boundary between two coherent units of speech is by inserting a pause. A pause occurring in a place that is logical in terms of linguistic factors, such as e.g. syntactic structure, serves the purpose of making understanding easier for the listener.

The dividing of speech into chunks with pauses in between is not only necessary for the perception of spoken language, but also needed in the production of speech, e.g. in order to provide the speaker with time to breathe. In the production of spontaneous speech, speech planned on-line, the pauses also provide the speaker with time for speech planning and choice of individual words [3]. However, pauses reflecting the speaker's planning of a sentence or search process for a specific word do not necessarily have to occur at syntactically and semantically motivated positions. Therefore, while pauses in read speech have been shown to coincide to a great extent with syntactic boundaries, such as sentence, clause and phrase boundaries [4], pauses in spontaneous speech are known to have a freer distribution [5].

Pauses in spontaneous speech are not distributed haphazardly, however. In this study, the aim is to describe the positions of pauses in a Swedish spontaneous speech material and determine what the acoustic correlates of pauses are. 'Pause' is defined perceptually, i.e. by pause is meant a perceived pause, regardless of whether it can be associated with a silent interval in the speech signal or not.

2. SPEECH MATERIAL AND ANALYSIS PROCEDURE

The speech material used in the study consisted of a nine-minute long monologue. A male speaker of a southern Swedish dialect was recorded while narrating the content of a fragment of a silent film (*The Last Laugh*) which was shown to him prior to the recording.

Three analyses of the material were carried out. First, a transcription of the monologue was used to identify the syntactic boundaries, i.e. potential pause positions in the material. However, performing a syntactic analysis on a spontaneous speech material is by no means an easy task. A modified version of the method of analysis developed within the research project 'Prosodic segmentation and structuring of dialogue' [6] for segmentation of spontaneous speech was used.

Subsequent to the syntactic analysis of the material, two phonetically untrained native speakers of Swedish were asked to listen to the monologue and independently mark the positions of the pauses they heard in an orthographic transcription. In order not to influence the listeners' perception of pauses, the transcription did not contain any punctuation marks, nor was it segmented. Of the total number of identified pause positions, the two listeners agreed in 80% of the cases (249 cases of 311). The pauses perceived by both listeners were included in the study.

Finally, a phonetic-acoustic analysis of the recorded material was conducted in the ESPS/waves+™ environment. The acoustic correlates of the perceived pauses were investigated with the help of spectrograms, the waveform and fundamental frequency curves. Measurements of pause duration and of F0 before and after silent intervals were made.

After listening to the monologue and visually inspecting the F0 curve, the transcription was completed with information about the distribution of accents. In the variant of Swedish spoken by the examined speaker, the difference between focal and non-focal accentuation is not clear, and therefore only one level of prominence (as distinct from no prominence) has been recognized.

3. RESULTS AND GENERAL DISCUSSION

3.1. Pause positions

By comparing the positions of the perceived pauses with the positions of the syntactic boundaries in the monologue, we can confirm that pauses in spontaneous speech have a freer distribution than pauses in read texts. Pauses in spontaneous speech are not restricted to occur only in positions where they are motivated in terms of syntactic structure. Although pausing frequently occurs between sentences or sentence-like utterances (incomplete sentences, sentence fragments, etc.), pauses also abound after discourse markers and conjunctions and before accented content words in the examined material.

As might be expected, the pauses occurring in sentence boundaries constitute a large part, 25%, of the analyzed perceived pauses (see example (1)). It is generally assumed that the sentence boundary is a strong pause position, possibly reflecting the externalisation of a thought or idea [3].

- (1) där hänger nåt som ser ut som en brudklänning va (p) han tittar på den
'something is hanging there that looks like a wedding dress right (p) he looks at it'

In texts read aloud, pausing between sentences has previously been observed to be obligatory in Swedish [4]. However, a sentence boundary does not appear to be an obligatory pause position in spontaneous speech. Although the sentence boundary constitutes the syntactic boundary at which pausing is most frequent in the material, only 57% of the total number of analyzed sentences are preceded by pauses. On the other hand, the numerous pauses after discourse markers could also be assumed to reflect the planning of upcoming sentences.

Of the total number of perceived pauses, 30% are situated after discourse markers (or 'cue words') and conjunctions (see example (2), the words *men* 'but' and *i alla fall* 'anyway' can function as indicators of discourse structure in Swedish [7]). Their position is interesting since pausing after a discourse marker or a conjunction, instead of pausing at the syntactic boundary preceding the marker or conjunction, can be seen as a means of holding the floor. The speaker can pause to plan the upcoming sentence or clause with no risk of giving the impression of having finished speaking. The commonness of this manner of pausing is reflected in the fact that pauses after subordinate conjunctions are as frequently occurring as pauses preceding subordinate clause boundaries in the examined material.

- (2) men (p) i alla fall (p) han går iväg (p) och går till arbetet
'but (p) anyway (p) he leaves (p) and goes to work'

Pauses before content words constitute 18% of the analyzed pauses (see example (3), *sprits* 'cake decorator' is accented). They occur immediately before an accented content word or precede the phrase in which the accented word occurs. It is not unreasonable to assume that these pauses reflect that the speaker is in the process of choosing an important word. It has also been proposed that they contribute to the prominence of the following accented word [5, 8].

- (3) och så så (p) skriver hon på den med en sån här (p) sprits
'and then then (p) she writes on it with one of those (p) cake decorators'

The remaining 27% of the examined pauses appear mainly at clause boundaries, before adverbial phrases and phrases expressing negations, between repeated words (usually a conjunction or the subject of the sentence), and before appositions and lexical hedges or fillers.

3.2. Acoustic correlates

The majority of the perceived pauses, 82%, can be associated with a silent interval of varying duration in the speech signal (ranging from 80 to 1790 ms in duration). Silent pauses can be found in all the pause positions mentioned above. They often co-occur with other acoustic correlates, such as vowel-like insertions (hesitation sounds), prepausal lengthening, changes in voice quality (glottalization or 'creak'), and specific F0 patterns. A gradual drop in intensity also appears to be a potential correlate of pauses, but it has not been studied in any detail here. In the cases of very short silent intervals, the pauses are most likely perceived only because of the fact that they occur in combination with these other correlates of pauses.

Not all the perceived pauses in the speech material can be related to a silent interval. Vowel-like insertions or hesitation sounds, although often co-occurring with silent intervals preceding and / or following them, were also by themselves perceived as pauses by the listeners. The majority of the vowel-like insertions (19 of 29) precede a sentence boundary or occur after a discourse marker. Insertions after discourse markers immediately follow the marker, while insertions in sentence boundaries are preceded by silent intervals. The hesitation sound inserted by the recorded speaker is, as has been reported for speakers of Swedish previously [5], the vowel [e].

Pauses that can neither be associated with a silent interval in the speech signal or a hesitation sound constitute 14% of the total number of perceived pauses. Other correlates of pauses can, in combination or by themselves, contribute to the perception of a pause. Prepausal lengthening (final lengthening) and F0 patterns involving a boundary-marking pitch movement seem to be sufficient for the perception of a pause. Glottalization, however, is usually combined with a drop in both F0 and intensity. In a few cases, the repetition of a word gives rise to the perception of a pause, despite the fact that no other correlate of pauses is present. The final lengthenings have the same distribution as the hesitation sounds not preceded by silent intervals, i.e. they are most frequently found after discourse markers and conjunctions. This tendency to prolong the duration of discourse markers and conjunctions by inserting a hesitation sound or by final lengthening might be seen as another means of holding the floor when planning an upcoming sentence or clause.

Measurements of the duration of the pauses reveal that silent pauses before sentence boundaries and after discourse markers and conjunctions are longer than pauses before accented content words and pauses at phrase boundaries (pauses before adverbial phrases and phrases containing an accented content word). Pauses containing clicks and pauses composed of one or two silent intervals in combination with a hesitation sound constitute, not surprisingly, the longest pauses in the material. The hesitation sounds preceded and / or followed by silent intervals are often found before sentences or after discourse markers in sentences introducing a new speech paragraph, as in (4).

- (4) och så hänger dom undan den här (p) uniformen (p) och äh (p) ja vad händer sen (p) äh (p) jo (p) i alla fall (p) dom det händer nåt
'and then they put away this (p) uniform (p) and uh (p) yes what happens next (p) uh (p) right (p) anyway (p) they something happens'

F0 patterns With the object to assign structure to the stream of speech, pauses are inserted in the boundaries between coherent units or groups of words belonging together. However, pauses in spontaneous speech do not always appear in positions where we would expect to find a prosodically marked boundary. We expect boundaries to be perceived at the major transition points in the discourse, e.g. between sequences of words corresponding to clauses, sentences and paragraphs in written discourse [9], but we have seen that pausing also frequently occurs in other locations in spontaneous speech, e.g. after discourse markers and conjunctions. The question is then: do these pauses occur without breaking the coherence of the prosodic phrase or utterance, and if so, how are they manifested acoustically as compared to those pauses coinciding with prosodic phrase and utterance boundaries?

Perceptually, a distinction can be made between four different kinds of silent pauses in the examined material. One type of pause is associated with a F0 pattern which appears to signal coherence over the silent interval. The F0 curve before the pause is flat, in some cases rising slightly just before the silent interval. Pitch is perceived as anticipating a continuation. There is no drop in F0 before the pause to signal a boundary, nor is there a reset of the F0 after the pause. The F0 after the pause is on the average of 0.8 Hz lower than before the pause (std deviation: 11.6 Hz, n=118). Pauses associated with this F0 pattern are not perceived as boundary signals. They are found between clauses, before content words, between repeated words, and after discourse markers and conjunctions. The finding that pauses between clauses are not perceived as terminal boundary signals is consistent with results

from perception experiments undertaken by Van Donzel [9], which show that clause boundaries in spontaneous Dutch are most frequently associated with 'non-finality'. Figure 1 shows an example of a pause associated with the F0 pattern described above.

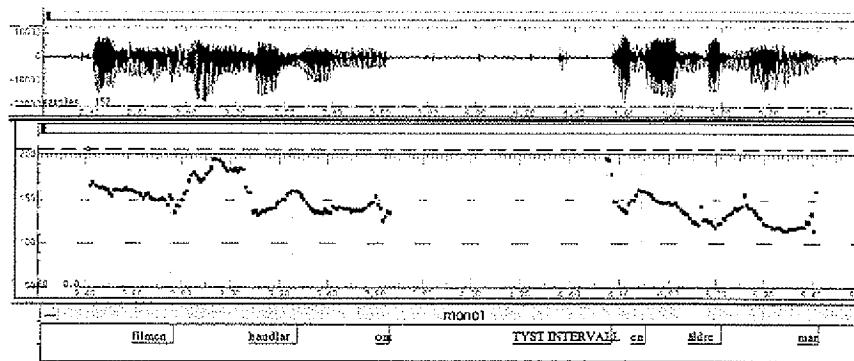


Figure 1: The utterance *filmen handlar om (p) en äldre man* 'the film is about (p) an older man'.

A second kind of pause in the material signals a weak boundary. The F0 curve before the pause drops in frequency to signal a boundary or remains flat and signals continuation, and after the pause there is a small reset of the F0 of 18.2 Hz (std deviation: 10.7 Hz, n=62). Pauses with F0 patterns signalling weak boundaries can be found before conjoined clauses and clause fragments, between sentences within the same speech paragraph, and when the speaker interrupts himself and starts over.

The third identified type of pause is fairly rare in the examined material. It constitutes a strong boundary signal and is found between speech paragraphs. The pause is preceded by a F0 fall and followed by an average reset of 38.0 Hz (std deviation: 18.4 Hz, n=9).

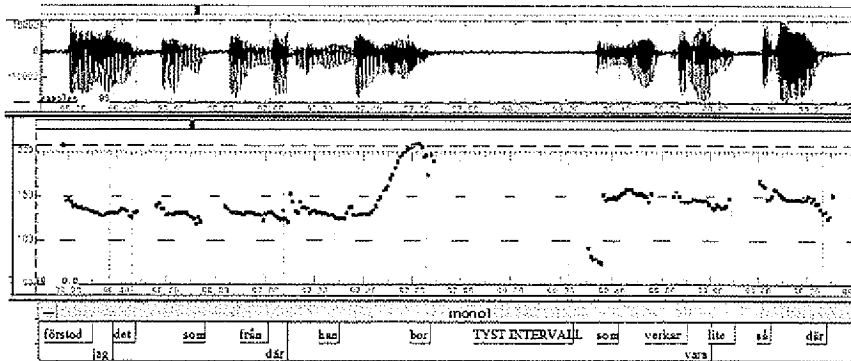


Figure 2: A fragment of the utterance *sen så (p) ser man nån (p) en (p) annan stadsvy liksom (p) som är (p) förstod jag det som från där han bor (p) som verkar vara lite så där (p) inte dom allra finaste delarna av stan* 'and then (p) you see some (p) a (p) another view of the city (p) which is (p) as I understood it from where he lives (p) that seems to be a bit (p) not the nicest parts of town'.

The fourth kind of pause noted in the monologue also functions as a terminal boundary signal. After a drop in F0 to a low value before the pause, the F0 curve rises rapidly and ends at a high value. After the pause, the F0 starts off at a lower value, but with a reset of the F0 if compared with a value measured before the final rise. The average reset of the F0 is 10.9 Hz (std deviation: 17.0 Hz, n=14) and the average rise is 50.0 Hz (std deviation: 28.7 Hz, n=14). The rise is, despite its resemblance to the

continuation rise often found in enumerations [1], not a signal of continuation and coherence over the pause, as in the case of the F0 pattern in figure 1. In the examined material the rise is found in clause final and sentence final positions and is perceived as a terminal boundary signal. Figure 2 shows an example of a terminal F0 rise in clause final position.

4. DISCUSSION

Pauses in spontaneous speech fill several functions. In addition to marking boundaries and providing the speaker with time to breathe, they also produce time for speech planning and choice of individual words. In the examined material, pauses occurred in mainly three positions: between sentences, after discourse markers and conjunctions, and before accented content words. While some pauses in spontaneous speech prosodically mark syntactic boundaries (e.g. the sentence boundary), it appears that others are situated in locations where the speaker can plan the upcoming sentence or clause without risking loss of the floor (after conjunctions and discourse markers). It also appears that speakers often pause when choosing an important word.

Silent intervals, hesitation sounds, final lengthening, glottalization and specific F0 patterns are some of the acoustic correlates of perceived pauses in spontaneous speech. The acoustic manifestation of a pause is to some extent dependent on its position and function. Pauses between sentences, presumably reflecting the planning of a sentence and / or the signalling of a boundary, and pauses after discourse markers and conjunctions, also assumed to reflect speech planning, have a longer duration than pauses preceding accented content words, reflecting the speaker's search process for an important word. Pauses which function to signal a boundary are associated with different F0 patterns than pauses which occur without breaking the coherence of the prosodic phrase or utterance. A pause is perceived as a boundary signal when preceded by a F0 fall or rise and / or followed by a F0 reset. A pause not associated with a movement or reset of the F0 will signal coherence over the pause. It occurs without breaking the coherence of the phrase or utterance and perhaps has the effect of avoiding interruptions during speech planning.

5. REFERENCES

1. Bruce, G., "Almän och svensk prosodi", *Praktisk lingvistik* 16, Department of Linguistics and Phonetics, Lund University, 1998.
2. Sanderman, A., *Prosodic phrasing, Production, perception, acceptability and comprehension*, Institute for Perception Research, Eindhoven University of Technology, 1996.
3. Goldman-Eisler, F., "Pauses, Clauses, Sentences", *Language and Speech* 15, 103-113, 1972.
4. Strangert, E., "Where Do Pauses Occur in Texts Read Aloud?", In Sigurdsson, A. H. (ed), *Papers from the Twelfth Scandinavian Conference of Linguistics, Reykjavik, June 14-16 1990*, 403-414. Linguistic Institute, University of Iceland, 1990.
5. Gårding, E., "Prosodiska drag i spontant och uppläst tal", in Holm, G. (ed), *Svenskt talspråk*, 40-85, Almqvist & Wiksell, Stockholm, 1967.
6. Bruce, G., Filipsson, M., Frid, J., Granström, B., Gustafson, K., Horne, M. and House, D., "Text-to-intonation in spontaneous Swedish", *Proceedings of Eurospeech 97, 5th European Conference on Speech Communication and Technology*, 215-218, Rhodes, Greece, 1997.
7. Horne, M., Bruce, G., Hansson, P. and Filipsson, M., "Boundary-signalling in discourse: the case of Swedish *men*, an ambiguous cue word", *Proceedings FONETIK 98, The Swedish Phonetics Conference, May 27-29 1998*, 138-141, Department of Linguistics, Stockholm University, 1998.
8. Strangert, E., "Speaking style and pausing", in Strangert, E., Heldner, M. and Czigler, P. (eds), *PHONUM 2*, 121-137. Department of Phonetics, University of Umeå, 1993.
9. Van Donzel, M. E., "Perception of discourse boundaries and prominence in spontaneous Dutch speech", in Svantesson, J.-O. (ed), *Working Papers* 46, 5-23, Department of Linguistics and Phonetics, Lund University, 1997.