Pausing strategies with regard to speech style

Dorottya Gyarmathy and Viktória Horváth

Department of Phonetics, Research Institute for Linguistics,

Hungarian Academy of Sciences, Budapest, Hungary

Abstract

Speech is occasionally interrupted by silent and filled pauses of various length. Pauses have many different functions in spontaneous speech (e.g. breathing, marking syntactic boundaries as well as speech planning difficulties, time for self-repair). The aim of the study was the analysis of the interrelation between the temporal pattern and the syntactical position of silent pauses (SP) on one hand. On the other hand, filled pauses (FP) were also analyzed according to their phonetic realization, as well as the combination of SPs and FPs. The effect of speech style on pausing strategies was also analyzed. A narrative recording and a conversational recording from 10 speakers (ages between 20 and 35 years, 5 male, 5 female) were selected from Hungarian Spontaneous Speech Database for the study. The material was manually annotated, silent pauses were categorized, then the duration of pauses were extracted. Results showed that the position of silent and filled pauses affects their duration. The speech style did not influenced the frequency of pauses. However, silent and filled pauses were longer in narratives than in conversations. Results suggest that pausing strategies are similar in general; however, the timing patterns of pauses may depend on various factors, e.g. speech style.

Introduction

Pauses serve various functions in speech, like breathing, grammatical function, providing time for speech planning processes and for perception as well (Levelt, 1989, Gósy, 2000). The realization of pauses depends on various factors, e.g. the speaker's age, the length and the complexity of the utterance or the speech style (Duez, 1982, Krivokapic, 2007). Researches revealed connection between the speech situation and the pauses. The more complex a speech task was—the greater cognitive effort it required—the longer and more frequent the pauses became (Goldman-Eisler, 1968; Kowal et al., 1975). Silent pauses were longer and more frequent in political speech, the longest pauses having a stylistic function. Filled pauses were not characteristic for this type of speech, whereas they were decidedly frequent in interview situations (Duez, 1982). Connection was found between the position and the duration of pauses, for example in the case of 'to+infinitive'

grammatical structures. There were significantly longer pauses before 'to' than after it during reading aloud; whereas the opposite was found in spontaneous speech, probably due to speech planning characteristics (Bada & Genç, 2008).

The effect of speech style on pausing were analyzed in Hungarian as well. Researches revealed differences in pausing strategies between spontaneous speech and reading aloud (Olaszy, 2005; Váradi, 2010). The ratio of pauses was less in conversations than in narratives, in addition the duration of pauses was shorter in conversations than in narratives (Markó, 2005). Re-telling a story was the most difficult speech task for young adults; therefore, the speakers produced pauses the most frequently in this task. The pauses realized with longer duration in re-telling a story than in conversations (Bóna, 2013).

Silent and filled pauses have several additional functions in conversations (Sacks et al. 1974). 'Pause' is defined as a signal break within a speech turn (we analyzed this type of silent pauses in the present study). Furthermore, pauses occur in conversations for thinking or for dramatic effect, the speaker can use them to highlight new information, and they can also be used to organize the discourse (Esposito et al., 2007).

The aim of the study is to analyze the occurrence and duration of silent pauses according to their position in conversations and narratives, on one hand. On the other hand, the realization of filled pauses and their combination with silent pauses were also analyzed. Our hypotheses were that (i) silent and filled pauses realize with different patterns according to speech style; (ii) the duration of silent and filled pauses is determined by their position.

Method and material

10 conversations and 10 narratives were selected for the study from a Hungarian Database called BEA (Neuberger et al., 2014). Three speakers participated in each conversation; the interviewer (Int) and one speaker (henceforth: the second speaker S2) were constant. S1 speaker was the third participant in each conversation. The S1 speaker was asked to tell his or her opinion on a given topic by the interviewer in the narrative sessions. The S1 participants were between 20 and 35 years old, half of them were male and half of them were female. Both the Int and S2 were 28 years old.

The total material was 175 minutes long (conversations altogether 131 minutes, *mean* = 13 min., range: 6.9–23.3 min.; while narratives altogether 44.5 minutes, *mean* = 4.45 min, range: 1.7–10.2 min). The annotation was carried out in Praat (Boersma & Weenink, 2018). The speech intervals, the silent and filled pauses were annotated, labelling the phonetic form of filled pauses as well. The duration of silent and filled pauses were automatically extracted.

1853 silent pauses occurred in the total material, 1185 in the conversations, while 668 in the narratives. Silent pauses were categorized based on the system developed by Gyarmathy (2018). The first distinction was whether the pause was related to disfluency (in these cases, the time span between the interruption of articulation and the beginning of correction was taken into account, as part of the editing phase – E), or it had a syntactical function (S) (Figure 1). Pauses with editing function (E) were further categorized based on whether the disfluency phenomena were due to the speaker's uncertainty or errors. Silent pauses with a syntactical function (S) were distinguished based on their position. Utterance onset pauses (S_Uo) occur when a speaker claims the turn; here the pause may only be preceded by a filler word or a discourse marker. Silent pauses at phrase boundaries (S_PhrB) are found between clauses of virtual sentences, often before or after a conjunction. Within phrase pauses (S_PhrW) are found within a grammatical unit ('phrase'). End of phrase pauses (S_PhrE) are silent pauses at the end of a virtual sentence, after which the speaker starts another virtual sentence that often represents a new thought unit. The frequency and the duration of silent pauses was also analyzed with regard to these categories. 555 filled pause occurred in the total material. The phonetic form, the frequency, the position and the duration of filled pauses were also analyzed, as well as the occurrences of FPs combined with silent pauses. Statistical analysis was conducted using SPSS 20.0 (GLM, GLMM).
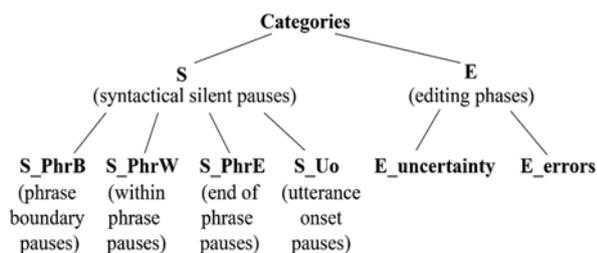

Figure 1. The categories of intra-speaker silent pauses.

Results

Results showed that the silent pause was the most frequent type, irrespectively of the speech style (5.8 items per 100 syllables in conversations on average, while 5.6 in narratives, see Figure 2).
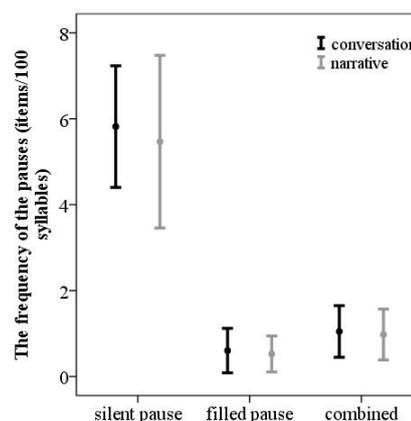

Figure 2. The frequency of pauses according to speech style.

The occurrence of combined pauses was about 1 item per 100 syllables on average, while the mean frequency of filled pauses was 0.5 item, irrespectively of the speech style.

The duration of silent pauses (Figure 3) were significantly longer in the narratives (523 ms on average) than in the conversations (466 ms on average) [GLMM: $F(1, 1851) = 10.057$ $p = 0.002$ pairwise: $t = 3.171$].
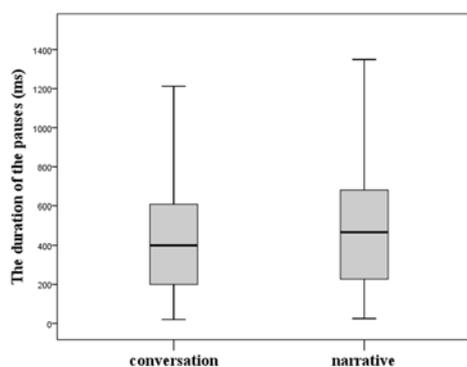

Figure 3. The duration of the silent pauses according to the speech style.

The occurrence and the duration of SPs were also analyzed according to the position. The most of the silent pauses were syntactical pauses (S) with the occurrences of 5 items per 100 syllables in the conversations and in the narratives as well. Pauses in editing phases (E) of disfluencies were considerably less frequent: 0.7 items per 100 syllables in conversations, while 0.5 items in narratives. The ratio of the subcategories of syntactical silent pauses were similar in both of the speech styles. Pauses occurred at phrase boundaries most frequently (S_PhrB): 2.6 items per 100 syllables in conversations and in narratives as well. 1.3 pauses

occurred per 100 syllables within phrases (S_PhrW)—the frequencies were the same in both of the speech styles. Pauses at the utterance onset (S_Uo) occurred the least frequently (only 0.05 items per 100 syllables in conversations and 0.04 in narratives), irrespectively of speech style. Silent pauses in editing phases occurred somewhat more frequently in the conversations: 0.58 items connected with uncertainty and 0.16 items connected with errors per 100 syllables (narratives: 0.46 items of E_unc and 0.09 items of E_errors).

We also analyzed the duration of pauses according to the subcategories (Figure 4.). The longest silent pauses were the S_PhrE and the S_Uo types, irrespectively of speech style. The S_Uo pauses were longer in conversations (697 ms, $SD = 999$ ms) than in narratives (500 ms, SD = 647 ms). In contrast, pauses at the end of the phrases (S_PhrE) were longer in narratives (754 ms, $SD = 479$ ms) than in conversations (677 ms, $SD = 430$ ms). The S_PhrB pauses were longer than S_PhrW pauses in conversations and in narratives as well (conversations: S_PhrB: 449 ms, $SD = 329$ ms; S_PhrW: 372 ms, $SD = 259$ ms; narratives: S_PhrB: 529 ms, $SD = 366$ ms; S_PhrW: 367 ms, $SD = 275$ ms). The pauses of editing phases connected to uncertainty phenomena (S_unc) realized with longer durations than pauses of editing phases connected to errors, irrespectively of speech style (S_unc in conversations: 388 ms, $SD = 324$ ms; in narratives: 418 ms, $SD = 340$ ms; E_error in conversations: 306 ms, $SD = 314$ ms; in narratives: 159 ms, $SD = 170$ ms). The statistical analysis revealed that the subcategories of silent pauses determine their duration in conversations [GLMM: $F(5, 1834) = 24.794$, $p < 0.001$] and in narratives [GLMM: $F(5, 1834) = 22.496$, $p < 0.001$] as well (Table 1 contains the results of pairwise contrasts).

The filled pauses realized with various phonetic forms. They consisted either of one speech sound or of two or three speech sounds. The single speech
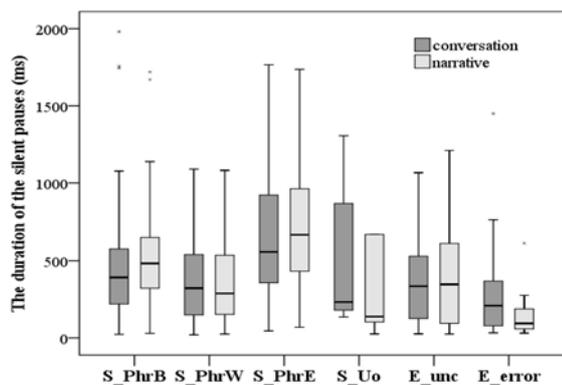

Figure 4. The duration of the silent pauses according to subcategories and speech style.

Table 1. The results of the pairwise contrast according to speech style.

speech style	subcategories of silent pauses	*t*-value	significance contrast
conversation	S_PhrB – S_PhrW	3.354	0.001
	S_PhrB – S_PhrE	8.108	0.000
	S_PhrB – E_unc	2.734	0.006
	S_PhrB – E_error	2.455	0.014
	S_PhrW – S_PhrE	9.889	0.000
	S_PhrE – E_unc	8.126	0.000
	S_PhrE – E_error	5.750	0.000
narrative	S_PhrB – S_PhrW	4.596	0.000
	S_PhrB – S_PhrE	6.375	0.000
	S_PhrB – E_unc	2.767	0.006
	S_PhrB – E_error	3.021	0.003
	S_PhrW – S_PhrE	9.391	0.000
	S_PhrE – S_Uo	3.039	0.002
	S_PhrE – E_unc	6.658	0.000
	S_PhrE – E_error	5.018	0.000

sound was a neutral vowel or a bilabial nasal-like consonant. Filled pauses consisting of more speech sounds were combinations of schwa and nasal with each other or with laryngeal consonant. The schwa was the most frequent form in eight speakers' speech (72% on average, 55–100%) in conversations, while the ratio was 72% on average in narratives (51–100%). We analyzed the combined occurrences of silent and filled pauses. Filled pauses occurred between two words (without any silent pauses) in a similar ratio than preceding a silent pause, irrespective of the speech style (Figure 5). The less frequent case was when filled pauses occurred between two silent pauses.
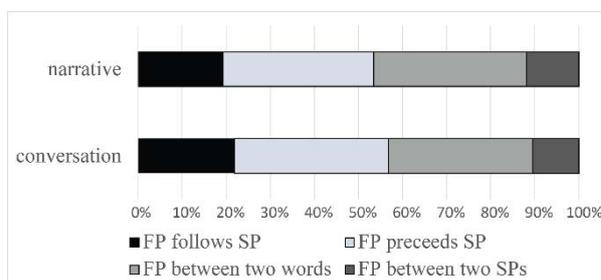

Figure 5. The combined occurrences of silent and filled pauses.

The duration of the most frequent filled pause (schwa) was analyzed with regard to its position and the speech style. The position of schwa had statistically significant effect on its duration [GLM: $F(1, 277) = 19.678$ $p = 0.001$ $\eta^2 = 0.664$]. Filled pauses were significantly longer between two silent pauses (*mean* = 434 ms, *SD* = 211 ms) than the filled pauses adhered to word(s) (*mean* = 296 ms, *SD* = 141 ms). The speech style also influenced the

duration of schwa-like FPs [GLM: $F(1, 277) = 5.322$ $p = 0.025$ $\eta^2 = 0.099$]. The duration of filled pauses were significantly longer (*mean* = 343 ms, *SD* = 154 ms) in the narratives than in the conversations (*mean* = 295 ms, *SD* = 155 ms).

Conclusions

Silent and filled pauses were analyzed in the study, according to the speech style. Our first hypothesis were partly confirmed: the frequency of SPs (and its subcategories as well) and FPs was similar in conversations and narratives of the certain speakers. However, SPs and FPs were significantly longer in narratives than in conversations, similar to earlier studies (cf. Markó, 2005). The task of speaking is easier in conversation than in narratives, due to the fact that partners help each other in managing the conversations on one hand. On the other hand, during the partner's speech, the following speaker has time for speech planning processes. Silent pauses appeared more often in grammatically functional positions in total (S_PhrB, S_PhrE, S_Uo) than within a phrase, irrespectively of speech style. Speakers usually do not create a break within a utterance; this indicates they not only plan the content and form of the utterance, but also the pauses (cf. Zellner, 1994). Within-phrase pauses can be a sign of a major speech planning problem. Data confirmed our hypothesis: the position of SPs determined their duration: pauses at boundaries were longer than pauses within a phrase. S_PhrBs were shorter in conversations than in narratives—their shorter duration may mark the speaker's intention of keeping the floor. The position of FPs influenced their realization. FPs occurred the least frequently between two silent pauses; however, with the longest durations in this position – they presumably indicate problems in planning processes.

Acknowledgements

References

Bada, E. & B. Genç. 2008. Pausing preceding and following *to* in *to*-infinitives: A study with implications to reading and speaking skills in ELT. *Journal of Pragmatics* 40(11): 1939–1949. https://doi.org/10.1016/j.pragma.2008.03.010

Boersma, P. & D. Weenink. 2018. *Praat: Doing phonetics by computer* (version 6.0.19). http://www.praat.org/ (accessed 24 May 2019).

Bóna J. 2013. A beszédszünetek fonetikai sajátosságai a beszédtípus függvényében [Phonetic features of pauses depending on the type of speech]. *Beszédkutatás* 21: 60–75.

Duez, D. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25(1): 11–25. https://doi.org/10.1177/002383098202500102

Esposito, A., V. Stejskal, Z. Smékal. & N. Bourbakis. 2007. The significance of empty speech pauses: Cognitive and algorithmic issues. In: F. Mele, G. Ramella, S. Santillo & F. Ventriglia (eds.), *Advances in Brain, Vision, and Artificial Intelligence*, 542–554. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-540-75555-5_52

Goldman-Eisler, F. 1968. Psycholinguistics: *Experiments in spontaneous speech*. London: Academic Press.

Gósy, M. 2000. A beszédszünetek kettős funkciója [The double function of pauses in speech]. *Beszédkutatás* 2000: 1–14.

Gyarmathy, D. 2018. The functions of silent pauses in spontaneous Hungarian speech. *The Phonetician* 115: 53–71.

Kowal, S., D. C. O'Connell & E. J. Sabin. 1975. Development of temporal patterning and vocal hesitations. *Journal of Psycholinguistic Research* 4(3): 195–207. https://doi.org/10.1007/BF01066926

Krivokapic, J. 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics* 35(2): 162–179. https://doi.org/10.1016/j.wocn.2006.04.001

Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. Cambridge (Massachusetts)–London (England): The MIT Press.

Markó, A. 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége* [Some suprasegmental features of spontaneous speech]. Ph.D. dissertation, ELTE, Budapest.

Neuberger, T., D. Gyarmathy, T. E. Gráczi, V. Horváth, M. Gósy & A. Beke. 2014. Development of a large spontaneous speech database of agglutinative Hungarian language. In: P. Sojka, A. Horák, I. Kopeček & K. Pala (eds.), *Proceedings of TSD 2014 Text, Speech and Dialogue, the 17th International Conference*, 424–431. https://doi.org/10.1007/978-3-319-10816-2_51

Olaszy, G. 2005. Prozódiai szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában [Prosodic features in news items, storytelling, short stories and commercials]. *Beszédkutatás* 2005: 21–50.

Sacks, H., E. A. Schegloff & G. Jefferson. 1974. A simplest systematics for the organization of turn taking for conversation. *Language* 50(4/1): 696–735. https://doi.org/10.1353/lan.1974.0010

Váradi, V. 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása [Comparing the temporal features of reading aloud and spontaneous speech]. *Beszédkutatás* 2010: 100–109.

Zellner, B. 1994. Pauses and the temporal structure of speech. In: E., Keller. E. (ed.), *Fundamentals of speech synthesis and speech recognition*, 41–62 Chichester: John Wiley.