THE FUNCTIONS OF SILENT PAUSES IN SPONTANEOUS HUNGARIAN SPEECH

Dorottya Gyarmathy
Research Institute for Linguistics,
Hungarian Academy of Sciences
gyarmathy.dorottya@nytud.mta.hu

Abstract

Speech is occasionally interrupted by pauses of various lengths that are an essential part of human speech production. It is currently a matter of debate what exactly can be considered to be a pause, what the diverse forms are in which pauses can occur in speech, what their acceptable minimum and maximum durations are, and what kinds of functions they can have in speech.

This study focuses on the analysis of the temporal structure of silent pauses in Hungarian spontaneous speech. We hypothesized that (1) silent pauses differ from each other in terms of their functions, (2) the various types of silent pauses exhibit different patterns of frequency and duration, and (3) the duration of silent pauses is partially determined by their syntactic positions.

We differentiated silent pauses (S) occurring either in phrases or at phrase boundaries from pauses functioning as editing phases (E). The former type of pauses will be called 'syntactical silent pauses' while the latter will be referred to as 'editing phase silent pauses'. We will subcategorize syntactical silent pauses as follows: (1) phrase boundary pauses, (2) within phrase pauses, (3) end of phrase pauses, and (4) utterance onset pauses. The subcategories of the editing phases depend on the type of disfluency surrounding the silent pause at hand.

Our results confirmed all three hypotheses. The temporal data supported the claim that the function of the silent pauses directly affected their frequency of occurrence and their duration.

Keywords: spontaneous speech, silent pauses, editing phrase, syntactical silent pauses, function of silent pauses

1 Introduction

Pauses, in an everyday sense, can simply be seen as silent portions of speech. Within a linguistic perspective, however, we are faced with a problem that is far more complicated. It is not quite clear what in fact can be considered to be a pause, how many types of pauses occur in speech, what its minimal and maximal duration may be, and what functions it can serve. During the first half of the twentieth century, the study of speech pauses focused largely on their use in rhetoric, or public speeches, both in the Hungarian and the international literature. The focus of analysis was the connection between the system of punctuation in written texts and the way it was rendered in speech (Mátray, 1861; Hevesi, 1908; Simonyi, 1903; Lindroth, 1933). The pause as a unit of the language system was first mentioned in the phonetics literature, as far as we know, by Sweet (1877, 1890), who approached the problem from a physiological angle. He related pauses to breathing, and used the term breath group for the portion of speech uttered within a single run of exhalation. This idea was further pursued with respect to Hungarian by Balassa (1886). A number of other scholars of the period (e.g., Viëtor, 1894; Jespersen, 1904) also connected silent pauses with the physiological necessity of breathing. Jones (1922), on the other hand, offered an account of the relationship between silent pauses and breathing that was quite different from previous ideas. He made a distinction between pauses that served the purposes of breathing and pauses that were meant to make the message clearer. He furthermore noted, with respect to the first type, that speakers normally paused for breathing at points in their speech flow where the meaning of what they were saying made it necessary or at least acceptable. In sum, scholars of the period interpreted pausing in terms of two things: the physiological necessity of breathing and the punctuation of written genres of language use (also Weiske, 1838; Bieling, 1880).

Within the Hungarian literature of phonetics, it was Hegedűs (1953) who first published a comprehensive paper based on empirical research that discussed pauses. Here, he claimed that it was evident that pauses were connected with breathing, but he also emphasised that breathing did not primarily serve a biological function in the speech process. Rather, it was subordinated to the way thoughts were conveyed. During evolution, the physiological necessity of breathing had been linked with a high level conscious form of activity, human speech. Based on his analyses of spontaneous speech and reading aloud, he claimed that the occurrence of pauses in speech could not be explained simply by biological needs, contrary to what previous phoneticians had thought.

The fact that the phenomenon of pauses was only tackled in detail from the mid-twentieth century onwards can be explained in two ways: 1) by the lack of technological equipment available for researchers, making it impossible for them to carry out objective measurements until then and 2) the fact that pauses were not taken to be part of the language system (Sallai & Szende, 1995) and instead were viewed as discontinuities in speaker activity. The literature in the second half of the twentieth century continued to offer a number of contradictions and classification problems. Descriptions of pauses and their functions were riddled by a misunderstanding of aspects of production and perception, as well as acoustic-phonetic parameters (Fónagy, 1967; Sallai & Szende, 1975; Szende, 1979; Váradi, 1988). An attempt to provide a comprehensive definition was made by Sallai and Szende (1995). Their definition embraced both silent and "filled" pauses, as well as pauses whose duration was zero, and cases of what they called "compensations for pauses" (like vowel lengthening preceding a silent pause). In the broader theoretical perspective they advocated, a pause was to be understood to be a break in the sequence (or serial structure) that carried or constituted information. It was not unambiguously clear, however, what exactly should be taken to be a pause. Indeed, speech contains instances where the signal breaks for a while as part of the production of certain speech sounds (as in the closure periods of voiceless plosives or voiceless affricates). This type of “no signal” period could not be taken to be pauses (Gósy, 2004). On the other hand, some authors identified various disfluency phenomena (like repetitions or false starts) as pauses, at least in English (Mahl, 1956).

People may stop speaking briefly for a variety reasons. One of these reasons can be simple physiological necessity (breathing), but pauses can also serve to organize speech into meaningful units (Esposito et al., 2007). Pause can provide time for the speaker to think of what to say next or for the listener to take in the effect of what has been said. It can signal the coming of new information or have a role in discourse organisation (Esposito et al., 2007). The literature lists several additional functions of pauses which appear to be dependent on the research paradigm utilized (Zellner, 1994).

The definition of pause currently agreed upon with respect to Hungarian tries to resolve the contradictions in the literature as presented above and takes both production and perception criteria into account. In the present research we use the definition by Gósy (2000:2):

> “A pause is a partially voluntary break in the speech flow that is either silent or filled with signal but that is independent of the articulation of any individual speech sound. Functionally, in speech production, it
> (i) enables the speaker to keep up the flow of breath necessary for articulation;
> (ii) helps to organize the message into units;
> (iii) serves the resolution of contradictions, false paths etc. in speech planning; and
> (iv) gives the speaker time for searching the mental lexicon and for modifying linguistic encoding while speaking.
>
> In speech comprehension, it
> (i) facilitates the processing of what has been heard;
> (ii) reduces entropy; and
> (iii) ensures the operation of the processes of comprehension and interpretation.”

Silent pauses constitute the most frequently occurring phenomenon in spontaneous speech, as confirmed by a number of studies both in Hungary and internationally (Boomer, 1965; Goldman-Eisler, 1958; Hargreaves & Starkweather, 1959; Levin et al., 1967; Tannenbaum et al., 1967; Verzeano & Finesinger, 1949; Misono & Kiritani, 1990; Gósy, 2000, 2003a; Menyhárt, 2003; Markó, 2005a; Bóna, 2007, 2013a; Neuberger, 2014). Their ratio to full speaking time is in general 20 to 30%, but their duration and frequency of occurrence are determined by a number of factors. These factors include the speaker’s general properties (Markó, 2005a; Gósy et al., 2011); her actual physical state, e.g. having consumed alcohol (Gyarmathy, 2007); factors of the speech environment, e.g. the effect of noise (Gyarmathy, 2008); the speaker's speaking skills, the speech situation, the topic (Markó, 2005b); the style of speech (Duez, 1982); the genre of speech (Imre, 2005; Olaszy, 2005); the register (Markó, 2005a; Váradi, 2010; Bóna, 2013b); the speaker's age (Laczkó, 2009; Bóna, 2010, 2012) and gender (Gocsál, 2001). In addition, pauses can be influenced by the specific language (Zwirner & Zwirner, 1937; Trouvain & Möbius, 2014; Trouvain et al., 2016); various syntactic factors like the length and complexity of the sentence (Volkskaya, 2003; Krivokapic, 2007); and the position of the pause within the utterance (Sallai & Szende, 1995; Gósy, 2004; Vallent, 2005; Menyhárt, 2010).

Ever since empirical investigations of speech pauses began, the determination of the minimal duration of a silent pause has been a matter of controversy. The assumed minimal value first varied between 100 and 250 ms, depending on author (Rochester, 1973), then on the basis of Goldman-Eisler's (1968) study it became accepted at 250 ms. The main aim of setting that lower limit was to be able to tell genuine pauses from breaks due to mere articulatory reasons (the silent phase of voiceless plosives). Technological developments later made it possible to identify pauses of extremely short duration; hence the limit was gradually decreased to as little as 30 ms. Today, it is up to the individual researcher whether or not she wants to establish a lower limit for pause length, and if she does, she must explain why she chose the value she did (Váradi, 1988; Gósy, 2000).

In his monograph on spontaneous speech, Levelt (1989) distinguishes three different reasons for a lack of signal. Specifically, the "silence" that occurs during speech may be

(i) a "pause" that always occurs within the current unit being said,
(ii) a "gap" that occurs between discourse units, and
(iii) a "lapse" that indicates the end of conversation.

Partly based on Levelt's terminology, Markó (2005a) introduces, along with pause, the notion of tacitness defined as a gap between two units of conversation, and stillness defined as silence preceding or following an act of speaking (be it a conversation or a monologue). The present paper investigates pauses to the exclusion of gaps and lapses (or of tacitness and stillness).

The aim of the present paper is to give a comprehensive temporal analysis of silent pauses in spontaneous Hungarian speech. Our initial hypotheses are that

(i) silent pauses occurring in spontaneous speech can be classified in terms of the functions they serve;
(ii) each type of pause exhibits a unique pattern both in frequency of occurrence and in duration, and
(iii) the duration of silent pauses is determined, in addition to their functions, by their positions within the utterance.

2 Method

In this paper, we investigate the occurrence of silent pauses in spontaneous narratives and opinion-giving monologues of 10 speakers taken from the BEA Hungarian Spontaneous Speech Database (Gósy et al., 2012). Each of the 5 male and 5 female speakers is a monolingual adult speaker of Standard Hungarian living in Budapest. The participants were between 20 and 40 years of age, their mean age was 27.4 years. The total length of the recorded material analysed here was 71 minutes 35 seconds, or 7 minutes 10 seconds of spontaneous speech per participant (on average). We found a total of 1603 silent pauses (160.3 pauses per person) making up one-fifth (20.76%) of the total speaking time. The mean length of the pauses was 554 ms, 22.4 pauses occurred per minute. This meant that pauses occurred once every 2142 ms on average. Table 1 shows that male and female participants differed from one another in this corpus. Men talked less and paused more (25.93%) than women did (16.46%), and men's average pause lengths were also higher. The table also illustrates individual differences. While participant f4 spoke for the longest time, pauses made up 12.5% of her total speaking time, and she produced the shortest silent pauses. We found the highest ratio of pausing (34.34%) and the longest pauses on average (945.4 ms) with speaker m2. The shortest monologue (4.04 min) was produced by speaker m5.

Table 1: Speaking time and pausing by participant

	speaking time (min)	**pause ratio (%)**	**mean pause duration (ms)**	**pause frequency (ms/pause)**
m1	4.80	25.76	543.33	1523
m2	6.75	34.34	945.40	1643
m3	8.61	30.26	735.50	1627
m4	5.52	18.35	518.10	2035
m5	4.04	20.97	445.84	1639
f1	10.63	20.09	639.92	2419
f2	6.93	19.46	517.84	2087
f3	4.58	16.35	376.55	1815
f4	11.43	12.50	367.76	2529
f5	8.29	13.92	450.14	2704
men	29.73	25.93	637.63	1693
women	41.87	16.46	470.44	2311
all	71.60	20.76	554.04	2142

Pauses were classified primarily in terms of whether they were related to some disfluency phenomenon or were used to increase the comprehensibility of the text. As disfluency phenomenon, pauses served as editing phases (provided time for detecting some covert or overt error and possibly correcting it). The second type of pause was used to segment speech into chunks, thereby facilitating listener comprehension. Silent pauses facilitating comprehension were labelled as **S** (syntactical silent pauses), while editing phases were labelled as **E.** Within both major groups, we defined subgroups. Editing phases **(E)** were subcategorized in terms of the type of disfluency phenomenon they represented. In the present investigation, we identified the following types of editing phases:

(a) **repetitions** (**E_rep**; *nagyon fontos hogy* E_rep *hogy mi veszi körül* 'it is very important what *E_rep* what surrounds it'),
(b) **restarts** (**E_res**; *nekem fontos volt ho-* E_res *hogy így legyen* 'I found it important th- *E_res* that is should be so'),
(c) **false starts** (**E_falseS**; *korábban teljesen kizár-* E_falseS *meghatározónak tűnt* 'earlier it appeared to be totally exclu- *E_falseS* crucial'),
(d) **false words** (**E_falseW**; *ennyi pénzér amennyiér* E_falseW *amennyibe egy békávé bérlet kerül* 'for so much money, the cost *E_falseW* price of a season ticket'),
(e) **anticipations** (**E_ant**; *mivel nem* E_ant *náluk nem volt szabad* 'as were *E_ant* they were not allowed to do that'), and
(f) cases of **pause in word** (**PinW**; *megismerek számos élet* PinW *vitelt* 'I get familiar with a number of life *PinW* styles').

For other reparable disfluency phenomena we did not find any examples in our corpus; but obviously new categories can be defined for any other disfluency phenomenon. We subcategorized syntactical silent pauses having a segmentation function (**S**) in terms of their positions within the utterances. We distinguished **utterance onset pauses (S_Uo)** occurring at turn-taking when the current speaker began talking; such pauses may be preceded at most by a contentless expletive or a discourse marker: Interviewer: *De most már annyira megemelték a bérlet árát is* 'But the price of season tickets have increased so drastically now'. Subject: *Hát* S_Uo *relatív, mert ha azt számolod, hogy…* 'Well *S_Uo* it depends, for if you consider...'. **Silent pauses at phrase boundaries (S_PhrB)** are ones that occur at phrase boundaries within virtual sentences (Gósy 2003b), often before or after conjunctions: *Személyes hobbinak is tekintem, és* S_PhrB *szerencsére vannak is lehetőségeim ebben a szakmában* 'I take it to be a personal hobby, and *S_PhrB* fortunately I also have possibilities in this trade'. We labelled pauses occurring within a grammatical unit ("clause") as **within phrase pauses (S_PhrW)**: *Egy havi nyolcezer forintos kiadás nem nagy* S_PhrW *összeg* 'Spending eight thousand forints per month is not a big *S_PhrW* sum'. Finally, **end of phrase pauses (S_PhrE)** are silent pauses occurring after virtual sentences, where the speaker begins a new virtual sentence, often a new unit of thought: *Előre nem közölt kritériumok alapján osztályoztak le.* S_PhrE *Egyébként a szakkal kapcsolatban azt gondolom, hogy…* 'They assigned marks based on criteria that were not stated previously. *S_PhrE* Otherwise, with respect to the major I think that...' (Figure 1).

Annotation and the determination of phrase durations were accomplished with Version 5.4.21 of Praat (Boersma & Weenink, 2013), and statistical analyses with SPSS 20. We built a general linear mixed model (GLMM) on our data where independent variables were the individual pause types, dependent variables were the duration values, and the random factors were speakers and gender.

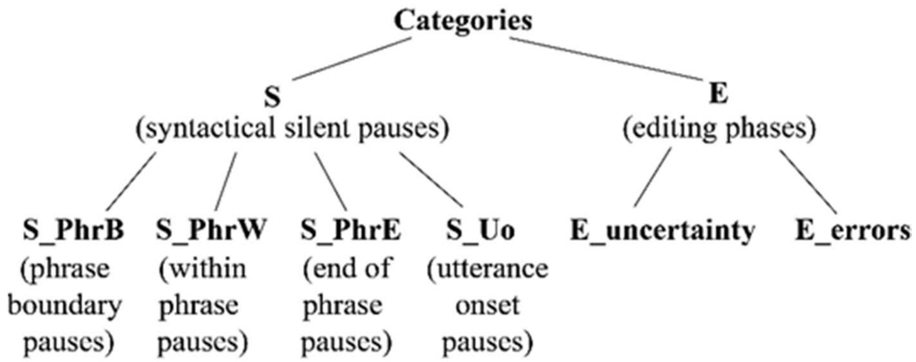

Figure 1.
The category system

3 Results

First we analysed our two main categories, syntactical silent pauses **(S)** and editing phases **(E)**. Their ratio was found to be as follows: 87.84% of the silent pauses we found belonged to group **S**, and 12.16% to group **E**. Of the total speaking time, syntactical silent pauses comprised 18.23% and editing phases a mere 2.53%. On the basis of our data, we can say that silent pauses occurring in spontaneous speech predominantly serve a segmentation purpose, while only slightly above one-tenth of them provide time for error repair. This is obviously related to the fact identified in several studies (Nooteboom, 1980; Gósy, 2008; Gyarmathy, 2010, 2011, 2012b, 2012c; Neuberger, 2010, 2011) that speakers repair roughly half of their errors. The two groups of pauses differ in their durations, but that difference was not significant (Figure 2).

Syntactical silent pauses had an average duration of 559 ms, while editing phases were 494 ms. Standard deviations and minimum and maximum values show that segmentation pauses are far more homogeneous: here, the data fall between 379 ms and 935 ms and deviate from the mean by 169 ms on average. Whereas editing phases (min: 264 ms, max: 980 ms) do so by 240 ms (Table 2). In the case of group **S,** the larger deviation and the wider interval are explained by the fact that error repair is associated with a variety of levels of planning, which take different amounts of time. When a given kind of disharmony occurs at a higher level, it takes more time to correct it, in general (Gyarmathy, 2012a).

Silent pauses used for segmentation occurred most often (54.05%) at phrase boundaries (**S_PhrB**), i.e., between virtual clauses. Pauses breaking grammatical structures, i.e., occurring within virtual clauses (**S_PhrW**), were found in 31.53% of the cases, those at the end of phrases (**S_PhrE**) in 12.71%, while those at the beginning of the utterance (**S_Uo**) in a mere 1.7% of cases. Most editing phases

(65.64%) contained repetitions (**E_rep**), 12.82% involved the phenomenon of pause in word (**PinW**), 8.21% were false word choice (**E_falseW**), 5.64% were restarts (**E_res**), 4.10% were false starts (**E_falseS**), and 3.59% involved anticipation (**E_ant**). That is, the primary source of editing phases occurring in the corpus was the speaker's uncertainty (84.1%), and only 15.9% were due to some kind of error.

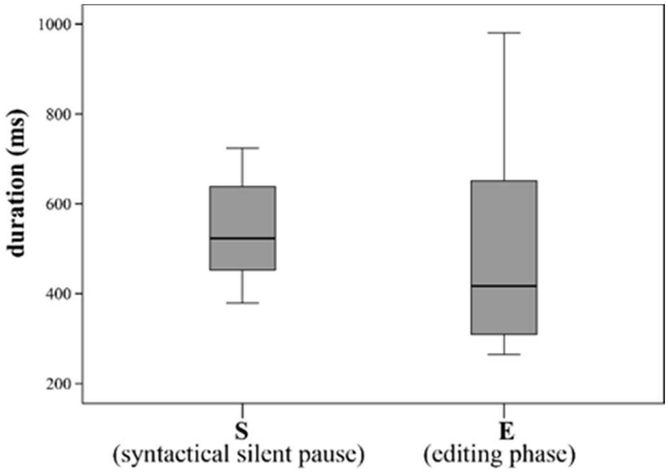

Figure 2.
The duration values of silent pauses serving segmentation **(S)** and realized as editing phases **(E)**.

Table 2: Data of silent pauses serving segmentation **(S)** and editing phases **(E)**

type	occurrence (%)	mean (SD) (ms)	min-max (ms)
S	87.84	559 (169)	379–935
E	12.16	494 (240)	264–980

In the category of silent pauses of segmentation, the longest average duration (899 ms; standard deviation: 598 ms) was found for utterance onset pauses (**S_Uo**); end of phrase pauses (**S_PhrE**) were of similar length (mean: 875 ms; standard deviation: 377 ms). The mean length of pauses at phrase boundaries (**S_PhrB**) was 587 ms (standard deviation: 215 ms), while that of within phrase pauses (**S_PhrW**) was 385 ms (standard deviation: 96 ms) (Table 3). The largest standard deviation was found in **S_Uo**. This was due to the fact that at the beginning of an utterance, speakers have to figure out the content of their utterance, and depending on how much the given topic falls within the given person's area of interest and how much background knowledge they have in it, planning may take more or less time. The relatively long duration of **S_PhrE** may have happend for a variety of reasons. The speaker, having finished a given train of thoughts, may offer their partner the possibility of turn taking, but if this does not take place, the original speaker may use the time for planning the continuation. The roughly 200 ms difference between the average lengths of **S_PhrB** and **S_PhrW**, together with the difference of the relevant values of standard deviation, revealed that speakers tend to produce longer pauses and ones with more variability in length between virtual clauses, helping the listener to segment their utterances and giving themselves time to think of what to say next. On the other hand, in the case of within phrase pauses, grammatical structures (may) get broken up, possibly even hampering comprehension. Hence the speaker instinctively tries to pause as little as possible in this situation.

The pause duration values we obtained may have been affected by the register used by the speaker. For instance, in an interview situation, the experimenter never cuts in and only rarely takes turns (even then, with the intention of helping the participant carry on). The participants are then allowed to formulate their ideas more at ease and with less pressure than in a real conversational situation (cf. Markó 2005a).

Statistical analysis of the data showed that the type of a pause determines its duration: $F(3, 31) = 16.355$; $p < 0.001$. Pairwise contrasts further showed that most groups significantly differed from one another. The duration of phrase boundary pauses (**S_PhrB**) differed from that of both within phrase (**S_PhrW**) and end of phrase (**S_PhrE**) ones, $p < 0.001$ in both cases. The length of within phrase pauses also significantly differed from that of end of phrase ($p < 0.001$) and from that of utterance onset ($p = 0.009$) ones.

Table 3: Data of subtypes of silent pauses serving segmentation

type	**occurrence (%)**	**mean (SD) (ms)**	**min-max (ms)**
S_PhrB	54.05	587 (215)	355–1113
S_PhrW	31.53	385 (96)	267–512
S_PhrE	12.71	875 (377)	537–1757
S_Uo	1.70	899 (598)	81–1921

From among editing phases, the longest average duration was found for restarts (**E_res**), it was 563 ms. This was followed by false word choices (**E_falseW**) at 504 ms. The average durations of pauses between repetitions (**E_rep**, 500 ms) and after anticipations (**E_ant**, 490 ms) were just below those values. The phenomenon of pause in word (**PinW**) reached an average duration of 351 ms, while that of false starts (**E_falseS**) was the shortest with 306 ms (see Table 4). From the average values we can see, among other things, which disfluency phenomena require the longest editing phases to resolve and which levels of planning may involve the most difficult problems for our speakers. Restarts are traditionally classified as phenomena pertaining to the speaker's uncertainty (see the introductory section above), although an earlier study showed that at least in some cases, they are akin to false starts and can be traced back to problems in lexical access (Gyarmathy, 2012c; Gyarmathy et al., 2015a, 2015b). The speaker has doubts concerning the correctness of the word she has chosen while she is pronouncing it and stops, but then her self-monitoring process decides that the word was correct after all, and the word is pronounced again, now in its entirety. The phenomenon of false word can be associated with the beginning of the speech planning process: lexical access. The speaker's lexical search comes up with the wrong word. This situation can only be resolved by starting the whole planning process from scratch (Horváth & Gyarmathy, 2010, 2012). Repetitions are also a type of disfluency suggesting the speaker's uncertainty. Speakers often make use of them as a strategy for gaining extra time in cases where a problem arises either in the portion of the utterance already said or in the part still to come (Gyarmathy, 2009). The higher average duration values of these three groups can be explained by the fact that higher levels of planning are involved in producing them. Statistical analyses again showed that the type of disfluency the given pause is associated with affects its duration ($F(5, 28) = 1.936$; $p = 0.040$), but in this case, the level of significance was weak. This may be partly due to the smaller number of tokens in each group. Pairwise contrasts confirmed statistically significant differences only between restarts and false starts ($p = 0.002$) and false starts and false words ($p = 0.027$).

Table 4: Data of subtypes of silent pauses serving as editing phases

type	occurrence (%)	mean (SD) (ms)	min–max (ms)
E_rep	65.64	500 (181)	258–779
PinW	12.82	351 (173)	124–605
E_falseW	8.21	504 (439)	68–1413
E_res	5.64	563 (262)	352–960
E_falseS	4.10	306 (267)	44–713
E_ant	3.59	490 (195)	331–772

Given that the less frequent phenomena in the system of categories pertaining to pauses associated with disfluency phenomena introduced above did not occur with all ten participants (for instance, anticipation occurred in the speech of four speakers only), these types of pauses could not be built into our general linear mixed model. In order to make up for missing values, our classification was modified in a way where we classified disfluencies in terms of whether they were of the uncertainty type or of the error type. The average duration of pauses associated with errors (including false starts, wrong words, and anticipations in the present case) was 445 ms (standard deviation: 383 ms). The range of data fell between 56 and 1413 ms. Silent pauses associated with the speaker's uncertainty (in the present corpus, repetitions, restarts, and the pause in word phenomenon) were realised with an average duration of 478 ms (standard deviation: 159 ms), and the range was between 290 and 721 ms. That is, average values of uncertainty disfluencies were somewhat longer on average, and this group turned out to be far more homogeneous.

Thus, we used four groups of silent pauses functioning for segmentation (**S_PhrB**; **S_PhrM**; **S_PhrF**; **S_Ui**) in our general linear mixed model, as well as including **uncertainty** disfluencies and **errors**. The average values of our two novel categories were: silent pauses associated with errors were realised with an average duration of 445 ms (standard deviation: 383 ms; min: 56 ms, max: 1413 ms), and uncertainty pauses with 478 ms (standard deviation: 159 ms; min: 290 ms, max: 722 ms). These six types constituted our independent variables and the dependent variables were the duration values. The random factors were speakers and gender, and we also studied the interaction of gender and types. The results showed that that the types of pauses strongly determined their durations: $F(5, 47) = 12.376$; $p < 0.001$. The speaker's gender also affected pause durations, but this effect was weaker: $F(1, 47) = 4.507$; $p = 0.039$ (Figure 3).

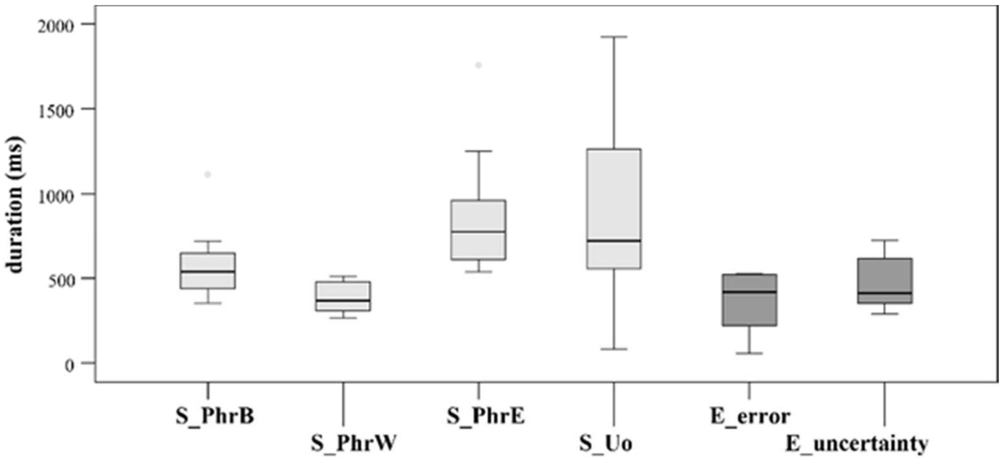

Figure 3.
The duration values of the various pause types

Pairwise contrasts confirmed the existence of significant differences between categories in a number of cases; see Table 5 for the details. It was only in six cases that comparisons did not yield statistically significant differences. The largest significant difference was found between pauses at phrase boundaries and within phrase pauses, between within phrase and end of phrase pauses, and between end of phrase pauses and those associated with uncertainty. Furthermore, large differences were found between end of phrase and error-bound pauses and between phrase boundary and end of phrase pauses. Differences of medium significance were found between within phrase and utterance onset pauses, as well as between within phrase and uncertainty-related ones. Weak significant differences were noted between utterance onset pauses and those associated with uncertainties or errors.

Considering our data in terms of gender, we can see that the groups of men and women differ in a number of respects. Men used both types of pauses with longer durations than women. The pausing strategies of both genders nevertheless showed similar tendencies (Figure 4), which were corroborated by our statistics. It was clear from both men's ($F(5, 47) = 10.207$; $p < 0.001$) and women's data ($F(5, 47) = 3.409$; $p = 0.010$) that pause type determined pause duration. Duration values were closest to one another in the case of within phrase silent pauses (**S_PhrW**) and editing phases of uncertainty disfluencies (**E_uncertainty**) where differences between genders were a mere 60–70 ms. Differences of 200–250 ms were found in the case of pauses at phrase boundaries and disfluencies of the error type, while the largest differences (300–400 ms) were noted with utterance onset (**S_Uo**) and end of phrase (**S_PhrE**) silent pauses.

Table 5: Pairwise contrasts for the various pause types

type	*p*-values
S_PhrB – S_PhrW	0.000
S_PhrB – S_PhrE	0.005
S_PhrW – S_PhrE	0.000
S_PhrW – S_Uo	0.008
S_PhrW – E_uncertainty	0.011
S_PhrE – E_error	0.003
S_PhrE – E_uncertainty	0.000
S_Uo – E_error	0.035
S_Uo – E_uncertainty	0.028

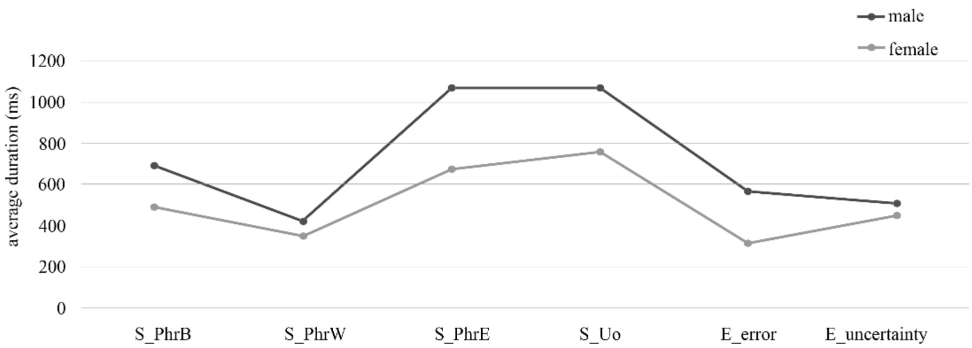

Figure 4.
Average durations of pauses with female and male participants.

Detailed data of men and women with respect to the individual types of pauses can be found in Table 6. These data show which groups of pauses can be considered homogeneous across the genders. On the basis of standard deviations, the most homogeneous group turned out to be **S_PhrW** for both genders; values of the individual speakers deviate from the mean by a mere 92 and 95 ms, respectively. Accordingly, minimum and maximum values were similar for the two genders.

Although average values were close to one another for editing phases of uncertainty disfluencies, this group was less homogeneous for both genders: minimum and maximum values differed, although the intervals between them hardly differed at all. In the remaining four categories, we can generally state that the women's data were far more homogeneous than the men's data: both their minima and maxima were lower and their intervals were narrower. The widest dispersion was found in the group of utterance onset pauses for both genders (women: 81–1264 ms, standard deviation: 482 ms; men: 255–1921 ms, standard deviation: 756 ms), but the women's values stayed below those of men.

Table 6: Mean values of pause types with the two genders

types	**genders**	**mean (SD) (ms)**	**min–max (ms)**
S_PhrB	male	689 (252)	509–1113
	female	486 (120)	355–649
S_PhrW	male	421 (92)	277–510
	female	349 (95)	266–512
S_PhrE	male	1073 (459)	538–1757
	female	678 (105)	607–855
S_Uo	male	1073 (756)	255–1921
	female	760 (482)	81–1264
E_error	male	570 (502)	74–1413
	female	319 (195)	56–522
E_uncertainty	male	507 (174)	290–695
	female	449 (158)	349–722

We calculated pairwise contrasts for both genders. In the case of women, we found significant differences for two pairs, whereas for men, seven pairs of categories were significant (Table 7). With female speakers, a medium strong difference was found between within phrase vs. end of phrase pauses, and a weak difference between within phrase and phrase boundary ones. With male speakers, marked differences were found between pauses at phrase boundaries and within phrase pauses, between end of phrase and uncertainty-related pauses, as well as between within phrase and end of phrase pauses. Somewhat weaker differences were found between pauses at phrase boundaries and end of phrase pauses, and between end of phrase pauses and editing phases of errors. The weakest significance values were found between within phrase and utterance onset pauses, as well as between pauses at phrase boundaries and those associated with uncertainty phenomena.

Table 7: Pairwise contrasts for the various pause types with the two genders

gender	type	*p*-value
female	**S_PhrB – S_PhrW**	0.048
	S_PhrW – S_PhrE	0.009
male	**S_PhrB – S_PhrW**	0.000
	S_PhrB – S_PhrE	0.007
	S_PhrB – E_uncertainty	0.030
	S_PhrW – S_PhrE	0.000
	S_PhrW – S_Uo	0.027
	S_PhrE – E_error	0.011
	S_PhrE – E_uncertainty	0.000

4 Conclusions

Silent pauses occurring in spontaneous speech serve a number of functions: they facilitate the comprehension of utterances by segmenting them into manageable chunks, they provide the speaker with time to resolve the occasional planning disharmony or to repair errors. Pauses have a crucial role both in speech production and in speech perception processes as communication unfolds. They provide time for speakers to select items from their mental lexicons and to produce the air flow that is necessary for phonation. For the listener, they contribute to a more fluid operation of speech processing (Gósy, 2000).

In this paper we studied Hungarian monologues to investigate silent pauses occurring in spontaneous speech. A number of previous studies, both for Hungarian and other languages, confirmed that silent pauses occurred often during speech production (Misono & Kiritani, 1990; Gósy, 2000, 2003a; Menyhárt, 2003; Bóna, 2007, 2013b; Neuberger, 2014). In general, they account for 20–30% of total speaking time. Data from the ten speakers we studied reinforced this conclusion. Our aggregate numbers showed that 20.76% of their utterances were silent pauses. In our analyses, we distinguished silent pauses whose role was segmentation of texts (**S** – syntactical silent pauses) from those associated with various disfluency phenomena (**E** – editing phases) and categorised both groups into subgroups of pauses. 87.84% of all silent pauses belonged to group **S**, and only 12.16% of them were used as editing phases of some disfluency phenomenon. This does not mean, of course, that speakers cannot utilise the time provided by pauses occurring in a segmentation function/position for resolving their occasional speech planning uncertainties; but unless an error appears in the surface structure of speech, they attempt to repair it without breaking the meaningful segmentation.

Within group **S**, we distinguished four subcategories in terms of where the pauses occurred in the utterance: the groups of utterance onset (**S_Uo**), phrase boundary (**S_PhrB**), within phrase (**S_PhrW**), and end of phrase (**S_PhrE**) pauses. The most frequently occurring type was **S_PhrB**, more than half of all data belonged here. Within phrase pauses comprised another one-third of the data, end of phrase pauses were one-tenth, while utterance onset pauses amounted to a mere 1%. If we analyse these ratios with respect to how many of the pauses occurred in a grammatically justified position, that is, without breaking the semantic and syntactic unity of the utterance (**S_PhrB**, **S_PhrE**, **S_Uo**), we can conclude that, in speech planning, speakers do not only plan the content and form of their utterances but also the pauses that occur in these utterances (Rochester, 1973; Zellner, 1994; Ramanarayanan et al., 2009). This proves that breathing operates during speech as a kind of subordinate process. In the background of within phrase pauses that break the grammatical structure, we can assume some kind of major planning disturbance whose discovery requires further study. The longest duration was characteristic of the **S_Uo** group, and the shortest time span belonged to **S_PhrW**. Duration values of the individual groups may refer to underlying processes going on in the background. In the case of utterance onset pauses, the speaker needs time to think over what she is going to say; the relatively longer time spent on end of phrase pauses may be due to the fact that, having finished one train of thought, speech planning has to start afresh for the next. In the case of within phrase pauses, their shorter duration may correlate with the communication need of intelligibility or processability. Statistical analyses have confirmed that the types of pauses determine their durations.

In group **E**, subcategories were determined by the types of disfluency found in the recordings analysed. We could distinguish pauses associated with repetitions (**E_rep**), restarts (**E_res**), false starts (**E_falseS**), false words (**E_falseW**), anticipations (**E_ant**), and the pause in word phenomenon (**PinW**). Frequency data showed that disfluency-related pauses are often be linked with phenomena based on the speaker's uncertainty than with actual errors. Average lengths of the individual subgroups corroborated the fact confirmed by a number of previous studies (see Gyarmathy, 2012a) that the higher level of speech planning where a given kind of disharmony occurs, the more time is needed to perform the correction. Statistical analyses confirmed that types of pauses determined pause durations in this group as well. The general conclusion we can draw on the basis of our results is that the duration of pauses is determined by the function they serve in communication.

Statistical analyses confirmed the existence of significant differences across gender. Pauses produced by male participants were longer in all cases, but the pattern of types was identical for both genders. Thus, we can assume that the strategies of pausing in spontaneous speech are determined by the structure of the language at hand, rather than by the individual speaker. Further proof of this claim, however, awaits future study.

Acknowledgement
This paper was supported by the János Bolyai Research Scolarship of the Hungarian Academy of Sciences.

References

Balassa, J. 1886. *A phonetika elemei, különös tekintettel a magyar nyelvre* [Elements of phontetics, with special regard to Hungarian]. Budapest: Magyar Tudományos Akadémia.

Bieling, A. 1880. *Das Prinzip der deutschen Interpunktion nebst einer übersichtlichen Darstellung ihrer Geschichte*. Berlin.

Boersma, P. & Weenink, D. 2013. *Praat: doing phonetics by computer* (Version 5.4.21) [Computer program]. http://www.praat.org (retrieved March 2016)

Boomer, D. S. 1965. Hesitation and grammatical encoding. *Language and Speech, 8,* 148-158.

Bóna, J. 2007. *A felgyorsult beszéd produkciós és percepciós sajátosságai*. [Features of the production and perception of accelerated speech]. PhD dissertation, Budapest.

Bóna, J. 2010. Beszédtervezési folyamatok az életkor és a beszédstílus függvényében [Speech planning processes across ages and registers]. *Magyar Nyelvőr, 134,* 332-341.

Bóna, J. 2012. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban [Features of spontaneous conversation in the young-old, old-old, and centenarian age groups]. In: Markó, A. (ed.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig* [Speech science: From acquisition to voice onset time]. Budapest: ELTE BTK–MTA Nyelvtudományi Intézet. 100-115.

Bóna, J. 2013a. *A spontán beszéd sajátosságai az időskorban* [Characteristics of old speakers' spontaneous speech]. Budapest: ELTE Eötvös Kiadó.

Bóna, J.2013b. A beszédszünetek fonetikai sajátosságai a beszédtípus függvényében [Phonetic properties of silent pauses depending on speech style]. *Beszédkutatás 2013*, 60-76.

Duez, D. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech, 25,* 11-25.

Esposito, A., Stejskal, V., Smékal, Z., & Bourbakis, N. 2007. The significance of empty speech pauses: Cognitive and algorithmic issues. In: Mele, F., Ramella, G., Santello, S., & Ventriglia, F. (eds.): *Advances in brain, vision, and artificial intelligence.* Berlin & Heidelberg: Springer. 542-554.

Fónagy, I. 1967. Áthajlás, szünet, szerkezet [Enjambement, pause, and structure]. *Nyelvtudományi Közlemények, 69,* 313-343.

Gocsál, Á. 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? [Do women speak faster than men do?] *Beszédkutatás 2001*, 61-72.

Goldman-Eisler, F. 1958. Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology, 10,* 96-106.

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.

Gósy, M. 2003a. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései [Disfluencies in spontaneous speech: Their frequency of occurrence and interrelations]. *Magyar Nyelvőr, 127(3),* 257-277.

Gósy, M. 2003b. Virtuális mondatok a spontán beszédben [Virtual sentences in spontaneous speech]. *Beszédkutatás 2003,* 19-43.

Gósy, M. 2004. *Fonetika, a beszéd tudománya*. [Phonetics: The science of speech]. Budapest: Osiris Kiadó.

Gósy, M. 2008. Önellenőrzési folyamatok a spontán beszédben [Self-monitoring processes in spontaneous speech]. *Magyar Nyelv, 104*(4), 402-426.

Gósy, M., Beke, A. & Horváth, V. 2011. Temporális variabilitás a spontán beszédben [Temporal variability in spontaneous speech]. *Beszédkutatás 2011*, 5-30.

Gósy, M., Gyarmathy, D., Horváth, V., Gráczi, T. E., Beke, A., Neuberger, T., & Nikléczy, P. 2012. BEA: Beszélt nyelvi adatbázis [BEA: A Hungarian spontaneous speech database]. In: Gósy, M. (ed.): *Beszéd, adatbázis, kutatások.* [Speech, database, research]. Budapest: Akadémiai Kiadó. 9-25.

Gyarmathy, D. 2007. Az alkohol hatása a spontán beszédprodukcióra [The effect of alcohol on spontaneous speech production]. *Beszédkutatás 2007*, 108-121.

Gyarmathy, D. 2008. Különböző zajok hatása a beszédprodukcióra [The effect of various noises on speech production]. *Alkalmazott Nyelvtudomány, 8(*1-2), 135-147.

Gyarmathy, D. 2009. A beszélő bizonytalanságának jelzései: ismétlések és újraindítások [Speaker uncertainty signals: repetitions and restarts]. *Beszédkutatás 2009,* 196-216.

Gyarmathy, D. 2010. A beszédellenőrzés működése alkoholos állapotban [Speech monitoring under the influence of alcohol]. In: Bárdosi, V. (ed.), *Világkép a nyelvben és a nyelvhasználatban* [World view in language and language use]. Budapest: Tinta Könyvkiadó. 125-135.

Gyarmathy, D. 2011. Strategies of disfluency repairs in spontaneous speech. *The Phonetician, 103/104,* 88-96.

Gyarmathy, D. 2012a. Az agyi monitorozás módosulása zajhatásra [The modification of cerebral monitoring under the influence of noise]. In: Navracsics, J. & Szabó, D. (eds.): *Mentális folyamatok a nyelvi feldolgozásban* [Mental processes in language processing]. *Pszicholingvisztikai Tanulmányok III.* 170-181.

Gyarmathy, D. 2012b. Rejtett és felszíni monitorozás a spontán beszédben. In: Bárdosi, V. (ed.), *A szótól a szövegig (konferenciakötet)* [From words to texts: Proceedings]. 99-109.

Gyarmathy, D. 2012c. Kétarcú újraindítás [Janus-faced restarts]. In: Markó, A. (ed.), *Beszédtudomány* [Speech science]. ELTE Bölcsészettudományi Kar, MTA Nyelvtudományi Intézet, Budapest. 50-67.

Gyarmathy, D., Neuberger, T., & Auszmann, A. 2015a. Ismétlések és újraindítások temporális mintázatai [Temporal patterns of repetitions and restarts]. In Gósy, M. (ed.), *Diszharmóniás jelenségek a beszédben* [Disharmonic phenomena in speech]. Budapest: MTA Nyelvtudományi Intézet. 131-149.

Gyarmathy, D., Gósy, M., Horváth, V., Neuberger, T., & Beke, A. 2015b. A szerkesztési szakaszok sajátosságai újraindításkor és téves kezdéskor spontán beszédben [Properties of editing phases in restarts and false starts in spontaneous speech]. In: Gósy, M. (ed.), *Diszharmóniás jelenségek a beszédben* [Disharmonic phenomena in speech]. Budapest: MTA Nyelvtudományi Intézet. 171-193.

Hargreaves, W. A. & Starkweather, J. A. 1959. Collection of temporal data with the duration tabulator. *Journal of the Experimental Analysis of Behavior,* 2, 179.

Hegedűs, L. 1953. On the problem of the pauses of speech. *Acta Linguistica Hungarica, 3,* 1-36.

Hevesi, S. 1908. *Az előadás művészete* [The art of performing]. Budapest: Stampfel.

Horváth, V. & Gyarmathy, D. 2010. „A lónak is négy nyelve van, mégis megbotlik": a mentális lexikon útvesztői ["Horses have four tongues, yet they stumble": In the maze of the mental lexicon]. *Beszédkutatás 2010*, 171-183.

Horváth, V. & Gyarmathy, D. 2012. Téves kezdések magyarázó elvei a spontán beszédben [Explanatory principles of false starts in spontaneous speech]. In: Gósy, M. (ed.), *Beszéd, adatbázis, kutatások* [Speech, database, research]. Budapest: Akadémiai Kiadó. 138-154.

Imre, A. 2005. Különböző műfajú szövegek szupraszegmentális jellemzői [Suprasegmental features of texts of various genres]. *Magyar Nyelvőr, 129,* 510-520.

Jespersen, O. 1904. *Lehrbuch der Phonetik*. Leipzig & Berlin: B. G. Teubner.

Jones, D. 1922. *An outline of English phonetics*. Leipzig: G. E. Stechert & Co.

Krivokapic, J. 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics, 35*(2), 162-179.

Laczkó, M. 2009. Középiskolai tanulók spontán beszédének temporális jellegzetességei [Temporal features of the spontaneous speech of secondary school students]. *Magyar Nyelvőr, 133*(4), 447-467.

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

Levin, H., Silverman, I., & Ford, B. 1967. Hesitations in children's speech during explanation and description. *Journal of Verbal Learning and Verbal Behavior, 6,* 560-564.

Lindroth, H. 1933. Sprachpsychologie und Interpunktion. *Archives Néerlandaises de Phonétique Expérimentale, VIII-IX.*

Mahl, G. 1956. Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology, 53,* 1-15.

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

Markó, A. 2005b. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben [Characteristics of temporal structure in various communicative situations]. *Beszédkutatás 2005*, 63-77.

Mátray, G. 1861. *A rendszeres szavalattan alaprajza* [The basics of a systematic study of poetry recitals]. Pest: Trattner–Károlyi.

Menyhárt, K. 2003. A spontán beszéd megakadásjelenségei az életkor függvényében [Disfluency phenomena in spontaneous speech across ages]. In: Hunyadi, L. (ed.), *Kísérleti fonetika – laboratóriumi fonológia a gyakorlatban*. [Experimental phonetics and laboratory phonology: practical aspects]. Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója,125-138.

Menyhárt, K. 2010. A beszédsebesség objektív mérési és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél [Objective measurements and subjective perception of speech rate today and sixty years ago]. *Beszédkutatás 2010,* 110-124.

Misono, Y. & Kiritani, Sh. 1990. The distribution pattern of pauses in lecture-style speech. *Logopedics and Phoniatrics, 2,* 110-113.

Neuberger, T. 2010. Korrekciós folyamatok gyermekek spontán beszédében [Correction processes in children's spontaneous speech]. In: Váradi, T. (ed.), *IV. Alkalmazott Nyelvészeti Doktorandusz Konferencia tanulmánykötete*. Budapest: MTA Nyelvtudományi Intézet. 112–123.

Neuberger, T. 2011. Az önkorrekciós folyamatok fejlődése gyermekkorban [The development of self-correction mechanisms in children]. In: Gecső, T. & Sárdi, Cs. (eds.), *Nyelvi funkciók – stílus és kapcsolat* [The functions of language, style, and connections]. Székesfehérvár & Budapest: KJF & Tinta Kiadó. 189-195.

Neuberger, T. 2014. *A spontán beszéd sajátosságai gyermekkorban*. [Features of children's spontaneous speech]. Budapest: ELTE Eötvös Kiadó.

Nooteboom, S. 1980. Speaking and unspeaking: detection and correction of phonological and lexical errors in spontaneous speech. In: Fromkin, V. A. (ed.), *Errors in linguistic performance. Slips of the tongue, ear, pen and hand.* New York & London: Academic Press. 87-97.

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.

Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L., & Narayanan, Sh. S. 2009. Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *Journal of the Acoustical Society of America, 126,* 160-165.

Rochester, S. R. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research, 2*(1), 51-81.

Sallai, J. & Szende, T. 1975. A szünet mint funkció [Pause as function]. *A Magyar Nyelvtudományi Társaság Kiadványai 143.*

Sallai, J. & Szende, T. 1995. Spontán közlések beszédszüneteinek pszicholingvisztikai értelmezése (egészséges és skizofrén közlők szövegeiben) [A psycholinguistic interpretation of speech pauses in spontaneous utterances produced by normal and schizophrenic speakers]. *Általános Nyelvészeti Tanulmányok XVIII*, 209-222.

Simonyi, Zs. 1903. *Iskolai helyesírás* [School orthography]. Budapest: Atheneum.

Sweet, H. 1877. *A handbook of phonetics*. Oxford: Clarendon Press.

Sweet, H. 1890. *A primer of phonetics.* Oxford: Clarendon Press.

Szende, T. 1979. A szünet és a junktúra [Pauses and junctures]. *Magyar Fonetikai Füzetek, 4*, 7-32.

Tannenbaum, P. H., Williams, F., & Wood, B. S. 1967. Hesitation phenomena and related encoding characteristics in speech and typewriting. *Language and Speech, 10,* 203-215.

Trouvain, J. & Möbius, B. 2014. Individuelle Ausprägung von Atmungspausen in der Mutter- und in der Fremdsprache als Anzeichen kognitiver Belastung. In: *Elektronische Sprachsignalverarbeitung 2014: Tagungsband der 25. Konferenz.* Dresden. 177-184.

Trouvain, J., Fauth, C., & Möbius, B. 2016. Breath and non-breath pauses in fluent and disfluent phases of German and French L1 and L2 read speech. *Proceedings of Speech Prosody 2016*. Boston. 31-35.

Vallent, B. 2005. A spontán beszéd ötven éve és ma. Esettanulmány. *Beszédkutatás 2005.* 99-111.

Váradi, T. 1988. A beszédszünet szubjektív és objektív regisztrálásának összevetéséről [On comparing the subjective and objective registration of speech pauses]. In: Kontra, M. (ed.), *Beszélt nyelvi tanulmányok* [Papers on spontaneous speech]. Budapest: MTA Nyelvtudományi Intézete. 44-59.

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.

Verzeano, M. & Finesinger, J. E. 1949. An automatic analyzer for the study of speech in interaction and in free association. *Science, 110,* 45.

Viëtor, W. 1894. *Elemente der Phonetik*. Leipzig.

Volkskaya, N B. 2003. Virtual and real pauses at clause and sentence boundaries. In: *Proceedings of the 15h International Congress of Phonetic Science*s. Barcelona. 499-502.

Weiske, J. 1838. *Theorie der Interpunktion aus der Idee des Satzes*. Leipzig.

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

Zwirner, E. & Zwirner, K. 1937. Phonometrischer Beitrag zur Frage der Lesepausen. *Archives Néerlandaises de Phonétique Expérimentale, XIII.* 111-128.