International Journal of Linguistics
ISSN 1948-5425
2024, Vol. 16, No. 5

Sociophonetic Study of Dysfluent Behaviour in Native English Speakers

Renáta Longauerová

Department of English Language and ELT Methodology

Charles University, Prague, Czech Republic

Received: May 8, 2024 Accepted: June 20, 2024 Published: October 22, 2024

doi:10.5296/ijl.v16i5.21891 URL: https://doi.org/10.5296/ijl.v16i5.21891

Abstract

The present project focuses on the study of the influence of two social factors, age and gender, on the frequency and variation of seven different types of dysfluencies (filled pauses, silent pauses, repairs, repetitions, false starts, vowel and consonant lengthening) in the spontaneous speech of native English speakers from England. The first part of the present article provides a general characterization of the relevant types of dysfluencies, together with the main issues concerning the production of speech dysfluencies. The empirical part presents the analyses of the recordings of 32 native English speakers from England. The overall results considering general influence of age and gender show that the only significant difference is between age groups, with older speakers producing more dysfluencies than younger speakers. Gender, on the other hand, does not make the difference significant, except for vowel lengthenings, that were produced significantly more by female speakers than male speakers, and repetitions, which is the only type of dysfluency where the difference is significant and influenced by both age and gender, with older male speakers producing significantly more dysfluencies than any other group.

Keywords: Dysfluencies, Dysfluent behaviour, Native English speakers, Sociophonetics, Frequency, Variation, Age, Gender

1. Introduction

1.1 Dysfluencies in Spontaneous Speech

Since the focus of the present study is on frequency and variation of different types of dysfluencies appearing in the production of spontaneous speech, it seems important to discuss what the notion of spontaneous speech actually refers to. Generally speaking, spontaneous speech can be defined as a type of speech that occurs in everyday conversations without previous planning and practicing. That is an essential difference from something that Clark

(2014) calls manufactured varieties. In spoken language, manufactured varieties can refer to read speeches or the speeches of people such as public speakers or actors, who make their living on spoken word and need to practice beforehand in order to create fluent utterances (Fox Tree, 1995). The presence of a certain amount of dysfluencies is a characteristic feature of spontaneous speech, or in general, spontaneous varieties, which are basically utterances formulated on the fly (Clark, 2014). Dysfluencies can be generally defined as "any phenomenon originated by the speaker, which changes the flow of the speaker's utterance" (McDougall et al, 2015; pg.2), or as phenomena which interrupt the flow of speech without adding a propositional content to an utterance (Fox Tree, 1995). It has been proven that dysfluencies are affecting around six words in every hundred words (Fox Tree, 1995). There are several types of dysfluencies and in the present study we will deal with the following: filled pauses, empty pauses, repetitions, repairs, false starts, vowel and consonant lengthening. Another important issue which is a subject of many studies dealing with dysfluent behaviour is that of discourse markers, such as 'you know', 'I mean', 'like', 'well', and 'so', also called editing expressions or lexical fillers (e.g. Clark, 2002). However, in the present study, discourse markers will not be addressed.

1.2 Disfluencies and Communication

For a long time, dysfluencies were seen as "unwanted elements or unfortunate by-products of speaking on the fly" (Fox Tree, 2000; pg. 376). However, the latest studies show that dysfluencies play an important role in communication. As Braun & Rosin (2015) suggest, they are important indicators of verbal planning processes and of the monitoring of one's speech. Dysfluencies therefore form an essential part of all language production theories and for explanatory purposes I will now provide a brief description of the language production processes.

One of the most influential models of language production processes is the one proposed by Willem Levelt (1983, 1989). According to this model, we can distinguish between three different stages of speech production. The first stage, called conceptualization, involves the speaker's intention to plan their utterance and the result is the so called "pre-verbal message", which is just a set of ideas that form part of the mental model of what the speaker wants to say (Warren, 2012). The second stage is the so-called "formulation" and involves transformation of the pre-verbal message into a verbal form. This is done by grammatical encoding, i.e. by selecting the appropriate word forms and putting them together, and by phonological encoding, i.e. selection of the appropriate phonological and articulatory forms. The final stage is the actual articulation of the message.

Levelt (1989) also points out that speakers can be simultaneously their own listeners, and they are able to listen to their overt speech as they can listen to speech of their interlocutors. This process involves an audition component and speech comprehension system, which speakers use to interpret their own speech sounds and transform them into meaningful words and sentences. The output of this process is the so-called parsed speech, i.e. an analysed string of words which form a sentence structure (Warren, 2012). But what happens if speakers experience difficulties while producing their speech? What happens if they have problems to

plan the pre-verbal message, or problems to select the appropriate word, phonological or articulatory forms? And what happens if there is an error in speech that was already articulated?

As Levelt (1983, 1989) suggests there is a specific cognitive system which helps speakers to recognize an error at any stage of language production and it is called a "verbal self-monitoring system". This idea was elaborated as a part of Levelt's language production theory, and it proposes that speakers can attend to, or self-monitor, both internal (inner) speech, i.e. prearticulatory speech, and external (overt) speech, i.e. articulated speech. These errors are detected via a double perceptual loop which consists of two loops:

1) internal loop (for the perception of internal speech) that helps speakers to prevent errors occurring in inner speech from being articulated, and

2) external loop (for the perception of external speech) that helps speakers to repair any damage that was caused by errors already articulated (Nooteboom, 2004).

This suggests that dysfluencies are, as it was already mentioned, closely related to language production and are true indicators of verbal planning processes and self-monitoring processes. I will now focus on the characterization of the selected types of dysfluencies.

1.3 Different Types of Disfluencies

As pointed out by Braun & Rosin (2015), phonetic manifestations of dysfluencies are varied and there are many classifications of different types of dysfluencies. In the present study we will focus on the seven most frequently occurring ones, starting with filled pauses and silent pauses, commonly known as fillers.

Filled pauses

In the present study, we will deal with two types of filled pauses: a pause filled by insertion of a vowel ('uh') and a pause filled by insertion of a vowel and a nasal ('um'). There are several factors influencing the presence of filled pauses in the production of spontaneous speech, the most apparent one being the occurrence of some type of uncertainty on the side of the speaker (Corley, 2007) or the need to gain more time to plan the upcoming utterance (Levelt, 1989). However, it has been suggested that the occurrence of filled pauses is not as arbitrary as it may seem. Filled pauses usually tend to occur:

- more frequently before lexical words than before function words (Maclay and Osgood, 1959),

- before low-frequency words, i.e. words used less commonly, or less predictable target words, i.e. key words (Corley et al., 2007),

- before longer and more complex phrases (Watanabe, 2008),

- before objects newly introduced in the discourse, i.e. discourse-new objects, as opposed to discourse-given objects (Arnold et al., 2003), and

- usually rather at the beginning of the major constituents such as phrases, clauses, and sentences than in other positions (Watanabe, 2008).

But what do filled pauses signal? Both ‘uh’ and ‘um’ signal an upcoming delay (Clark, 1994; Fox Tree 2001) but some of the studies found out that the length of the delay depends on the preceding filler (e.g. Fox Tree, 2001). If the speaker’s utterance is preceded by the vocalic filler ‘uh’, the delay is going to be short. In the case of the nasal filler ‘um’, its presence signals that the upcoming delay is going to be long. As Fox Tree (2001) further asserts, different filled pauses might also have different effects on the on-line comprehension. Levelt (1983, 1989) suggests that filled pauses, as well as silent pauses, tend to occur as a part of a repair dysfluency which consists of three components. The first phase is called reparandum, and it corresponds to the item that needs to be repaired. However, as Li & Tilsen (2015) point out, dysfluencies that were caused by problems with word-retrieval, i.e. with accessing the word in one’s mental lexicon, do not contain reparandum, such as in the sentence we had – uh the dog first. After reparandum, there is a moment of interruption, which is the moment of the problem detection and the interruption of the speech flow (Li & Tilsen, 2015). The second component is optional and is called editing phase. Speakers can use filled pause, empty pause, or the phase is entirely skipped. The last component is called repair, and this is represented by the original target word, or, "the correct version of what was wrong before" (Levelt, 1983; pg. 44).

Silent pauses

Our speech is naturally filled with pauses which are associated with our respiratory system and which we make in order to breathe. While speaking, we also produce pauses that can be referred to as grammatical pauses, also called juncture pauses. These pauses mark boundaries between syntactic units such as phrases, clauses, and sentences (Cenoz, 1998) and are necessary for the production of intelligible speech. However, there are many pauses occurring in the middle of phrases, clauses, and sentences and these are considered to be non-grammatical. Also called non-juncture pauses (Cenoz, 1998), these are classified as a type of dysfluency. As previously mentioned, in co-occurrence with repairs, silent pauses, as well as filled pauses, are usually part of the editing phase of the three-component structure of a repair dysfluency suggested by Willem Levelt (1983). However, the real question is if there is a difference in the usage of filled pause and silent pause. Some of the findings suggest these characteristics that silent pauses share with filled pauses:

- they occur when the speaker needs to gain more time during language production (Tissi, 2000),

- they occur more frequently before lexical words than before function words (Maclay and Osgood, 1959), and

- they tend to occur before low-frequency words and less predictable words (Goldman-Eisler, 1961).

However, there is still a specific context in which the usage of filled pauses is preferable. Wingate (1984) suggests that when the speakers are aware of the upcoming difficulty during speech production, they usually use a filled pause. On the other hand, when the dysfluent speech is unplanned, they tend to use silent pauses.

Repetitions

Repetitions as a type of dysfluency occur when the speaker repeats words or phrases (Fox Tree, 1995), but as Maclay and Osgood (1959) point out, it is important to distinguish between repetitions that are semantically significant and can change the meaning of the utterance. This can be illustrated on an example provided by Maclay and Osgood: "I I saw a very very big boy". Both "I" and "very" are repeated, but only the repeated "I" can be considered as a dysfluency repetition. Repetition of "very" intensifies the following adjective "big" and thus changes its meaning. These repetitions will not be considered in the present study.

According to the research conducted by Maclay and Osgood (1959), it is usually function words rather than lexical words that tend to be repeated and they usually occur as antecedents to lexical words, by which we can assume that the most important function they serve is to provide time for selection of the right lexical item, the same function provided by pauses. As suggested by McDougall et al. (2015), in the production of spontaneous speech we can come across four different types of repetitions: part-word repetitions (1), whole word repetitions (2), phrase repetitions (3), and multiple repetitions (4). In the empirical part of the study, however, the focus will be on repetitions as a group, covering all four types.

Repairs

As it was already mentioned in the previous section, repairs, also called corrections, is a type of dysfluency usually formed out of three components: reparandum, editing phase, and repair. In the present study, we will cover two different types of repairs distinguished by Levelt (1983): covert and overt repairs. Covert repairs are those that happen before the actual overt articulation and thus do not change, delete, or add anything new to the utterance. Covert repairs are usually accompanied by a certain type of editing term (such as 'uh' or 'um'), e.g. "I saw, uh, twelve people at the party", or can be manifested by repetition of the same word, which might be preceded by an editing term, but does not have to, e.g. "go to red, red node" (Levelt, 1983). These will not be included in the empirical part of the present study. Overt repairs, on the other hand, fix an error after it was already articulated, e.g. "I am trying to lease, or rather, sublease my apartment" (Levelt, 1983) and in the present study I will deal only with those. As one can see, overt repairs can be accompanied with different types of discourse markers such as 'or rather', 'I mean', 'you know', 'pardon', 'sorry', 'no', or 'well' which serve as clues for the addressee that the speaker has an intention to correct a preceding item (Clark, 2002). According to Levelt (1983), overt repairs can be divided into three subgroups. The first type of overt repair occurs when speakers realize that the formulations of their ideas are not appropriate and thus choose more suitable forms. These repairs are called A-repairs, i.e. appropriateness-repairs. The second subgroup of overt repairs is called E-repairs, i.e. error repairs, which are produced when speakers realize that their utterance contain a certain error, e.g. phonetic, syntactic, lexical and even suprasegmental error. They are probably a result of the right input message but the activation of wrong lexical item. The third group is so-called D-repairs, and these occur when speakers realize that they should express another idea before the one that they already articulated and thus start again. However, this type of repair is quite infrequent, and in Levelt's corpus they represent only 1% of the total number of repairs. Levelt

also distinguishes between different subgroups of A-repairs, E-repairs, and D-repairs, but this distinction is too detailed and for the purpose of this study irrelevant. Moreover, in the empirical part of the present study, also the basic subgroups of A-repairs, E-repairs, and D-repairs will be unified under the term repairs.

False starts

There are several studies that can serve as evidence of the fact that false starts and repairs are basically the same phenomenon (see Maclay & Osgood, 1959; Levelt, 1983, 1989). However, the study by Maclay and Osgood (1959) implicitly proves that there is a certain difference between these two terms. For the authors, false starts are "all incomplete or self-interrupted utterances," incomplete being non-retracted false starts (1) and self-interrupted being retracted false starts (2).

(1) I saw a very ...

(2) I saw a very big || very small boy

From the examples provided, we can see that the difference between these two clauses is in the speaker's attempt to correct the wrong word. Example (2) looks like a repair dysfluency (as discussed earlier in this section). Therefore, in the present study, retracted false starts will be included under the term "repairs", and the term "false starts" will be used only for non-retracted "false starts".

Vowel and consonant lengthening

Lengthening of phonemes is another type of dysfluency often to be found in the course of the production of spontaneous speech. As Clark (2002) points out, prolongations in both function and content words are mostly used by speakers to mark a temporary suspension to deal with a production problem, a function that they share with both filled pauses and silent pauses. The most common form of lengthening is lengthening of reduced vowels, such as [ə] in words like 'to' or 'the', which consequently become a non-reduced vowel and is pronounced as [tu:] and [ðɪ:] (Clark, 2002). However, lengthening can affect any speech sound and in the present study, I will therefore focus on lengthening of vowels and consonants in any position. The criterion I use to distinguish between phonemes of normal length and lengthened phonemes is adopted from the study by McDougall et al. (2015): if the duration of the phoneme is ≥ 200 ms, the phoneme will be considered to be lengthened.

1.4 Hypotheses

As it was previously mentioned, speakers are generally dysfluent when they experience some difficulties with language production. They can monitor them either before the actual articulation, which suggests problems with conceptualization or formulation, or after the production of the overt speech. But what are some other factors that can influence the occurrence of dysfluencies in spontaneous speech of speakers? The empirical part of this study examines the influence of two social factors, age and gender, on the frequency and variation of selected dysfluencies in the speech of native English speakers. We distinguish between four different social groups: young female speakers, young male speakers, older female

speakers, and older male speakers. According to the previous studies examining the effects of age and gender on speech and production of dysfluencies, we propose three different hypotheses:

H1: Older speakers are more dysfluent than younger speakers

H2: Male speakers are more dysfluent than female speakers.

H3: The most dysfluent group is the group of older male speakers

2. Method

Recordings of 32 native English speakers were obtained in two ways: most of the participants were recorded in the sound-proof studio of the School of Psychology and Clinical Language Sciences at the University of Reading, using the software Audacity and an AKG D80 studio microphone. Other participants, who were not able to be physically present at the studio, were recorded by using a portable handheld recorder Tascam DR-07mkll. In both cases we used the sampling rate of 48-kHz. 32 native English speakers from different areas of England currently living in the South East region or The Greater London in the period of data collection, were chosen according to affiliation to different social groups based on their gender and age, forming 4 groups of 8 participants: male speakers aged 17-30, female speakers aged 17-30, male speakers aged 35-72, and female speakers aged 35-68. To test their spontaneous speech, the recording consisted of a simple interview, designed to make the participants comfortable and forget they were being interviewed. We asked questions about their hometowns, childhood and school memories, the place they live in, their jobs, as well as their hobbies and interests.

2.1 Data Processing

Since the recordings were of different lengths, we orthographically transcribed only five-minute sections of all the recordings using a conventional English orthography. The recordings were subsequently analysed by using a computer programme designed for phonetic analyses, Praat, version 5.3.56 (Boersma & Weenink, 2013). In order to establish boundaries between the segments of speech, we consulted Machač & Skarnitzl (2009). The following dysfluencies were identified: filled pauses, empty pauses, false starts, repairs, repetitions, vowel and consonant lengthening. After we coded all types of dysfluencies, we ran two Praat scripts: the first script was designed to extract the total amount of dysfluencies for every single speaker into Microsoft Excel tables and figures. The second script was designed to calculate the total amount of words pronounced by speakers by excluding words pronounced by the interviewer. Afterwards, we used Microsoft Excel to count the amount of dysfluency rate per every 100 words, which also included filled pauses, repeated words, and fragmented words, by applying the following formula:

total amount of dysfluencies / total amount of words x 100

Using this method, we obtained all the counts necessary to run the statistical analyses and test the statistical significance of our results.

2.2 Statistical Analysis

The present study is focused on the comparison of four different social groups and this type of comparison and the statistical significance of the difference between the groups are usually tested by using the statistical method of analysis of variance (ANOVA). We also used ANOVA to examine differences in the usage of every single type of dysfluency.

However, to better understand which factor is driving the main results, we also ran so-called post hoc t-tests and examined differences between two same age groups but of different gender, and two same gender groups but of different age.

As suggested by Volín (2007), results can be considered statistically significant if the value of p is lower than 0.05, i.e. $p < 0.05$. However, values of p may differ. If the values of p are lower than 0.001, i.e. $p < 0.001$, the results are considered highly significant, while the values found between $0.05 < p < 0.1$ are referred to as marginally significant. The next chapter will reveal if there are any statistically significant differences among the groups tested as well as among the groups' preferences for specific dysfluencies.

3. Results

First, we will look at the dysfluency rates of all four social groups and examine the effects of both age, i.e. differences between older and younger speakers, including both male and female speakers, and gender, i.e. differences between male and female speakers, including both older and younger speakers. We will also look at the interaction between gender and age. Second, we will take a look at the results of the post-hoc tests and compare the differences between the speakers of the same age group but of different gender, as well as those between the speakers of the same gender group but of different age. Last, we will compare the results of ANOVA we ran for every type of dysfluency and also look at group and individual preferences.

3.1 Overall Results

All dysfluency rates I consider in the present study are per 100 words, including filled pauses, repeated words, and fragmented words. When considering the influence of age, the ANOVA test, which was performed by comparing dysfluency rates of all the members of the two groups, shows that older speakers produced more dysfluencies than younger speakers and the difference between the two age groups was proven to be statistically significant: $F(1,30) = 4.46$; $p < 0.05$. On the other hand, the difference between the two gender groups is not statistically significant: $F(1,30) = 1.37$; $p > 0.2$. We also ran ANOVA to measure the interaction between gender and age, i.e. how the effect of one variable changes in relation to the other variable. If the interaction were significant, the amount of dysfluencies produced by the two genders would vary at different ages. For instance, female speakers might produce more dysfluencies than male speakers when they are younger, but when they are older, the pattern might change and it is no longer female speakers producing more dysfluencies, but male speakers. However, the results proved that the interaction is not statistically significant: $F(1,28) = 0.97$; $p > 0.3$. This suggests that the pattern of dysfluent behaviour when comparing genders is same across ages.

As mentioned above, we also ran so-called post hoc t-tests to examine differences between two same age groups but of different gender, and two same gender groups but of different age groups. We found out that the age effect is driven only by male speakers: older male speakers produced more dysfluencies than younger male speakers and the difference between the two groups was proven to be statistically significant $t(14) = 2.07$; $p = 0.05$. The results show that women, on the other hand, do not drive the age effect. Even though older female speakers were more dysfluent than younger female speakers, the difference between these two groups is not statistically significant: $t(14) = 0.85$; $p = 0.4$. Considering gender, there is an overall absence of gender effect. The difference between the groups of younger male speakers and younger female speakers is very small and thus it is shown to be not statistically significant: $t(14) = - 0.19$; $p = 0.85$. When comparing the groups of older male speakers and older female speakers, older men were slightly more dysfluent than older women but the difference between these two groups is not statistically significant: $t(14) = - 1.22$; $p = 0.24$.

3.2 Influence of Gender and Age on Individual Types of Dysfluencies

In the previous general analysis, both ANOVA and post-hoc t-tests were considering dysfluency rates by summing up all the types of dysfluencies included in the present study. However, I also ran a series of ANOVA tests to investigate the effects of gender and age on different types of dysfluencies in isolation. Among the seven types of dysfluencies I examined in the present study, there are four dysfluencies whose usage revealed to be influenced either by gender or age. There were no statistically significant differences in the production of repairs, filled pauses, and silent pauses, but there were statistically significant differences for false starts, consonant lengthening, vowel lengthening and repetitions.

Differences in the usage of false starts and vowel lengthening are, in relation to the overall results, much unexpected. The overall results show that older speakers produce more dysfluencies than younger speakers and the difference between the two groups was found to be statistically significant. Here, the results show that false starts are produced the most by the group of younger speakers with the difference being statistically significant: $F(1,28) = 4.77$; $p < 0.05$. I already mentioned that both ANOVA and post-hoc tests proved an overall absence of gender effect. However, vowel lengthening and repetitions are the only types of dysfluencies where the difference in the production is influenced by gender. Regarding vowel lengthening, it was the group of female speakers that produced more than the group of male speakers. The difference was proven to be statistically significant: $F(1,28) = 3.82$; $p < 0.1$.

The difference in the production of consonant lengthenings is, on the other hand, influenced by age. The group of older speakers produced a higher amount of consonant lengthenings than the group of young speakers, with the difference being statistically significant: $F(1,28) = 7.11$; $p < 0.05$. The last type of dysfluency, repetitions, is the only type influenced by both gender and age. The group producing the highest amount of repetitions is older male speakers. Older speakers produced a significantly higher amount of repetitions than the group of younger speakers: $F(1,28) = 5.35$; $p < 0.05$ and males speakers produced a significantly higher amount of repetitions than the group of female speakers: $F(1,28) = 3.75$; $p < 0.1$.

3.3 Individual Preferences

As it was previously stated, all the cognitive processes connected with language production cannot be consciously controlled and are speaker-specific (Braun & Rosin, 2015), which suggests that the production of dysfluencies and their variation are also speaker-specific and different individuals would use different patterns of dysfluencies (e.g. Maclay & Osgood, 1959; Shriberg, 2001; King et al., 2013; Braun & Rosin, 2015; McDougall et al., 2015).

However, as we also mentioned, the combination of different social factors plays an important role in the variation of the occurrence of dysfluencies, as well as in language variation in general. To see if there are any outstanding individual preferences, I tried to even out all the factors (regional, socio-economic, and educational) as much as possible so results would not be compromised in any way.

In this section, I will look at the dysfluencies as they were produced within the selected five-minute sections. Even though the utterances still differ in lengths (speech rate and thus the amount of words pronounced varies among speakers), from the figures provided we can still understand the individual preferences. In the first part I will consider the female speakers and in the second part) will look at the preferences of the male speakers. Older speakers will be marked with (o), younger speakers with (y).

Regarding female speakers, neither F04 (o) nor F13 (o) produced any false starts, and the speakers F02 (y), F12 (o), and F15 (o) used only one in the chosen 5-minute section. In comparison with other types of dysfluencies, false starts actually proved, together with corrections and repetitions, to be the least common type of dysfluency among women as a group. Instead, the most common dysfluencies are lengthenings, specifically vowel lengthening, and pauses, silent pauses more than filled. One speaker partially violated this pattern: the speaker F10 (o), strongly preferred consonant lengthening over vowel lengthening, using it 42 times in the 5-minute section, making the rate of 9.40 per 100 words. This particular speaker also used much more filled pauses than silent pauses. She produced 26 silent pauses, which makes the rate of 5.81 silent pauses per 100 words, while 35 filled pauses creating a rate of 7.83 per 100 words (including both types). It also should be stated that among filled pauses, women used in general more of ‘um’ filler than ‘uh’ filler. Some female participants displayed a much skewed preference in this direction: both F04 (o) and F13 (o), for instance, used only ‘um’ and did not use any ‘uh’. The only exceptions were speakers F05 (y), F06 (y), and F16 (y), who produced a slightly higher amount of ‘uh’. False starts and repairs are the least produced dysfluencies also by male speakers. In fact, M12 (o) did not produce any false starts and M15 (o) did not produce any repairs. On the other hand, the most common dysfluencies are lengthening, specifically consonant lengthening, silent pauses, and the majority of male speakers produced also a fair amount of filled pauses. The highest amount of silent pauses were produced by M01 (y), M11 (o), M13 (o) and also M09 (y), whose dysfluency rate for this specific type of dysfluency was the highest out of all male participants and also in comparison with other dysfluency produced by himself. In his speech, he produced 100 silent pauses, which makes the rate of 11.26 of pauses per 100 words. M11 (o) also produced a fair amount of repetitions, specifically 39, which makes the rate of 5.71

of repetitions per 100 words and is thus the highest amount of repetitions among all male participants. Regarding filled pauses, all male speakers generally produced more of a filler 'um'. Speaker M7 (y) used exclusively this type of filler and none of the 'uh' filler. Only four speakers, M1 (y), M6 (y), M11 (o), and M12 (o) used more 'uh' fillers than 'um', M11 (o) using it notably more than any other male participant, specifically 45 times in the course of his 5-minute section, which makes the dysfluency rate of 6.60 of fillers 'uh' per 100 words.

4. Discussion

In the first part of the research, I looked at the influence of age, gender and the interaction between these two factors. I compared younger and older speakers, including both male speakers and female speakers, and then I compared male speakers and female speakers, including both age groups. The results show that older speakers are more dysfluent than younger speakers and the difference between these groups is statistically significant, which confirmed my first hypothesis. This proved to be the case particularly for consonant lengthening and repetition dysfluencies. However, there is one exception among all the types of dysfluencies and that is false starts. This type of dysfluency was produced more by younger speakers and the difference was also proven to be significant.

The second hypothesis, however, was not confirmed. Even though the dysfluency rate for male speakers was slightly higher, the difference between them and the group of female speakers was not statistically significant. The only two types of dysfluencies which were influenced by gender were vowel lengthening and repetition and the difference was found to be significant. In the case of vowel lengthening, it was surprisingly the group of female speakers who were driving this result. Repetitions, on the other hand, were produced more by the group of male speakers, and, as we mentioned above, also by older speakers. From this we can conclude that the usage of repetition is the only type of dysfluency, which is significantly influenced by both age and gender, specifically by older men. For this type of dysfluency, my last hypothesis was confirmed. I also tested the interaction between gender and age, which was not found to be significant.

The results from the post hoc tests showed that the age effect is found only when comparing male speakers. The difference between older male speakers and younger male speakers was proven to be statistically significant, while the difference between female speakers was not influenced by their age and was not significant. Gender effect was, once again, proven to be missing. There was almost no difference between younger male speakers and younger female speakers and the slight difference found between older male speakers and older female speakers was not significant.

The reasons behind these results might be difficult to determine. As we mentioned in the theoretical part, there is no reason to assume why male speakers should be more dysfluent than female speakers. I offered a suggestion made by Shriberg (1996), who says that this might be due to men's desire to hold the floor of the conversation and speak to a person or a group of people for a long time without allowing them to take turn.

However, this might be true only for everyday conversations, where both sides get a chance to ask and answer the questions. In this study, we examined spontaneous speech of our participants by asking them simple questions about their lives, so the cues for their turn were very obvious. Participants were also aware of the fact that they could take as much time as needed to answer the interviewer's question without worrying of being interrupted. Even though I mentioned some studies that proved men to be more dysfluent and our study found a difference too, statistics shows that this difference is not significant. However, there was a significant difference in the usage of repetitions, which was driven by our male speakers. The same results were found also in the study by Bortfeld et al. (2001), where the difference in the production of dysfluencies between male and female speakers was mostly due to the production of repetitions, in their case also fillers.

The reason for this phenomenon is probably due to the fact that repetitions, as mentioned in the theoretical part, provide speakers with some extra time to select the following word, especially when it comes to repetitions of functional words which occur as antecedents to lexical words (Maclay & Osgood, 1959). This function is also shared with other types of dysfluencies such as pauses, both filled and silent, and vowel and consonant lengthening and since the results show that female speaker indeed produced significantly more vowel lengthenings and male speakers consonant lengthenings, we can presume that there might be a gender preference for a specific type of dysfluency when in need for some extra time to retrieve the word. However, to be truly able to determine the real cause of the preference for this specific type of dysfluency, we might need to consider the possibility of the presence of other factors, for example, a higher level of anxiety, not being familiar with the interviewer, as well as the opposite gender of the interviewer. The interaction of other factors thus remains a relevant issue for future research.

The reason for the age being a factor that is statistically significant is probably due to difficulties with the word retrieval that comes with ageing. The results show that the significance is driven mostly by consonant lengthening and, once again, repetition. The main function of both dysfluencies is, indeed, as mentioned above, to provide speakers with some extra time when they are experiencing difficulties to select the following word. Yet, there was one exception when considering age effects: false starts. The results show that it is, in fact, the younger speakers who produce more of them and this difference in the usage of false starts was also proven to be significant. One possible explanation for this phenomenon could be that apart from causing some difficulties with word retrieval, ageing was proven to improve people's speech and since their vocabulary is richer and their conceptual development increased, they might have a clearer idea about the content they want to communicate without producing self-interrupted and unfinished utterances.

In conclusion, we can say that the production of dysfluencies was found to be affected by age but not gender. There are specific types of dysfluencies, however, for which this statement proves to be incorrect.

5. Conclusion

The aim of the present research was to study the influence of two social factors, age and gender, on the frequency and variation of seven most common types of dysfluencies (filled pauses, silent pauses, repetitions, repairs, false starts, and vowel and consonant lengthening) on spontaneous speech of native English speakers from England. We presented 3 hypotheses:

H1: Older speakers are more dysfluent than younger speakers

H2: Male speakers are more dysfluent than female speakers.

H3: The most dysfluent group is the group of older male speakers.

The first hypothesis was confirmed with a statistically significant difference between older speakers and younger speakers. The second hypothesis was not confirmed, as there was no statistical difference between male speakers and female speakers in general. The third hypothesis was confirmed only with one type of dysfluencies: repetitions.

To my knowledge, this is the first study that includes all the most common types of dysfluencies in one analysis. Most of the studies tend to focus only on one or two types of dysfluencies. Bortfeld et al. (2001) included 4 types, however, one of them was editing expressions (e.g. I mean, rather, that is), often referred to as lexical fillers or discourse markers, which we decided to omit from our study. The reason behind is that these kinds of expressions carry meaning and fulfil a variety of very specific pragmatic functions that are beyond the functions of the phonetic dysfluencies presented in this study.

Branigan et al. (1999) included more types of dysfluencies (repeats, deletes, inserts, substitutes and their combinations, plus words in the reparandum and filled pauses). However, they concentrated on the overall measures of dysfluency and did not distinguish between different categories, which they themselves considered to be simplistic and presented it as one of their limitations. This study, on the other hand, provides an elaborate overview of the most common types of dysfluencies as individual phenomena, looks at the group preferences as well as individual preferences and discusses the potential factors that may influence these preferences.

6. Limitations and Future Research

Even though our findings are consistent with some findings from other studies, there are certain limitations to it. For example, Bortfeld et al. (2001) studied the influence of same social factors as I did in my study and they proved that older speakers were more dysfluent than younger speakers. However, the age range of the participants included in my groups of older speakers, both male and female, was probably too wide. Some of the speakers from both older male speakers and older female speakers groups could be easily classified as ‘middle-aged.’ The reason behind the inclusion of these speakers in the older speakers groups was simply the lack of participants. Future studies focusing on the influence of age on the production of dysfluencies should undoubtedly include enough participants to create 3 testing groups of young speakers, middle-aged speakers, and older speakers. There might be a

significant change in the production of dysfluencies between middle-aged speakers and older speakers that we might be missing in our research.

Another potential limitation of our study is the design of the task. As mentioned in our methodology, we wanted to test the production of dysfluencies in spontaneous speech and therefore prepared simple questions about participants' hometowns, childhood and school memories, the place they live in, their jobs, as well as their hobbies and interests. These questions were intended to make it feel less like an interview and more like a casual chat between friends, so the participants feel comfortable and forget they were being interviewed. However, most of the participants were not familiar with the interviewer prior to the recording session. Considering that some of the participants might not have particularly positive childhood or school memories, they might not have felt comfortable sharing them with a stranger, which, consequently, could have affected the fluency of their speech and causing many more dysfluencies than they would under different circumstances or if they were asked different questions. A task that is a little more structured and omits potentially sensitive topics might be a better choice for the future research.

References

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies Signal Theee, Um, New Information. *Journal of Psycholinguistic Research, 32*(1), 25-36. https://doi.org/10.1023/A:1021980931292

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.56.). Retrieved from http://www.praat.org/

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech, 44*(2), 123-147. https://doi.org/10.1177/00238309010440020101

Braun, A., & Rosin, A. (2015). On the Speaker-Specificity of Hesitation Markers. *Proceedings of the 18th International Congress of Phonetic Sciences. (ICPhS 2015).* Glasgow: University of Glasgow.

Cenoz, J. (1998). Pauses and Communication Strategies in Second Language Speech. Retrieved from http://files.eric.ed.gov/fulltext/ED426630.pdf

Clark, H. H. (2002). Speaking in Time. *Speech Communication, 36,* 5-13.

Clark, H. H. (1994). Managing problems in speaking. *Speech Communication, 15,* 243-250.

Clark, H. H. (2014). Spontaneous discourse. In Goldrick, M., Ferreira, V., & Miazzo, M. (Eds.), *The Oxford Handbook of Language Production* (pp. 292-308). Oxford: Oxford University Press.

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way you, er, say it: Hesitations in speech affect language comprehension. *Cognition, 105,* 658-668.

Fox Tree, J. E. (1995). The Effects of False Starts and Repetitions on the Processing of

Subsequent Words in Spontaneous Speech. *Journal of Memory and Language, 34,* 709-738.

Fox Tree, J. E. (2000). Coordinating spontaneous talk. In Wheeldon, L. R. (Ed.), *Aspects of Language Production* (pp. 375-406). Philadelphia: Psychology Press.

Fox Tree, J. E. (2001). Listeners' uses of *um* and *uh* in speech comprehension. *Memory & Cognition, 29*(2), 320-326. https://doi.org/10.3758/BF03194926

Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech, 4,* 18-26. https://doi.org/10.1016/j.cognition.2006.10.010

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41-104.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.

Li, J., & Tilsen, S. (2015). Phonetic evidence for two types of disfluency. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015).* New York: Cornell University.

Maclay, H., & Osgood, C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. *Word, 15,* 19-44.

McDougall, K., Duckworth, M., & Hudson, T. (2015). Individual and Group Variation Dysfluency Features: A Cross-Accent Investigation. *Proceedings of the 18th International Congress of Phonetic Science (ICPhS 2015).* Glasgow: University of Glasgow.

Nooteboom, S. (2004). Self-Monitoring of Inner Speech and Over Speech. In *Poster session abstract included in the conference proceedings of From Sound to Sense, 50.*

Shriberg, E. (1996). Disfluencies in Switchboard. Addendum. *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 1996)* (pp. 11-14). Philadelphia.

Shriberg, E. (2001). To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association, 31*(1), 153-169. https://doi.org/10.1017/S0025100301001128

Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. In Riccardi, A., & Viezzi, M. (Eds.), *The Interpreters' Newsletter* (pp. 103-128). Trieste: Scuola Superiore di Lingue Moderne per Interpreti e Traduttori.

Vol ń, J. (2007). *Statistick émetody ve fonetick ém v ýzkumu.* Praha: Epocha.

Warren, P. (2012). *Introducing Psycholinguistics.* Cambridge: Cambridge University Press.

Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of the upcoming phrases of native and non-native listeners. *Speech Communication, 50,* 81-94. https://doi.org/10.1016/j.specom.2007.06.002

Wingate, M. E. (1984). Pause loci in stuttered and normal speech. *Journal of Fluency*

Disorders, *9*(3), 227-235.

Copyrights