MDPI

Article

Occurrences and Durations of Filled Pauses in Relation to Words and Silent Pauses in Spontaneous Speech

Mária Gósy [1,2]

1 Hungarian Research Centre for Linguistics, 1068 Budapest, Hungary; gosy.maria@nytud.hu
2 Eötvös Loránd University, 1088 Budapest, Hungary

Abstract: Filled pauses (i.e., gaps in speech production filled with non-lexical vocalizations) have been studied for more than sixty years in different languages. These studies utilize many different approaches to explore the origins, specific patterns, forms, incidents, positions, and functions of filled pauses. The present research examines the presence of filled pauses by considering the adjacent words and silent pauses that define their immediate positions as well as the influence of the immediate position on filled pause duration. The durations of 2450 filled pauses produced in 30 narratives were analyzed in terms of their incidence, immediate positions, neighboring silent pauses, and surrounding word types. The data obtained showed that filled pauses that were attached to a word on one side were the most frequent. Filled pauses occurring within a word and between two silent pauses were the longest of all. Hence, the durations of filled pauses were significantly influenced by the silent pauses occurring in their vicinity. The durations and occurrence of filled pauses did not differ when content or function words preceded the filled pause or followed it. These findings suggest that the incidence and duration of filled pauses as influenced by the neighboring words and silent pauses may be indicative of their information content, which is related to the processes of transforming ideas into grammatical structures.

Keywords: filled pause positions; functions; spontaneous speech planning; durations; silent pauses; content words; function words

1. Introduction

Various units of the speech signal possess characteristics that carry semantic, emotional, pragmatic, and other types of information (e.g., Brodbeck et al. 2018; Gwilliams and Davis 2022; Jaeger 2010; Rastall 2006). The speech flow, however, also contains gaps, i.e., various kinds of pauses, resulting from speech planning processes. Compared to meaningful lexical units and phrases, these gaps may appear to be simply disruptive events in fluent speech production. Indeed, many types of gaps (certain pauses, coughs, laughs, smacks, etc.) are natural phenomena in speech and disrupt fluent speech (e.g., Bachorowski et al. 2001; Li and He 2011; Trouvain 2014). Looking at the gaps that are pauses from a functional aspect, we note that these gaps are not accidental, instead they fulfill various important tasks in the speech production process. In this paper, we focus on one particular type of gap, the vocalic filled pause that consists only of a vowel (henceforward FP).

Researchers have amassed a great deal of knowledge on FPs during the past decades. FPs have been studied in different languages and from different perspectives. There are various explanations provided for filled pause production, their functions in speech, and their perception by listeners (Corley et al. 2007; Corley and Stewart 2008; Ferreira and Bailey 2004; Finlayson and Corley 2012; Fox Tree 2002; Lickley 2015; Navarretta 2015; O'Connell and Kowal 2005; Shriberg 2001; Clark and Fox Tree 2002; Stouten et al. 2006; Tottie 2016; Watanabe et al. 2008; etc.). Investigations have studied the incidence and characteristics of FPs during both L1 and L2 language acquisition, with typical speakers and also with atypical speakers (e.g., Bortfeld et al. 2001; Gayraud et al. 2011; Gósy et al. 2014; Hlavac 2011;

de Jong and Bosker 2013; Merlo and Barbosa 2010; Searl et al. 2002). Research has demonstrated the frequent clustering of discourse markers and filled pauses (Crible et al. 2017; Kosmala and Crible 2022). In sum, researchers have generally agreed that FPs provide extra time for the speaker to plan and execute speech, signal the delay, monitor their speech, and repair errors and flaws, as well as signal conversational turns and serve as discourse markers (e.g., Kosmala and Crible 2022; Levelt 1989; Lickley 2015; Postma 2000; Clark and Fox Tree 2002; Shriberg 2001; Tottie 2016).

Before going into the details of this study, let us have a look at the forms of FPs in Hungarian, where [ø]-like or [ə]-like vowels are frequently used to fill pauses in spontaneous utterances. These vowels are responsible for more than 70% of all vocalic FPs (Gósy et al. 2014). Other vocalizations, such as *m* [m], *öh* [øh], *öm* [øm], and *öhm* [øhm] are also commonly used (the consonant [h] is phonetically existent in [øh] and [øhm]). The Hungarian phoneme inventory contains both the phonemically short and long /ø/ vowel, which is a rounded front mid vowel whose phonemically long version is also the 3rd person singular personal pronoun (ő/øː/ = 'he/she/it'). (There is no gender marking in Hungarian pronouns.) The neutral vowel is not a member of the Hungarian phoneme inventory and not an allophone, either. However, speakers produce [ə]-like vowels randomly in specific phonetic contexts, particularly in fast speech (Siptár and Törkenczy 2000).

The occurrences of FPs are influenced by many factors, like the age and mental state of the speaker, speakers' behavior, phonetic context, the syntactic structures involved, the length of the utterance, topic, communication situations, etc. (Bortfeld et al. 2001; Navarretta 2015; Roberts et al. 2009; Shriberg 2001; Yaruss et al. 1999; Watanabe et al. 2008; etc.). Typically, authors put the mean incidence of FPs at about 6% of all the words uttered (Eklund 2010). When FPs were examined in a large corpus, they were reported to be from one-third to over one-half of all disfluencies (Shriberg 2001). An average of four filled pauses per about 70 words was found in Canadian English-speaking subjects' narratives (Roberts et al. 2009). In an investigation on Hungarian FPs, speakers produced 3.82 filled pauses per minute on average with large individual differences that ranged between 0.8 to 9.5 occurrences in narratives (Horváth 2010).

The durations of FPs reported in the literature range from about 100 ms to about 750 ms or even much longer (e.g., Clark and Fox Tree 2002; de Jong and Bosker 2013; Merlo and Barbosa 2010; Shriberg 2001). The mean duration of FPs produced by young Hungarian speakers was reported to be 344 ms, and FPs that occurred between two silent pauses were significantly longer than those in other positions (Gósy et al. 2017). In an earlier study, the mean duration of FPs preceded by silent pauses and attached to an adjacent word was 599 ms, while those that were followed by a silent pause had a mean duration of 554 ms (Horváth 2010). Note, that in this latter research study, all types of vocalic FPs were considered, and both young and middle-aged speakers participated in the recordings.

A fairly large number of studies have focused on the positions of FPs in relation to certain words, sentences, phrases, and turns, as well as their alignment to the adjacent word. As early as the 1950s, Maclay and Osgood (1959) found that the majority of FPs were located at word boundaries. These authors also analyzed the interrelations between the occurrences of FPs and unfilled pauses and various word types. They found that both FPs and unfilled pauses occurred more frequently before content words than before function words. Comparing the incidence of FPs and unfilled pauses, they found that FPs were more likely to appear before function words. Later, Cook found the opposite tendency (1971). In his study, FPs occurred more often before pronouns, but less often before nouns, verbs, and adverbs. Boomer (1965) reported that the most frequent FP position was after the first word of phonemic clauses. Similarly, Cook (1971) found that FPs tended to occur before the first, second, or third word of a clause. Finally, Eklund and Shriberg (1998) analyzed the occurrences of FPs in sentence-initial and in sentence-medial positions. They also found that FPs were used more frequently in sentence-initial than in sentence-medial positions. They argued that FPs occurring at the start of an utterance or turn indicated that the speaker wanted to hold the floor. The question then arose whether

filled pauses were distributed at random in spontaneous speech or was there an explanation for their distribution? Lounsbury (1954) suggested that FPs occurred when the speaker faced the highest statistical uncertainty as to what to say next (i.e., the highest entropy). Several studies have confirmed that the production of FPs was affected by the number of options available to the speaker (e.g., Schachter et al. 1991). About 50% of all FPs occur in utterance-initial and/or phrase-initial positions. Beattie and Barnard reported 55.3% for initial positions of FPs (Beattie and Barnard 1979); Eklund found 45.5% (Eklund 2004); 45% was reported by O'Connell and Kowal (2005), and 63% by Pfeifer and Bickmore (2009). Since these studies were concerned with the observation of FP positions in phrases (sentences, turns, etc.), little is known about the immediate positions of FPs, i.e., whether they were attached (on one side) to an adjacent word or not. In addition, these studies considered various types of FPs not only vocalic FPs.

FPs can be produced differently depending on adjacent words and silent pauses that define their immediate position. Theoretically, FPs may occur in the following five positions in relation to their immediate context: (i) FP occurs between two silent pauses (silFPsil), for example, *és akkor* silFPsil *kifestették a szobát* 'and then silFPsil/they/painted the room'; (ii) FP is preceded by a silent pause and attached to a word on the other side (FPword), for example, *bezárták a* sil FP*boltot* '/they/closed/the/sil FPshop'; (iii) FP is attached to a word on one side and followed by a silent pause on the other side (wordFP), for example, *énekelnek egy dalt*FP sil '/they/sing a songFP sil'; (iv) FP occurs between two words (without any silent pause), for example, *iskolába*FP*mentek* 'to schoolFP/they/went'; and (v) FP occurs within a word inserted into diverse parts of the word, for example, *zsaro*FP*lás* 'blackFPmail'. It is the speaker who decides on the positions as they utter FPs, but their selection appears to be characteristic of the language, too (e.g., Korotaev et al. 2020; de Leeuw 2007; Swerts 1998).

The immediate position of FPs in relation to either silent pauses or adjacent words is usually left unspecified in studies analyzing FPs, and currently there is no consensus even on the terms that could be used to identify these positions. Clark and Fox Tree (2002) termed the position where an FP is attached to the last segment of a word as cliticization. The problem is with the term of cliticization that it is used in linguistics with a different meaning, which may cause misunderstanding and uncertainty when used in FP-contexts. If one wants to put emphasis on the articulation gestures, FPs attached to a word can be termed as a kind of co-articulation with the preceding or following word, particularly where the filler is vocalic, as in the case of Hungarian. However, the term coarticulation is used in phonetics in a well-defined phonetic context, therefore, it is doubtful whether the term 'coarticulation' is really appropriate for this phenomenon. In analyzing Russian monologues, Korotaev and his colleagues (Korotaev et al. 2020) used the terms 'clusters' and 'quasi-clusters' to describe the immediate positions of FPs, as opposed to the context referred to as 'isolation'. According to their definition, FPs occur in 'quasi-clusters' when they are preceded or followed by silent pauses. 'Clusters' are identified when two different disfluency types occur one after another or separated by only one word. Authors describing FP-contexts in Hebrew (Silber-Varod et al. 2016) used the terms enclitic FPs (for the FPword position) and proclitic FPs (for the wordFP position), borrowing these terms from morphology and syntax. Considering all of the possible terms, the present author decided not to use them because of their well-defined meanings in various fields of linguistics and will refer to these positions using the simple verb 'attach'. Attaching a vocalic FP to an adjacent word means that it is coarticulated either with the first or with the last segment of an adjacent word. The coarticulated sound is the same as the realizations of the /ø/ and /øː/ phonemes in regular Hungarian words, such as *öreg* 'old', *örvény* 'maelstrom', *őzike* 'fawn', *selejtező* 'qualifier', *temető* 'cemetery'. Speakers produce words in which an [ø] acts as an FP either at the beginning or at the end of the word without any difficulty. In the following examples, the letters *ö* and *ő* (denoting the short and long phonemes /ø/ and /øː/) exemplify the use of this FP: *öretek* (ö+*retek*) 'FPradish', *önárcisz* (ö+*nárcisz*) 'FPdaffodil', *üvegő* (*üveg*+ő) 'bottleFP', *társalgáső* (*társalgás*+ő) 'discourseFP'.

Speakers have been reported to attach an FP onto a previous word, but not onto a following word in native (British) English speech (Clark and Fox Tree 2002). These researchers found that *uh* and *um* were often cliticized onto prior words but never onto following words (2002). In addition, FPs were found to be cliticized onto functional words (Clark and Fox Tree 2002). FPs were found to be more frequent between a lexical item and a silent pause than between two silent pauses in another study also in (British) English speech (de Leeuw 2007). Dutch and German speakers seemed to behave differently in positioning FPs. Attached FPs were found to be common in Dutch but not in German (de Leeuw 2007). Silber-Varod et al. (2016) found that in Hebrew, attached FPs were more common than FPs between silent pauses, and more specifically, FPs in wordFP positions are more common than those in FPword positions. In Russian monologues, no FPs were found followed by a silent pause (Korotaev et al. 2020). Guaïtella (1993) found that FPs occurred in various positions in French spontaneous speech. She reported that 50% of all FPs occurred in wordFP positions, while only 5% occurred in FPword positions. FPs between two silent pauses and those occurring between words shared with similar frequencies, were found in 22.5% of all cases. Four FP positions with different incidences were identified in Dutch (Swerts 1998): 15.5% of all FPs occurred in FPword positions, 34% in wordFP positions, 27% in positions where FPs were surrounded by silent pauses and 23.5% occurred between words. FPs within a word are rare in Swedish and no cases were found in American English (Eklund and Shriberg 1998). FPs and also silent pauses can be found inserted into a word by many speakers in Hungarian spontaneous speech (Gósy and Krepsz 2017).

This study intends to describe the occurrences and durations of FPs focusing on their immediate positions in Hungarian spontaneous speech. We seek answers to the following questions. (i) What is the distribution of the immediate surroundings of FPs and SPs? (ii) Are FPs attached to the first segment of a word more frequently than to the last segment of a word? (iii) Are the durations of FPs dependent on the immediate positions and on the type of word preceding and following them?

The present research aims to shed light on the temporal properties and incidence of [ø]-like and [ə]-like vocalic FPs according to their immediate positions in spontaneous speech produced by young Hungarian-speaking subjects. The present study will not analyze the positions of FPs with respect to either their places in the phrases or their combinations with discourse markers (Kosmala and Crible 2022; Navarretta 2015).

Based on findings in the literature, we assumed that (i) the incidence of FPs would show significant differences in various immediate positions, (ii) the durations of FPs would show significant differences depending on their immediate positions, (iii) the durations of silent pauses in the vicinity of FPs would show close interrelationships with the durations of FPs, (iv) the type of word (content vs. function) would have an effect on the duration of FPs, and (v) the speakers would show differences both in the incidence and durations of FPs in their speech samples. Finally, we hypothesized that (vi) the implicit information of FPs about speech planning would be revealed by their immediate positional properties.

2. Materials and Methods

The speech material consisted of spontaneous narratives produced by 30 native speakers of Hungarian. Young speakers were randomly selected from the BEA Hungarian speech database (Gósy 2012). Speaker age ranged between 23 and 29 years, with a mean age of 25 years and half of the participants were women. All of them were monolingual speakers of Hungarian, had normal hearing, and none of them had any speech defect. All the participants lived in Budapest speaking colloquial Hungarian. They either studied at the university at the time of the recording or had a university degree and worked in various professions (teacher, engineer, researcher, computer analyst, economist, actor, physician, chemist, social worker, musician, consultant, etc.). The speakers' mean articulation rate (excluding pauses) was 4.6 syllables/s, ranging from 4.2 syllables/s up to 4.9 syllables/s.

According to the BEA database protocol, all the participants were asked to speak about their life and their views on topics of current interest suggested by the interviewer (who

was the same person across all recordings). Recordings were made in the same sound-attenuated room under identical technical conditions using an Audiotechnica AT4040 cardioid condenser microphone connected directly to a computer using GoldWave to record samples at 44.1 kHz, 16 bits, monaurally. For the present study, more than 6.4 h of speech samples from the BEA database were used. The duration of recording per subject was about 13 min (SD = 0.3 min).

The speech material was manually annotated, focusing on vocalic FPs (variants of the vowel [ø] and the neutral vowel [ə]). In the annotations, FPs were marked by the letters *öö* (the *ö* letter sounds as [ø] in Hungarian) while silent pauses were indicated by SIL. SIL was defined as a silent interval longer than 100 ms since minimum duration of silent pauses in the vicinity of FPs was 100 ms in our material. The duration of most silent pauses in the vicinity of filled pauses was around 500 ms. There were only two incidents where 30-ms-long silent pauses were found; they preceded FPs and they were excluded from analysis. Although silent periods may contain breathing noise (Trouvain et al. 2016), these were ignored in annotating the speech material. Parts-of-speech for the words preceding or following the FPs were identified as either content words or function words. Annotations were done in Praat (Boersma and Weenink 2015). No disagreement was found in the identification of filled pauses in the speech samples between the two annotators (who were both phoneticians).

The FPs identified in the speech material were coded according to the following four most frequent types (FPs occurring between words were relatively infrequent in the speech material, therefore this position was not considered.) There were two positions in which FPs were attached to a lexical item on one side (and were preceded or followed by a silent pause on the other side). An FP may be attached to the first segment of the word (this is the FPword position) after a silent pause and it may be attached to the last segment of the word (this is the wordFP position) followed by a silent pause. An FP may occur within a word without a silent pause preceding or following it (i.e., the within-word position). No further analysis was carried out as to the interruption point within the word. FPs may occur surrounded by silent pauses on both sides (this is the silFPsil position). Examples (FPs are marked by **öö** for better recognition):

(i) Type FPword: *meg akarta mutatni az új* SIL **öö***lakást* ('/he/wanted to show the new SIL ööflat');

(ii) Type wordFP: *és vonattal***öö** SIL *megyünk a tóhoz* ('and/we/travel by trainöö SIL to the lake');

(iii) Type FP within word: *ott láttuk a kom***öö***ponistát* ('there/we/saw the comööposer');

(iv) Type silFPsil: *kiskorom óta* SIL **öö** SIL *hegedültem* ('from my early childhood SIL **öö** SIL/I/played the violin').

Measurements were also made in Praat (Boersma and Weenink 2015). The duration was measured as the interval (i) between the onset and offset of the second formant of the vocalic FPs occurring between silent pauses, and (ii) between the onset/offset of the second formant of the vocalic FP and the onset/offset of the preceding and following segment based on traditional criteria. Duration of an FP inserted into a word was segmented and measured according to traditional acoustic-phonetic methods (Stevens 1999). Figure 1 shows two FP positions: in one of them, the vocalic FP is attached to the last segment of a word, and in the other one, it is attached to the first segment of the word.
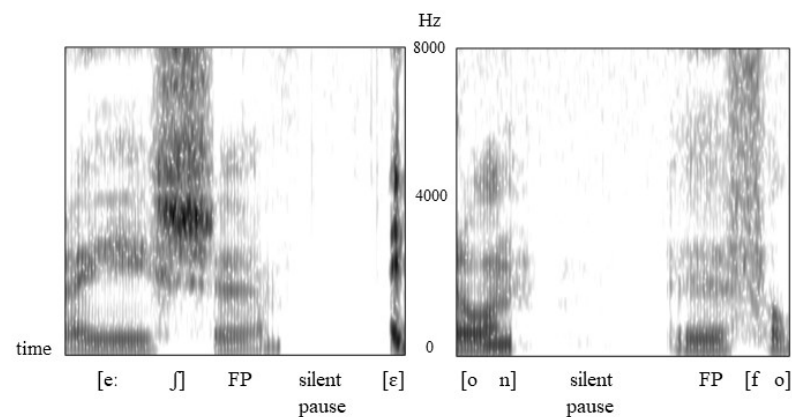
Figure 1. Speech fragments where FP is followed (**left**) and preceded (**right**) by a silent pause and is attached to the last consonant [ʃ] of the preceding word *és* 'and' (**left**) and to the first consonant [f] of the following word *fogunk* '/we/will' (**right**).

Durations were extracted automatically using a specific Praat script (developed in the Phonetics Laboratory of the Hungarian Linguistics Institute) based on annotations. A total of 2450 FPs were found in the speech material. The total number of silent pauses was 2873. The incidence of FPs varied according to the speaker, ranging from 52 to 126 items. The positions wordFP, FPword, and silFPsil were found in all speakers, while no within-word FP was found in four participants. Since the speakers' mean speech rate was very similar, the incidence of FPs was expressed in items/minute.

FPs were analyzed according to (i) incidence, (ii) duration, and (iii) the effects of position, silent pauses, as well as preceding and following words. In addition, analysis focused on the interactions of filled and silent pause durations, as well as of FP durations and the type of the preceding and following words (function vs. content words).

All statistical analyses were performed using R (R Core Team 2022). A goodness-of-fit test was performed for variable FP positions. The hypothesis tested was whether the population probabilities were equal (R Core Team 2022). To test the relationship between speaker and FP position, a Pearson's chi-square test was used (also using test package R (R Core Team 2022). To assess the effects of FPposition, SILbefore, SILafter, word type (content vs. function word before and after) on FPduration, a linear mixed effects model (LMM) with random-effect intercepts was run, with FP duration (in ms) as the dependent variable, FPposition, SILbefore, SILafter, word type before and after as the fixed effects, as well as participant (speaker) as random effects. The LMM analysis was performed using the lme4 and lmerTest packages (Bates et al. 2015; Kuznetsova et al. 2017). A Kruskal–Wallis test was also performed. Post-hoc tests with Bonferroni correction were performed in the emmeans package (Length 2021). Satterthwaite approximations to degrees of freedom were used in analyses.

3. Results

The data obtained in this study will be discussed in terms of incidence, duration, and position. The effects of silent pauses and content vs. function words preceding and following FPs will be presented together with their statistical analyses.

3.1. Incidence of FPs

When considering all FPs, we noted 6.3 occurrences of FPs per minute in our speech samples. The occurrence of FPs in various positions demonstrated a specific pattern. FPs that were attached to a word were the most frequent one (5 occurrences/minute), particularly those that were attached to the last segment of the preceding word (3.1 occurrences/minute). FPs that were attached to the first segment of the word occurred more infrequently (1.9 occurrences/minute). FPs that were articulated within words were even less frequent (0.2 occurrences/minute).

The distribution of FPs according to various positions expressed in percentages, with total occurrences amounting to 100%, revealed that the majority of FPs occurred in the wordFP position (49.39%), followed by the FPword position (30.49%), while the proportions of FPs between two silent pauses (16.65%) were lower and even lower within words (3.47%). Thus, the speakers who provided our speech material preferred to use FPs mainly attached to the end of a word (wordFP position), and they used FPs least frequently within a word. The great majority of FPs within a word occurred between root morphemes and suffixes, between prefixes and root morphemes, and between two suffixes. There were only two words found where FPs occurred within the root morpheme.

Statistical analysis showed significant differences in FP incidence depending on their positions ($\chi^2(3) = 1135.0$, $p < 0.001$). Regarding the relatively low incidence of FPs within words, in subsequent statistical analyses, we eliminated these occurrences. Even so, the results showed significant differences in FP incidence depending on position ($\chi^2(2) = 411.2$, $p < 0.001$).

3.2. Words before and after FPs

An analysis was carried out to obtain information about the incidence of content and function words around pauses. The content words were mainly nouns and verbs, while the majority of function words were various conjunctions. In sum, 2387 content words (50.5%) and 2343 function words (49.5%) were involved with filled and silent pauses. The word type was not systematically analyzed in cases where an FP occurred within a word, although it was observed that most within-word FPs occurred in nouns.

The distributions of content and function words according to FP position do not show large differences ($\chi^2(2) = 3.2719$, $p = 0.1846$). Figure 2 (left) shows no difference in the incidence of FPs between content and function words when the latter preceded the FP. However, when FPs were surrounded by two silent pauses, the proportion of content words preceding FPs significantly increased ($\chi^2(2) = 13.1451$; $p = 0.0014$).
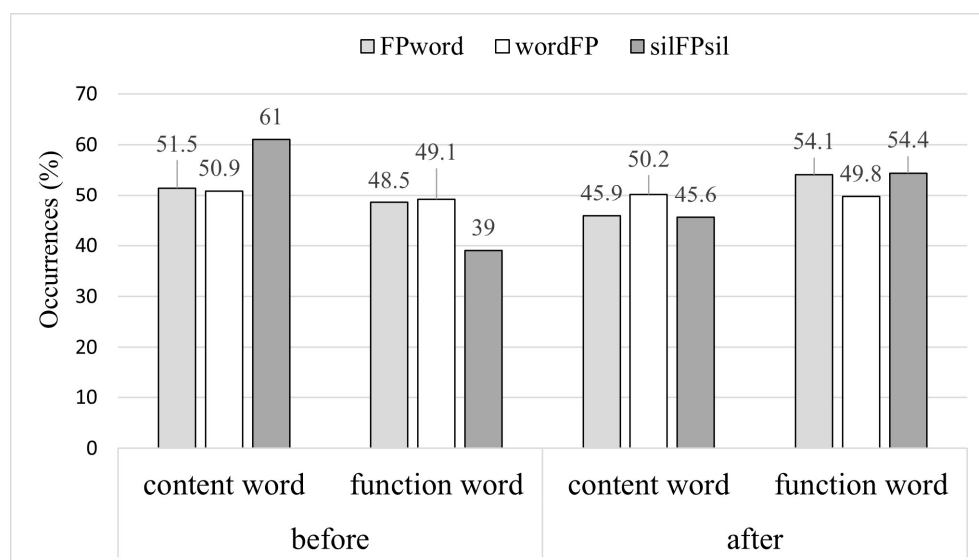

Figure 2. Incidence of content and function words according to FP positions.

The distributions of content and function words that followed FPs were similar depending on position (Figure 2, right). Statistical analysis revealed no differences depending on FP positions when content and function words followed FPs ($\chi^2(2) = 4.5209$, $p = 0.1043$). Although the difference in the incidence of content and function words that followed FPs in silFPsil positions was somewhat (by 8.8%) higher than in other positions, it did not reach a significant level. Although we expected differences depending on word type, the data

showed that there was no difference, except in the case of content words preceding the FP in the silFPsil position.

3.3. Durations of FPs

The mean duration of all FPs in our speech material was 412 ms (SD = 287 ms). The shortest one was 50 ms long while the longest duration that was found in our material was 2708 ms. The longest durations occurred in within-word positions (mean value = 652 ms) while they tended to be shorter between two silent pauses (mean value = 597 ms). FPs were much shorter when attached to a word on one side and were followed or preceded by a silent pause. FPs were shorter when attached to the first segment of the word (mean value = 329 ms in FPword position) than when they were attached to the last segment of the word (mean value = 383 ms in wordFP position). Figure 3 demonstrates FP durations according to the four positions.
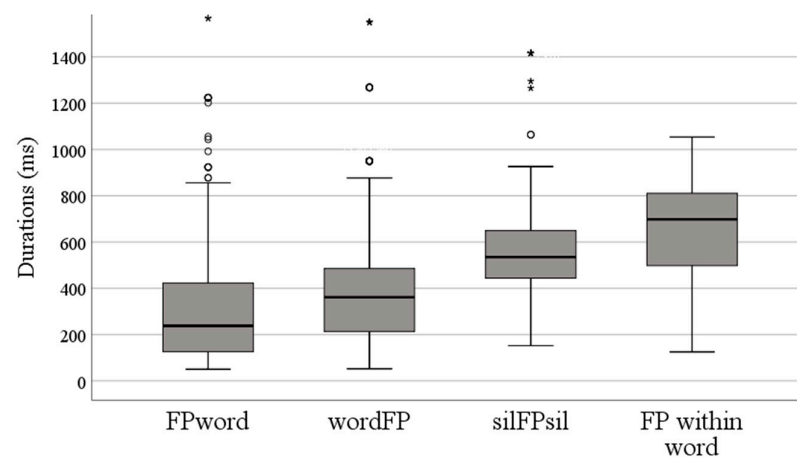

Figure 3. Durations of FPs according to their four positions (medians and interquartile ranges). (The asterisks and circles mark the outliers.)

Statistical analysis revealed that durations of FPs were significantly different depending on their positions (F(3, 2425.6) = 107.26, p < 0.001). In addition, pairwise comparisons were conducted that also confirmed significant differences (p < 0.05 in all cases).

3.4. Effects of Neighboring Words on FP Durations

A statistical analysis was carried out to obtain information on whether FP durations varied depending upon word type (content vs. function words) preceding or following the FP. The mean duration of FPs preceded by content words was 398 ms (SD = 273 ms), while it was 408 ms (SD = 298) when preceded by function words. Although FPs were longer after function words than after content words by about 10 ms, it raises the question about the importance or real effect of this difference. The mean duration of FPs when followed by either content or function words was the same, 403 ms (SD = 276 ms and 294 ms, respectively).

Statistical analysis showed that the type of word (content vs. function) preceding FPs had a significant effect on their duration (F(1, 2341.1) = 3.8648, p = 0.0494). Although the effect is significant, the difference, as noted above, is small. There was no significant effect on FP durations when either content or function words followed them (F(1, 2344.1) = 0.0079, p = 0.929). We think that it is safe to say that word type has practically no effect on FP durations.

3.5. Durations of Silent Pauses

The mean duration of silent pauses preceding FPs was 618 ms (SD = 562 ms), the shortest duration was 100 ms while the longest 4000 ms. The mean duration of silent pauses

following FPs was 586 ms (SD = 741 ms); the shortest was 159 ms while the longest 8764 ms. As explained above, silent pauses may precede FPs, may follow them, and FPs may occur between two silent pauses.

Silent pauses preceding FPs (FPword position) had a mean duration of 526 ms (SD = 558 ms) while those following FPs (wordFP position) had a mean duration of 536 ms (SD = 782 ms). Silent pauses that surrounded FPs showed a mean duration of 788 ms (SD = 531 ms) when preceding FPs and a mean duration of 736 ms (SD = 578 ms) when following FPs.

Statistical analysis was carried out to determine whether the durations of silent pauses influenced the durations of FPs. The durations of FPs in FPword positions are significantly influenced by the preceding silent pause ($F(1, 745) = 19.902$, $p < 0.001$). Similarly, the durations of FPs in wordFP positions are significantly influenced by the following silent pause ($F(1, 1012.4) = 165.86$, $p < 0.001$). In silFPsil positions, the durations of silent pauses had significant effects on FP durations (when silent pauses precede FPs: $F(1, 404.90) = 31.641$, $p < 0.001$; when silent pauses follow FPs: $F(1, 404.66) = 62.030$, $p < 0.001$). The effect of silent pauses in the latter cases, however, proved to be the opposite. The longer the silent pause preceding the FP, the longer the duration of the FP. The longer the silent pause following the FP, the shorter the duration of the FP.

3.6. Speakers

As expected, both the incidence and the durations of FPs showed characteristic differences among speakers. Results of the statistical analysis confirmed that the incidence of FPs in various positions were significantly different ($\chi^2(87) = 514.6$, $p < 0.001$) for different speakers. Figure 4 demonstrates FP and silent pause durations according to position for all speakers. The durations of FPs were significantly different among speakers in FPword, wordFP, and silFPsil positions ($\chi^2(29) = 121.42$, $p < 0.001$; $\chi^2(29) = 140.22$, $p < 0.001$; $\chi^2(29) = 101.4$, $p < 0.001$, respectively). No significant differences were found among speakers, however, when FPs occurred within words ($\chi^2(25) = 30.153$, $p = 0.2186$). This outcome can be explained by the relatively low number of FPs in this position.
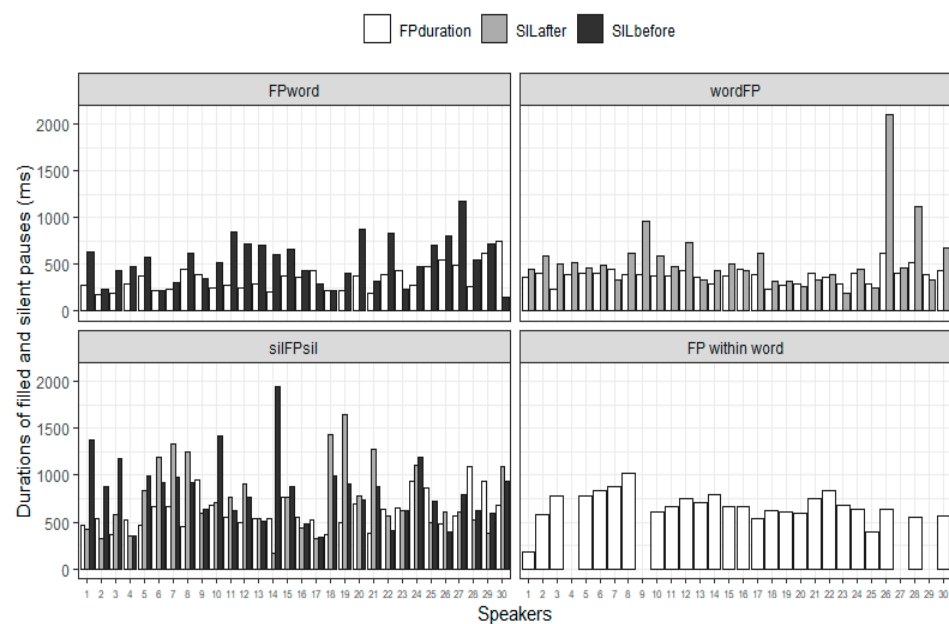

Figure 4. Durations of FPs and silent pauses according to the four FP positions analyzed for all the speakers. The white bars stand for the durations of FPs, grey bars stand for durations of silent pauses that follow FPs, while black bars stand for durations of silent pauses that precede FPs.

4. Discussion

Each unit of speech from segments to intonation curves contains information. Listeners process both speech and non-speech information at the same time during speech perception and speech comprehension either consciously or unconsciously (Brodbeck et al. 2018; Finlayson and Corley 2012; Frank and Jaeger 2008; Gwilliams and Davis 2022; Jaeger 2010; etc.). Specific gaps like FPs occurring in the speech flow convey information about speech planning, speaker behavior, language properties, topic content, etc., as well as various messages for the listeners (e.g., Cossavella and Cevasco 2021; Fox Tree 2002; Fraundorf and Watson 2011; Kirjavainen et al. 2021; Levelt 1989; Local and Kelly 1986; O'Connell and Kowal 2005; Tottie 2016). Although FPs appear to interrupt the speech flow, they are in fact important components of speech production that are audible to the listener. Most studies define the main functions of FPs as providing extra time for the speaker to overcome speech planning or execution difficulties and providing pragmatic signals (e.g., turn-taking) for listeners and supplement various information contents (e.g., Arnold et al. 2003; Ferreira and Bailey 2004; Fox Tree 2002; Finlayson and Corley 2012; Levelt 1989; Roberts et al. 2009; Roggia 2012; Swerts 1998; Tottie 2014).

The positional patterns of FPs, in general, seem to be characteristic of both the speaker and the language, and may contain information about speech planning (Beattie and Barnard 1979; Christenfeld 1994; Clark and Fox Tree 2002; Eklund and Shriberg 1998; de Leeuw 2007; Eklund 2010; Silber-Varod et al. 2016; etc.). The positionally anchored patterns shown and analyzed in this paper describe relationships with adjacent words and their possible functional representations discussed further in this section. Data revealed that the more frequent the FPs, the shorter they are, and more infrequent FPs are longer. The combination of an FP attached to a word and a silent pause was the most common and the shortest condition. If the FP was surrounded either by silent pauses or by segments of a word, they were the most infrequent and longest ones. This finding appears to be characteristic of FPs occurring in Hungarian spontaneous speech (see also e.g., Clark and Fox Tree 2002; Korotaev et al. 2020; de Leeuw 2007; Silber-Varod et al. 2016 for other languages). Therefore, further investigations are needed to confirm or fine-tune these relationships as characteristic of specific languages where patterns of FPs provide an opportunity to analyze their positions in immediate contexts.

Most of our assumptions were confirmed. The incidence and durations of FPs differed significantly depending on immediate positions. Silent pauses in the vicinity of FPs showed statistically confirmed interrelationships with the durations of FPs. Thus, positions and durations of both silent pauses and FPs together account for the speakers' underlying speech planning strategies. In the following paragraphs, we present our conception of the information content of FPs depending on their immediate positions.

(i) FPs within words are the longest, though they are relatively rare, and are related to the activation of the mental lexicon. Speakers interrupt word articulation when they need confirmation of whether they have accessed the lexeme correctly and when they are uncertain, for various reasons, about suffixes (Levelt et al. 1999; Özdemir et al. 2007; Slevc and Ferreira 2006). These FPs may also signal insecurity about the correctness of the word. Hungarian has a rich morphology and words can easily contain 3–4 suffixes that may explain the frequency with which FPs are inserted into words. The information content of FPs in this position may be confirmation of the correctness of lexical access and selection of suffixes.

(ii) FPs surrounded by silent pauses are long and relatively rare; they seem to contain information on the speaker's serious difficulties concerning the continuation of their speech. These difficulties may concern the level of thoughts, selection from thoughts, and also problems with grammatical formulation. The function of FPs in this position is associated with the selection of thoughts and grammatical formulation (Hartsuiker et al. 2005; Levelt 1989; Watanabe et al. 2008; etc.).

(iii) FPs preceded by a silent pause and attached to a word (FPword) suggest heightened control over previous stretches of speech. These FPs are the shortest and are relatively

frequent, which indicates that speakers who encountered problems in speech planning of various kinds could successfully solved them covertly. FPs in this position show a kind of "retrospective" control (Hlavac 2011; Levelt 1983; Postma 2000).

(iv) Use of the wordFP strategy suggests that the speaker is uncertain about upcoming parts of speech. This may concern lexical access (Navarretta 2015), morphological formulation, or uncertainty in articulation planning. These FPs are the most frequent in our material and they are shorter than those occurring between two silent pauses but longer than FPs in FPword positions. FPs in this position are supposed to show that some sort of search is taking place in the mental lexicon (for various reasons). According to Eklund and Shriberg (1998), in these cases, the speaker has already committed themself to the semantic content, but not yet to grammatical and lexical encoding.

The occurrence of FPs attached to a word may raise questions concerning word structure and word recognition. The *ö*-like attachment to a word does not violate Hungarian phonotactics or syllable structure. Hungarian words are relatively long (words consist of 3.5 syllables in spontaneous utterances, in general), so one more vowel either at the beginning or at the end of the word does not disrupt pronunciation.

Based on our data, we assume that FPs carry different kinds of information dependent upon their positions in immediate contexts. Although speakers continuously monitor their speech planning processes (Levelt et al. 1999; Postma 2000), they face speech planning problems more frequently than they can control during their speech flow. To solve speech planning problems during covert speech, planning takes more time and the demand for solving difficulties occurs relatively more frequently than in the case of control processes (e.g., Postma 2000). Control of phrases that have been uttered is in general less frequent. It happens when speakers, for one reason or another, become suspicious of the correctness of previously uttered words and repeat themselves. This control process requires significantly less time than online planning, so the FPs are shorter.

In Hungarian, speakers attach the majority of vocalic FPs either to the beginning or to the end of the words. This is true in close to 80% of all occurrences. Our explanation is that FPs that are co-articulated with words are less conspicuous. Listeners focus on lexemes as the primary semantic units. They may recognize the attached *ö*-like ([ø]-like) vowel or neutral vowel as a kind of noise that they can easily ignore. Experimental data have confirmed that listeners were not able to recognize FPs in spontaneous speech accurately (Horváth 2014; Kirjavainen et al. 2021). However, FPs inserted into a word are striking, particularly because their durations are the longest among all FPs.

FPs are stigmatized in colloquial Hungarian speech (Gósy et al. 2014). People are instructed not to use them too frequently in various verbal communication situations (schools, interviews, official talks, etc.). Therefore, speakers subconsciously try to "hide" unwanted FPs. Apparently, one of the strategies used to avoid isolated and longer FPs is to attach them to a word where they are not too conspicuous. In addition, in the cases of the attached FPs, one silent pause disappears, thereby increasing the fluency of speech.

5. Conclusions

The findings of this research project added support to the well-known claims made about the various functions of FPs (Corley et al. 2007; Hlavac 2011; Kosmala and Crible 2022; Levelt 1989; O'Connell and Kowal 2005; Postma 2000; Tottie 2016; etc.). In this study, we systematically focused on occurrences and durations of FPs in four different positions in relation to adjacent words and silent pauses. We think that this approach makes it possible to conduct fine-grained investigations that may widen our knowledge about the strategies speakers use while talking. Most of the FPs in the speech material we analyzed behaved as parts of the phonological construction of the words to which they were attached. Their incidence and temporal properties may signal various functions.

The data obtained in this study confirm the everyday experience that speakers vary significantly in the strategies that they use in producing FPs, both in respect of their incidence and durations (see also Clark and Fox Tree 2002). Although incidences of FPword,

wordFP, and silFPsil positions were used by all speakers, FPs occurring within a word were not. The variation observed suggests that different speakers have different strategies and put different amounts of cognitive effort into their speech production. Although the immediate context of FPs may comprise information on speech planning, correction, and control, a critical aspect of this idea is the well-known fact that there are a number of factors and specific interrelations among them that influence the production of FPs in all languages.

References

Arnold, Jennifer E., Maria Fagnano, and Michael K. Tanenhaus. 2003. Disfluencies signal *theee, um*, new information. *Journal of Psycholinguistic Research* 32: 25–36. [CrossRef] [PubMed]

Bachorowski, Jo-Anne, Moria J. Smoski, and Michael J. Owren. 2001. The acoustic features of human laughter. *The Journal of the Acoustical Society of America* 110: 1581–97. [CrossRef] [PubMed]

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

Beattie, Geoffrey W., and P. J. Barnard. 1979. The temporal structure of natural telephone conversations (directory enquiry calls). *Linguistics* 17: 213–29. [CrossRef]

Boersma, Paul, and David Weenink. 2015. Praat: Doing Phonetics by Computer. Available online: http://www.fon.hum.uva.nl/praat/download_win.html (accessed on 15 October 2019).

Boomer, Donald S. 1965. Hesitation and grammatical encoding. *Language and Speech* 8: 148–58. [CrossRef]

Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober, and Susan E. Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44: 123–47. [CrossRef]

Brodbeck, Christian, Elliot H. Hong, and Jonathan Z. Simon. 2018. Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology* 28: 3976–83. [CrossRef]

Christenfeld, Nicholas. 1994. Options and *um*s. *Journal of Language and Social Psychology* 13: 192–99. [CrossRef]

Clark, Herbert H., and Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84: 73–111. [CrossRef]

Cook, Mark. 1971. The incidence of filled pauses in relation to part of speech. *Language and Speech* 14: 135–39. [CrossRef]

Corley, Martin, and Oliver W. Stewart. 2008. Hesitation disfluencies in spontaneous speech: The meaning of *um*. *Language and Linguistics Compass* 2: 589–602. [CrossRef]

Corley, Martin, Lucy J. MacGregor, and David I. Donaldson. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105: 658–68. [CrossRef]

Cossavella, Francisco, and Jazmín Cevasco. 2021. The importance of studying the role of filled pauses in the construction of a coherent representation of spontaneous spoken discourse. *Journal of Cognitive Psychology* 33: 172–86. [CrossRef]

Crible, Ludivine, Liesbeth Degand, and Gaëtanelle Gilquin. 2017. The clustering of discourse markers and filled pauses. *Languages in Contrast* 17: 69–95. [CrossRef]

de Jong, Nivja H., and Hans Rutger Bosker. 2013. Choosing a threshold for silent pauses to measure second language fluency. Paper presented at the 6th Workshop on Disfluency in Spontaneous Speech (DiSS 2013), Stockholm, Sweden, August 21–23; Stockholm: KTH Royal Institute of Technology, pp. 17–20.

de Leeuw, Esther. 2007. Hesitation markers in English, German, and Dutch. *Journal of Germanic Linguistics* 19: 85–114. [CrossRef]

Eklund, Robert. 2004. Disfluency in Swedish Human–Human and Human–Machine Travel Booking Dialogues. Dissertation thesis, Linköping Studies in Science and Technology—Unitryck, Linköping, Sweden.

Eklund, Robert. 2010. The effect of directed and open disambiguation prompts in authentic call center data on the frequency and distribution of filled pauses and possible implications for filled pause hypotheses and data collection methodology. Paper presented at the DiSS-LPSS Joint Workshop 2010, 5th Workshop on Disfluency in Spontaneous Speech and 2nd International Symposium on Linguistic Patterns in Spontaneous Speech, Tokyo, Japan, September 25–26; pp. 23–26.

Eklund, Robert, and Elizabeth Shriberg. 1998. Crosslinguistic disfluency modelling: A comparative analysis of Swedish and American English human–human and human–machine dialogues. Paper presented at the ICSLP 98, Sydney, Australia, November 30–December 5; pp. 2631–34.

Ferreira, Fernanda, and Karl G. D. Bailey. 2004. Disfluencies and human language comprehension. *Trends in Cognitive Sciences* 8: 231–37. [CrossRef]

Finlayson, Ian R., and Martin Corley. 2012. Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin and Review* 19: 921–28. [CrossRef]

Fox Tree, Jean. 2002. Interpreting pauses and *um*s at turn exchanges. *Discourse Processes* 34: 37–55. [CrossRef]

Frank, Austin F., and Florian T. Jaeger. 2008. Speaking Rationally: Uniform information density as an optimal strategy for language production. Paper presented at the 30th Annual Meeting of the Cognitive Science Society (CogSci08), Washington, DC, USA, July 23–26; pp. 939–44.

Fraundorf, Scott H., and Duane G. Watson. 2011. The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language* 65: 161–75. [CrossRef]

Gayraud, Frédérique, Hye-Ran Lee, and Melissa Barkat-Defradas. 2011. Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics and Phonetics* 25: 198–209. [CrossRef]

Gósy, Mária. 2012. BEA—A multifunctional Hungarian spoken language data base. *The Phonetician* 105–106: 50–61.

Gósy, Mária, and Valéria Krepsz. 2017. Szünet a szóban: Típusok, jellemzők, időtartamok. [Pause within a word: Types, characteristics, durations]. In *Morfémák időzítési mintázatai a beszédben [Temporal Patterns of Morphemes in Speech]*. Budapest: MTA Nyelvtudományi Intézet, pp. 199–225. [CrossRef]

Gósy, Mária, Judit Bóna, András Beke, and Viktória Horváth. 2014. Phonetic characteristics of filled pauses: The effects of speakers' age. Paper presented at the 10th International Seminar on Speech Production (ISSP), Cologne, Germany, May 5–8; pp. 150–53.

Gósy, Mária, Dorottya Gyarmathy, and András Beke. 2017. Phonetic analysis of filled pauses based on a Hungarian-English learner corpus. *International Journal of Learner Corpus Research* 3: 151–76. [CrossRef]

Guaïtella, Isabelle. 1993. Functional, acoustical and perceptual analysis of vocal hesitations in spontaneous speech. Paper presented at the ESCA Workshop on Prosody, Lund, Sweden, September 27–29; pp. 128–31.

Gwilliams, Laura, and Matthew H. Davis. 2022. Extracting language content from speech sounds: The information theoretic approach. In *Speech Perception. Springer Handbook of Auditory Research*. Edited by Lori L. Holt, Jonathan E. Peelle, Allison B. Coffin, Arthur N. Popper and Richard R. Fay. Cham: Springer, vol. 74, pp. 113–39. [CrossRef]

Hartsuiker, Robert J., Herman H. H. Kolk, and Heike Martensen. 2005. The division of labor between internal and external speech monitoring. In *Phonological Encoding and Monitoring in Normal and Pathological Speech*. Edited by Robert J. Hartsuiker, Roelien Bastiaanse, Albert Postma and Frank Wijnen. New York: Psychology Press.

Hlavac, Jim. 2011. Hesitation and monitoring phenomena a in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation? *Journal of Pragmatics* 43: 3793–806. [CrossRef]

Horváth, Viktória. 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57: 288–306. [CrossRef]

Horváth, Viktória. 2014. *Hezitációs jelenségek a magyar beszédben [Hesitation Phenomena in Hungarian Speech]*. Budapest: ELTE Eötvös Kiadó.

Jaeger, Florian T. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61: 23–62. [CrossRef] [PubMed]

Kirjavainen, Minna, Ludivine Crible, and Kate Beeching. 2021. Can filled pauses be represented as linguistic items? Investigating the effect of exposure on the perception and production of *um*. *Language and Speech* 65: 1–27. [CrossRef]

Korotaev, Nikolay A., Vera I. Podlesskaya, Katerina V. Smirnova, and Olga V. Fedorova. 2020. Disfluencies in Russian spoken monologues: A distributional analysis. In Paper presented at the Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2020", Moscow, Russia, June 17–20; pp. 1–13.

Kosmala, Loulou, and Ludivine Crible. 2022. The dual status of filled pauses: Evidence from genre, proficiency and co-occurrence. *Language and Speech* 65: 216–39.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. Christensen. 2017. lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software* 82. [CrossRef]

Length, Russel V. 2021. Emmeans: Estimated Marginal Means, Aka Least-Squares Means. Available online: https://cran.r-project.org/web/packages/emmeans/index.html (accessed on 23 October 2022).

Levelt, Willem J. M. 1983. Monitoring and self-repair in speech. *Cognition* 14: 41–104. [CrossRef]

Levelt, Willem J. M. 1989. *Speaking. From Intention to Articulation*. Cambridge: MIT Press. [CrossRef]

Levelt, Willem J. M., Ardi Roelofs, and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22: 1–75. [CrossRef] [PubMed]

Li, Yanxiong, and Qianhua He. 2011. Detecting laughter in spontaneous speech by constructing laughter bouts. *International Journal of Speech Technology* 14: 211–25. [CrossRef]

Lickley, Robin J. 2015. Fluency and disfluency. In *The Handbook of Speech Production*. Edited by Melissa Redford. Hoboken: Wiley-Blackwell, pp. 445–69. [CrossRef]

Local, John, and John Kelly. 1986. Projection and 'silences': Notes on phonetic and conversational structure. *Human Studies* 9: 185–204. [CrossRef]

Lounsbury, Floyd G. 1954. Transitional probability, linguistic structure, and systems of habit-family hierarchies. In *Psycholinguistics. A Survey of Theory and Research Problems*. Edited by Charles E. Osgood and Thomas A. Sebeok. Baltimore: Waverly Press, pp. 93–101.

Maclay, Howard, and Charles E. Osgood. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15: 19–44. [CrossRef]

Merlo, Sandra, and Plínio Almeida Barbosa. 2010. Hesitation phenomena: A dynamical perspective. *Cognitive Processing* 11: 251–61. [CrossRef]

Navarretta, Constanza. 2015. The functions of fillers, filled pauses and co-occurring gestures in Danish dyadic conversations. Paper presented at the 3rd European Symposium on Multimodal Communication, Dublin, Ireland, September 17–18; pp. 55–61.

O'Connell, Daniel C., and Sabine Kowal. 2005. *Uh* and *um* revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research* 34: 555–76. [CrossRef]

Özdemir, Rebecca, Roelofs Ardi, and Willem J. Levelt. 2007. Perceptual uniqueness point effects in monitoring internal speech. *Cognition* 105: 457–65. [CrossRef]

Pfeifer, Laura M., and Timothy Bickmore. 2009. Should agents speak like, *um*, humans? The use of conversational fillers by virtual agents. In *Intelligent Virtual Agents*. Edited by Zsófia Ruttkay, Michael Kipp, Anton Nijholt and Hannes Högni Vilhjálmsson. IVA 2009. Lecture Notes in Computer Science. Berlin and Heidelberg: Springer, vol. 5773, pp. 460–66. [CrossRef]

Postma, Albert. 2000. Detection of errors during speech production: A review of speech monitoring models. *Cognition* 77: 97–131. [CrossRef] [PubMed]

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rastall, Paul. 2006. Language as communication, pattern and information. *La Linguistique* 42: 19–36. [CrossRef]

Roberts, Patricia M., Ann Meltzer, and Joanne Wilding. 2009. Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders* 42: 414–27. [CrossRef] [PubMed]

Roggia, Aaron B. 2012. *Eh* as a polyfunctional discourse marker in Dominican Spanish. *Journal of Pragmatics* 44: 1783–98. [CrossRef]

Schachter, Stanley, Nicholas Christenfeld, Bernard Ravina, and Frances Bilous. 1991. Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology* 60: 362–67. [CrossRef]

Searl, Jeffrey P., Rodney M. Gabel, and J. Steven Fulks. 2002. Speech disfluency in centenarians. *Journal of Communication Disorders* 35: 383–92. [CrossRef]

Shriberg, Elisabeth. 2001. To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetics Association* 31: 153–69. [CrossRef]

Silber-Varod, Vered, Hamutal Kreiner, Ronen Lovett, Yossi Levi-Belz, and Noam Amir. 2016. Do social anxiety individuals hesitate more? The prosodic profile of hesitation disfluencies in Social Anxiety Disorder individuals. Paper presented at the Speech Prosody 2016 (SP2016), Boston, MA, USA, May 31–June 3; pp. 2016–49.

Siptár, Péter, and Miklós Törkenczy. 2000. *The Phonology of Hungarian*. Oxford: Oxford University Press. [CrossRef]

Slevc, Robert L., and Victor S. Ferreira. 2006. Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language* 54: 515–40. [CrossRef]

Stevens, Kenneth. 1999. *Acoustic Phonetics*. Cambridge: MIT Press. [CrossRef]

Stouten, Frederik, Jacques Duchateau, Jean-Pierre Martens, and Patrick Wambacq. 2006. Coping with disfluencies in spontaneous speech recognition: Acoustic detection and linguistic context manipulation. *Speech Communication* 48: 1590–606. [CrossRef]

Swerts, Marc. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30: 485–96. [CrossRef]

Tottie, Gunnel. 2014. *Uh* and *um* in British and American English: Are they words? Evidence from co-occurrence with pauses. In *Linguistic Variation: Confronting Fact and Theory*. Edited by Nathalie Dion, André Lapierre and Rena Torres Cacoullos. New York: Routledge, pp. 38–54. [CrossRef]

Tottie, Gunnel. 2016. Planning what to say: *Uh* and *um* among the pragmatic markers. In *Outside the Clause. Form and Function of Extra-Clausal Constituents*. Edited by Gunther Kaltenböck, Evelien Keizer and Arne Lohmann. Amsterdam: John Benjamins, pp. 97–122. [CrossRef]

Trouvain, Jürgen. 2014. Laughing, breathing, clicking—The prosody of nonverbal vocalisations. Paper presented at the Speech Prosody 2014, Dublin, Ireland, May 20–23; pp. 598–602.

Trouvain, Jürgen, Camille Fauth, and Bernd Möbius. 2016. Breath and non-breath pauses in fluent and disfluent phases of German and French L1 and L2 Read Speech. Paper presented at the Speech Prosody 2016, Boston, MA, USA, May 31–June 3; pp. 31–35. [CrossRef]

Watanabe, Michiko, Keikichi Hirose, Yasuharu Den, and Nobuaki Minematsu. 2008. Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication* 50: 81–94. [CrossRef]

Yaruss, J. Scott, Robin M. Newman, and Tracy Flora. 1999. Language and disfluency in non-stuttering children's conversational speech. *Journal of Fluency Disorders* 24: 185–207. [CrossRef]