

A Comparative Study of Pauses in Dialogues and Read Speech

†Sofia Gustafson-Čapková and ‡Beáta Megyesi

†Department of Linguistics
Computational Linguistics
Stockholm University, Sweden
sofia@ling.su.se

‡Centre for Speech Technology
Department of Speech, Music and Hearing
Royal Institute of Technology, Sweden
bea@speech.kth.se

Abstract

This study aims to investigate the length, frequency and position of various types of pauses in three different speaking styles: elicited spontaneous dialogues, professional reading and non-professional reading.

1. Introduction

In the last decade several studies have been done on the role and features of pausing. One reason is that pauses indicate groups of words in speech, i.e. prosodic phrases which highlight the information structure of the utterance [1], [2], [6], [9], [13].

Speakers in monologue discourse, for instance, typically vary the length and position of pauses on the basis of the information structure; Pauses occur between all topical units, and directly after the topic-introducing phrase or clause [13].

It has been shown that pausing varies not only within a speaking style but also across speaking styles. Spontaneous dialogues and the read version of the same text have been compared for Swedish in [1] and for English in [6]. These studies reported that the number and the distribution of pauses as well as the speech rate differs across the speaking styles.

In [10] and [11], the distribution and features of pauses in professional news announcement, non-professional news reading and monologues have been compared. The results show that spontaneous speech contains long and frequently occurring pauses, while professional announcing is characterized by shorter and fewer pauses. Non-professional announcing is placed in between those two polarities. The pauses occurs mainly in places relevant to the underlying message, e.g. at syntactic boundaries, and at semantically important words. However, pauses also occurs in other positions. In those cases there seems to be a preference for sites as e.g. in connection to conjunctions.

Spontaneous speech differs from read speech in that the speaker often incrementally plan the content of what to say during the utterance, a fact that cause more frequently occurring and longer pauses [2] and [3]. In news reading, on the other hand, the text is already present and the planning phase is excluded, why pauses can be assumed to be less frequent and shorter [4], [10].

2. Pauses in different speaking styles

The purpose of this study is to investigate the function of pauses from several perspectives in three different speaking styles: elicited spontaneous dialogues, and news read by both professional announcers of radio news and non-professional readers.

2.1. Data

The material consists of recordings of Swedish radio news [12] read by four professional readers. The same material was also read by four non-professional readers. In both conditions there were two male and two female speakers. The spontaneous speech material consists of recordings of two Swedish map task dialogues, each with two dialogue participants, totally one male and three female speakers. The materials consist of 920 words each.

2.2. Method

In order to investigate the position and the duration of pauses, the speech data was processed automatically by a pause detector. Silent intervals longer than or equal to 100 ms was used as acoustic correlate for pausing. In this study, filled pauses may include breathing and swallowing intervals if they appear in conjunction with silent intervals equal to or longer than 100 ms. The reason for including those sounds in the pauses is, that they are natural physical phenomena. Humming, on the other hand, are interpreted as conversational support, hence, they are not allowed inside pauses.

The pause data is analyzed with regard to the length, frequency, types and position of pauses and this is described in the next section.

3. Results

3.1. Length and frequency of pauses

Not surprisingly, there are differences in duration as well as in the ratio of pausing between the speaking styles.

The mean duration of the acoustic pause length, see Figure 1, is lowest in professional reading (271 ms), highest in non-professional reading (561 ms) followed by dialogue (538 ms). The reason for the short duration of the pauses in skilled

reading might be that professional readers are under time pressure which is not the case in the other two speaking styles.

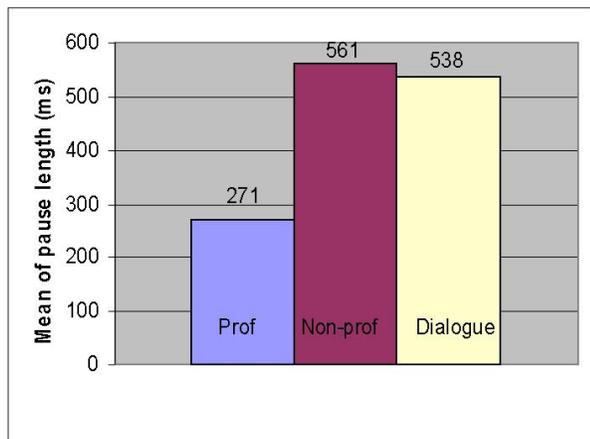


Figure 1: Mean acoustic pause length in professional and non-professional reading, and in dialogue.

Considering the frequency of pausing, the ratio of word per acoustic pause is highest in professional reading (77 words/pause), while non-professional reading (8.4 words/pause) gets a slightly higher rate than dialogue (5.5 words/pause), see Figure 2.

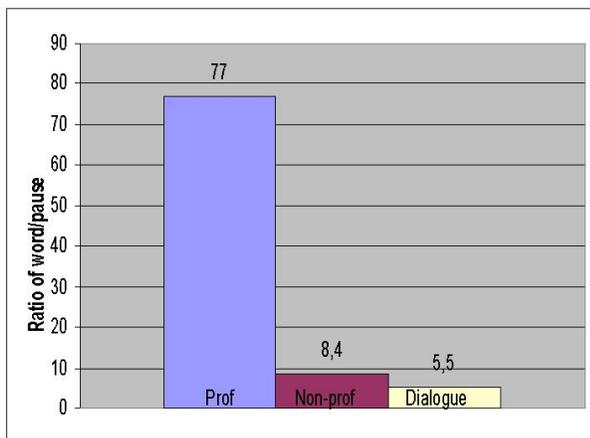


Figure 2: Ratio of word/pause in professional and non-professional reading, and in dialogue.

Although there are differences in the duration and frequency of pauses between the styles, the total length of the speech files is approximately the same for the reading styles: 372 seconds for professional and non-professional readings and 357 seconds for the dialogues. Hence, the time it takes to pronounce a word in average differs between the speaking styles: The average length of a word in professional announcing is 405 ms, in non-professional 337 ms and in dialogues 292 ms.

Thus, the non-professional readers in our study read as fast as professionals but the reading technique differs in the styles. Non-professionals use longer pauses and higher amount of reductions while professionals articulate clearly without the usage of silent sequences. In dialogues, the amount of reductions and

pauses are highest but the pauses have shorter duration than in non-professional reading.

3.2. Types of pauses

When examining pauses, we can distinguish between different types such as silent pause, and complex pause with breathing and/or swallowing. Silent intervals in stops, i.e. closures, were not included in the study. The usage of the types differs across the speaking styles, as it is shown in Table 1.

Across the speaking styles, silent pauses are most frequent in dialogues, while complex pauses are frequently occurring in professional reading, see Figure 3.

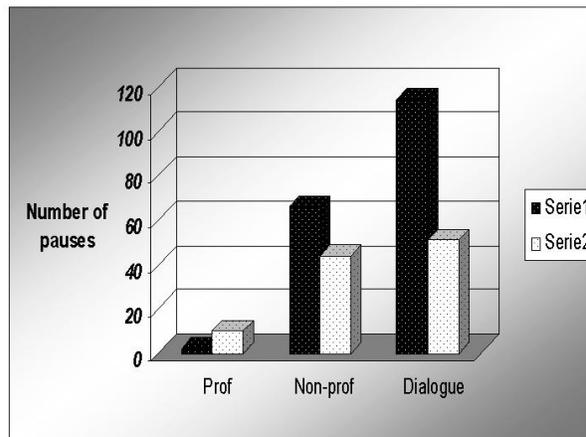


Figure 3: The amount of silent and complex pauses, illustrated as serie 1 and serie 2 respectively, in professional and non-professional reading and in dialogue.

There are also interesting differences between the types of pause used within each style. In dialogues, 69% of the pauses are silent, and 31% are complex. Similar results can be found in non-professional reading where 60% of pauses are silent and the rest of the pauses are complex. In professional reading, on the other hand, complex pauses are the most commonly occurring type; 83% of the pauses are complex.

Table 1: The amount of pauses and the duration (in seconds) of different pause types in the different speaking styles are given.

PAUSES	Professional		Non-professional		Dialogue	
	NO	SEC	NO	SEC	NO	SEC
PAUSE	2	0.67	66	29.49	114	45.63
COMPLEX	10	2.58	44	32.21	51	43.08
TOTAL	12	3.25	110	61.70	165	88.71

3.3. Position of pauses

In order to examine the position of the pauses, the texts have been analyzed according to their phrase structure, such as noun phrase, adjective phrase, adverb phrase, etc. Also, each clause was annotated with discourse labels: theme continuation (TC) or theme shift (TS). We use the notion of theme in the same sense as Vonk, et al. [14]. Theme continuation means that the attention is still centered around the same topic of discourse, even though new information may be added about the topic. Theme

shift means that the attention is directed towards a new topic of discourse or a new aspect of the topic already under consideration. The discourse markers do not appear inside phrases, but only in connection to phrase, clause, sentence and turn boundaries.

In addition to the notion of theme, the notion of turn was used in the analysis. Turn is used in the dialogue to mark up the points for turn-taking which is a fundamental phenomenon in dialogues and can be defined as speaker switch from speaker A to speaker B in the conversation.

Furthermore, each turn and theme was classified according to the type of their following constituent: phrase, clause or sentence.

In cases where a pause appears inside a phrase, the PoS of the word was marked as well as whether the word is a phrasal head or not.

In case of hesitations attached to a word we have annotated the hesitation, but classified the position according to the constituent following the hesitation. The reason for doing this was to keep the more detailed information, and not just incorporate it in a super-ordinate category of hesitation.

To give a picture of the granularity of the annotation, consider the following example taken from a turn in the dialogue.

“...Då ska vi se då har vi en
(PAUSE/DETERMINER-NOT_HEAD) en
så karta här framför oss och jag har
(PAUSE/HESITATION-HEAD) eeh landstigit
(PAUSE/TC-PHRASE_BOUNDARY) på en plats
(PAUSE/TC-PHRASE_BOUNDARY) på den här
ön...”

“...Lets take a look here we have a
(PAUSE/DETERMINER-NOT_HEAD) a so
map here in front of us and I have
(PAUSE/HESITATION-HEAD) eeh arrived
(PAUSE/TC-PHRASE_BOUNDARY) to a place
(PAUSE/TC-PHRASE_BOUNDARY) on this
island...”

The results show that different types of pauses are to a certain extent favored in different positions.

3.3.1. Silent pauses

Since the large deviation in the results between the speaking styles, it is not possible to give account for the results in terms of the position with regard to the categories across the speaking styles. Thus, we will describe the main features for each style.

Our results show that in professional reading, silent pauses are rare (17%) but when they occur they are always in connection to sentence boundaries and theme continuation.

In non-professional reading, silent pauses are also connected to theme continuation (65.2%) but they can be found primarily at phrase boundaries, and secondarily at clause and sentence boundaries. 34.8 % of the silent pauses occurs in connection to theme shift, mainly at sentence boundaries.

In dialogues, pure pausing is highly frequent between turns; 37% of the silent pauses belong to that category. In connection to turn boundaries (i.e. where turn taking occurs) conversational particles, such as *mm*, *okay* and *aha*, are most common; 48.8% of the silent pauses by turn boundaries appear before conversational particles.

Inside turns, theme continuation is more frequently associated with silent pauses than theme shift (23% and 12%, respectively). Theme shift occurs in the first place at sentence

boundaries, but also at phrase and clause boundaries. Theme continuation, on the other hand, is common mostly at phrase boundaries, in the second place at clause boundaries and lastly at sentence boundaries.

Furthermore, in dialogue silent pauses also occur in conjunction to head nouns and adverbs which is not the case in either of the reading styles.

3.3.2. Complex pauses

In professional news announcing 70% of complex pauses, i.e. pauses with breathing, swallowing, etc, can be found before a theme shift at the beginning of a sentence. The remaining part is found at theme continuation; mostly at sentence boundaries but also between noun phrases in a list.

In non-professional reading, 61% of complex pauses correlates with theme continuation at sentence and clause boundaries, and also, but to a lesser extent in connection to noun phrases. Furthermore, 38% of pauses occur in theme shifts at sentence boundaries.

In dialogues, the distribution of turns, theme shift and theme continuation in connection to pauses are relatively even: 28%, 22% and 26%, respectively. The rest of pauses can be found in phrasal heads (8%): nouns, or adverbs preceded by hesitation particles. The remaining part (16%) are pauses in connection to overlapping speech, conversational particles, hesitations, and others.

At theme shift, the duration of pauses are longer compared to theme continuation which may indicate that the speaker is signalling a new theme.

4. Discussion

As shown in the results, the length and frequency of pauses differed greatly across the speaking styles in our material. They form a continuum with dialogue on one end, professional announcing on the other end, and non-professional reading in between.

Spontaneous speech has the most frequently occurring pausing with longest durations, a fact that might partly be explained by the planning component involved in spontaneous speech. The speaker has to plan incrementally what to say. This is not the case in reading where the text is already present. In professional announcing, the reading is furthermore performed under a time constraint, which may cause few and short pauses. Non-professional readers are not limited by time constraint and do not have to plan the content, but rather how to execute the message.

The compression of speech is however not present when examining the duration of words; the more thorough articulation in professional reading may be explained by the need of clarity in speech since no repair is possible. In dialogues on the other hand, the speaker is free to correct him/herself and also utilize the advantage of feedback.

The results show that certain types of pauses tend to be favoured in certain positions.

Complex pauses have a multi-functional use in both allowing the reader/speaker to breath, but also at the same time shape the discourse. They tend to appear in connection to theme shift as a way for the speaker to mark the introduction of a new topic or a new aspect of a message.

Silent pauses, on the other hand, are found before theme continuation. They are shorter in general, marking additional information about the same topic.

We have seen that the pattern of pausing differs across speaking styles. In some cases, it is a result of a speech situation per se, e.g. there are no turns in read speech. In other cases it might be a result of the genre within a particular type of speaking style (such as reading and spontaneous dialogue). We assume that for example fairy-tale reading differs from news announcing just as spontaneous conversation from elicited task oriented dialogue.

Further research on pauses should include the study of the variations within different speaking styles. Also, the perceptual aspect of pauses – how pauses are perceived by listeners – should be investigated; More specifically the issue on the type and position of pauses that are actually perceived by listeners.

5. Conclusion

Three varieties of speech has been investigated according to type, frequency, length and position of pauses. The different speaking styles show varying patterns in pausing.

Professional announcers has few and short pauses with a multi-functional usage combining pausing with breathing, swallowing, etc. Pauses appear mainly at sentence boundaries where the theme of the message changes. Non-professional readers use longer and more frequently occurring pauses, that in most cases are silent. The position of pauses varies to a larger degree than in professional reading. We find pauses at sentence, clause and phrase boundaries as well as in head words. In dialogues the frequently occurring pauses with long durations are mostly silent, similarly to non-professional reading. However, the pauses occur in a more varying context (such as turn takings, in connection to conversational particles, modifiers, etc) compared to the reading styles.

6. Acknowledgements

We would like to express our great thanks to those who contributed to our study: to Petur Helgasson for the dialogue corpus, to Swedish Radio for the Swedish news recordings, to Mattias Heldner for his help with the recordings of the non-professional readings, and to Harald Berthelsen and Jens Edlund for the pause detector. Last but not least, we would like to thank Rolf Carlson and Kjell Gustafson for their valuable suggestions and helpful comments.

7. References

- [1] Bruce, G., “Modelling Swedish Intonation for Read and Spontaneous Speech”, *Proceedings of International Congress on Phonetic Sciences*, Vol. 2 pp. 28-35, 1995.
- [2] Deese, J., “Pauses, prosody and the demands of production in language”, *Temporal Variables in Speech, Studies in Honour of Frieda Goldman-Eisler*, Hans & Raupach, Manfred (eds.), Mouton Publishers, 1980.
- [3] Goldman-Eisler, F., “Pauses, Clauses, Sentences”, *Language and Speech*, 15:2, 1972.
- [4] Garman, M., “*Psycholinguistics*”, Cambridge University Press, 1990.
- [5] Helgason, P., “*Stockholm Corpus of Spontaneous Speech*”, Department of Linguistics, Stockholm University.
- [6] Hirschberg, J., “Prosodic and other acoustic cues to speaking style in spontaneous and read speech”, *Proceedings of International Congress on Phonetic Sciences*, Vol. 2, pp. 36-43, 1995.

- [7] Hirschberg, J., “Prosodic variation and discourse structure across speaking styles”, *Prosody: Theory and Experiment, Studies presented to Gösta Bruce*, Kluwer Academic Publisher, 1997.
- [8] Hirschberg, J., “Communication and Prosody: Functional Aspects of Prosody”, *Speech Communication: Special Issue on Dialogue and Prosody*, Terken, J., & Swerts, M. (eds.), 2001.
- [9] Ostendorf, M., “Prosodic Boundary Detection” *Prosody: Theory and Experiment, Studies presented to Gösta Bruce*, Kluwer Academic Publisher, 1997.
- [10] Strangert, E., “Speaking style and pausing”, *PHONUM*, Reports from the Department of Phonetics, University of Umeå, 1993.
- [11] Strangert, E., “Clause Structure and Prosodic Segmentation”, *FONETIK-93 Papers from the 7th Swedish Phonetics Conference*, John Sören Petterson (ed), Uppsala, May 12-14, 1993.
- [12] Recordings of Swedish Radio News, Swedish Radio, 1999-2000.
- [13] Swerts, M. & Geluykens, R., “Prosody as a marker of information flow in spoken discourse”, *Language and Speech* 37, 21-45, 1994.
- [14] Vonk, W., Hustinx, L. & Simons, W., “The Use of Referential Expressions in Structuring Discourse”, *Language and Cognitive Processes* 7, 3/4, 301-333, 1992.