

Univerzita Karlova v Praze

Filozofická fakulta

Ústav anglického jazyka a didaktiky

BAKALÁŘSKÁ PRÁCE

Renáta Longauerová

**Sociofonetická studie dysfluentního chování u rodilých mluvčích
angličtiny**

**Sociophonetic study of dysfluent behaviour in native English
speakers**

Praha, 2016

Vedoucí práce: doc. Mgr. Radek Skarnitzl, Ph.D.

Acknowledgements

First and foremost, I would like to express my deepest gratitude and appreciation to my supervisor, doc. Mgr. Radek Skarnitzl, PhD., for all his guidance, valuable comments, and encouragement, as well as for his immense help with the process of data analysis.

My thanks also go to Dr Vesna Stojanovik from the School of Psychology of the University of Reading for allowing me to use their studio for the process of recording the speech material, to Dr Linda Shockey from the Department of English Language and Applied Linguistics of the University of Reading for providing me with the handheld recorder, and to all the native English speakers who were willing to participate and allowed me to record their speech.

Last but not least, I would like to thank all my family and my friends for their incredible support throughout.

Declaration of authorship

Herby I declare that the following BA thesis is my own work for which I used only the sources and literature mentioned.

Prague, August 13, 2016

Signature

Abstract

The present thesis 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 speakers were divided into four different social groups according to their age and gender. The first part of the present thesis 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.

Key words: dysfluencies, dysfluent behaviour, native English speakers, sociophonetics, frequency, variation, age, gender

Abstrakt

Táto práca je zameraná na výskum vplyvu dvoch sociálnych faktorov, veku a pohlavia, na frekvenciu výskytu a variabilitu siedmich typov dysfluencií (vyplnené pauzy, tiché pauzy, opakovania, opravy, falošné začiatky, predlžovanie samohlások a predlžovanie spoluhlások) v spontánnej reči rodených hovoriacich angličtiny z Anglicka. Účastníci boli rozdelení do štyroch skupín podľa veku a pohlavia. Prvá časť práce ponúka všeobecnú charakteristiku všetkých siedmich typov dysfluencií spolu s hlavnými témami týkajúcich sa tvorby dysfluencií. Praktická časť prináša analýzu 32 nahrávok rodených hovoriacich angličtiny. Výsledky zameriavajúce sa na všeobecný vplyv veku a pohlavia ukazujú, že významný rozdiel sa nachádza jedine medzi vekovými skupinami, kde starší sú tí, ktorí produkujú viac dysfluencií. Pohlavie, na druhej strane, nespôsobuje žiadny významný rozdiel, s výnimkou predlžovania samohlások, kde rozdiel je významný, s výskytom vyšším u žien, a taktiež opakovanie, čo je jedinou dysfluenciou, kde je rozdiel významný u veku aj pohlavia, s výskytom vyšším u starších mužov.

Kľúčové slová: dysfluencie, dysfluentné chovanie, sociofonetika, rodení hovoriaci angličtiny, frekvencia výskytu, variabilita, vek, pohlavie

Table of contents

1. Introduction	7
2. Theoretical part	8
2.1 Dysfluencies in spontaneous speech	8
2.1.1 What are dysfluencies?	8
2.1.1.1 Language production.....	9
2.1.1.2 Trouble detection.....	10
2.1.2 Different types of dysfluencies	11
2.1.2.1 Filled pauses and silent pauses.....	11
2.1.2.2 Repetitions, repairs, and false starts	14
2.1.2.3 Vowel and consonant lengthening	16
2.2 Production of dysfluencies	16
2.2.1 General factors influencing dysfluent behaviour	16
2.2.2 Functions of dysfluencies.....	19
2.2.3 Dysfluencies and their effects on listeners.....	21
2.3 Dysfluencies in a sociophonetic context	23
2.3.1 What is sociophonetics?.....	23
2.3.2 Social factors conditioning language variation.....	24
2.3.3 Influence of gender and age	29
2.4 Hypotheses	31
3. Method and material	32
3.1 The process of recording and the participants.....	32
3.2 Data analyses	32
3.3 Statistical analyses.....	33
4. Results	35
4.1 Overall results	35
4.2 Influence of gender and age on individual types of dysfluencies	38
4.3 Individual preferences	39
5. Discussion	43
6. Conclusion	46
Bibliography	48
Zhrnutie.....	53

1. Introduction

The occurrence of various types of dysfluencies is a natural part and pervading feature of spontaneous speech. Dysfluencies are usually described as any phenomenon which interrupts the flow of speech without adding a propositional content to the utterance (Fox Tree, 1995). There are many questions accompanying the investigation of speakers' dysfluent behaviour. Some of the studies try to determine the functions or the effects of different types of dysfluencies, while other studies focus on general factors that might have a potentiality to influence dysfluent behaviour. Among them, there are many social factors that influence the frequency of their occurrence as well as their variation.

The first section of the theoretical part presents general characterisations of the seven types of dysfluencies we consider in the present thesis: filled pauses, silent pauses, repetitions, repairs, false starts, vowel and consonant lengthening. The next section provides descriptions of the main issues concerning production of dysfluencies, such as general factors influencing dysfluent behaviour, functions of dysfluencies, as well as their effects on listeners. The last section deals with the concept of sociophonetics and different social factors influencing dysfluent speech as well as language variation in general. Two of these social factors, gender and age, are factors studied in the present thesis and therefore dealt with in a separate subsection.

The empirical part of the present thesis is devoted to the description of the process of recruiting the participants, the recording, and the analysis of the speech material – 32 recordings of native English speakers from England divided into 4 social groups according to their gender and age. The aim of this research is to investigate the frequency and variation of seven types of dysfluencies occurring in the speakers' spontaneous speech. The problem is approached from the sociophonetic point of view, and as mentioned before, two important factors influencing language variation, gender and age, are considered. To ascertain whether the influence of gender and age on the difference in the usage of the seven types of dysfluencies in the speakers' spontaneous speech is significant, several statistical analyses were carried out.

2. Theoretical part

2.1 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, this 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).

2.1.1 What are dysfluencies?

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 seven types of dysfluencies: 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 be disregarded.

For a long time, dysfluencies were often seen only 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 the communication. As Braun & Rosin (2015) suggest, they are important indicators of verbal planning processes and of monitoring of one's speech. Dysfluencies therefore form an essential part of all language

production theories and for the explanatory purposes we will now provide a brief description of all the language production processes.

2.1.1.1 Language production

One of the most influential models of language production processes is the one constructed 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 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 called *formulation* and involves transformation of the pre-verbal message into a verbal form. This is done by grammatical encoding, i.e. selection of 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. The Levelt's language production model is demonstrated in Figure 1.

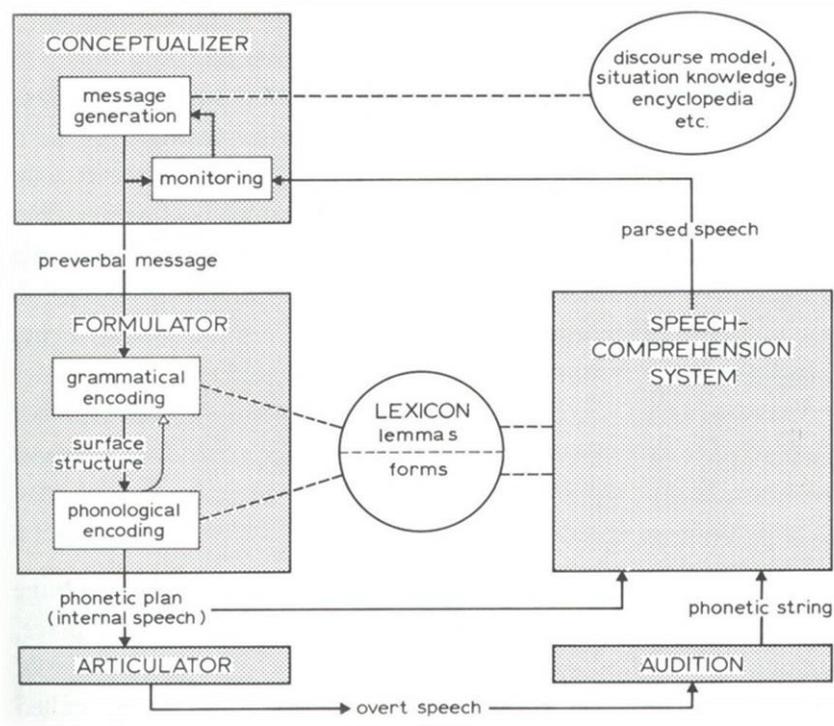


Figure 1. Levelt's model of speech production (Levelt, 1989; pg. 9)

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. As we can see from Figure 1 above, this involves an *audition* component and *speech comprehension system*, which speakers use to interpret their own speech sounds into meaningful words and sentences. The output of this process is 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 appropriate word, phonological or articulatory forms? And what happens if there is an error in speech that was already articulated?

2.1.1.2 Trouble detection

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*. It 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).

The actual model of double perceptual loop in relation to Levelt's language production model is demonstrated in Figure 2.

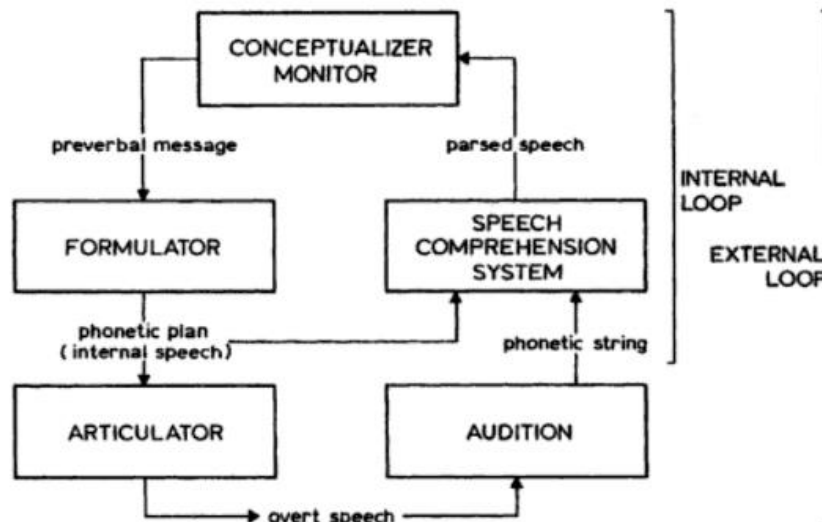


Figure 2. Double perceptual loop in the relation to the language production model (Levelt, 1989; pg. 470)

This proves that dysfluencies are, as it was already mentioned, are closely related to language production and are true indicators of verbal planning processes and self-monitoring processes. However, apart from that they serve other important function in the communication, but these will be discussed in subchapter 2.2 together with general factors influencing dysfluent behaviour and effects dysfluencies have on listeners. Now we will focus on the characterization of the selected types of dysfluencies.

2.1.2 Different types of dysfluencies

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

2.1.2.1 Filled pauses and silent pauses

Filled pauses

In the present study, we will deal with two types of filled pauses: a pause filled by insertion of vowel ('uh') and a pause filled by insertion of vowel and nasal ('um').¹ There are several

¹ Filled pauses, silent pauses and also vowel and consonant lengthening are often unified under the umbrella term *hesitations* (Collard, 2009).

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 when the speaker needs to gain more time to plan the upcoming utterance (Levelt, 1989). However, it has been suggested that the occurrence of filled pauses is not that arbitrary as it may seem. They 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 signalize? Both ‘uh’ and ‘um’ signalize 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’, it 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. However, this will be covered in an individual chapter dealing with effects of dysfluencies, i.e. the perception of dysfluencies from the listener’s perspective (see section 2.2.2).

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 (see Figure 3).

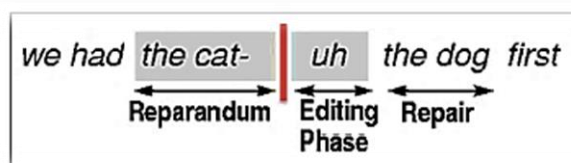


Figure 3. Levelt’s model of a repair dysfluency (Li & Tilsen, 2015; pg. 2)

As we can see from Figure 3, the first phase is called *reparandum*, the item that needs to be repaired. However, as Li & Tilsen (2015) point out, dysfluencies that were caused by the 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). On the figure it is indicated by the red vertical line. 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 represents the original target word, or, "the correct version of what was wrong before" (Levelt, 1983; pg. 44). Repairs as a type of dysfluency will be considered later in this section (see subsection 2.1.2.2).

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*. They 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) and 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 the language production (Tissi, 2000),
- more frequently before lexical words than before function words (Maclay and Osgood, 1959), and
- 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.

2.1.2.2 Repetitions, repairs, and false starts

Repetitions

Repetitions as a type of dysfluency occur when the speaker repeats words or phrases (Fox Tree, 1995), but as Maclay & 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 & Osgood: *I I saw a very very big boy*. Both *I* and *very* are repeated, but only repeated *I* can be considered as a dysfluency repetition. Repetition of *very* intensifies following adjective *big* and thus changes its semantic meaning. These repetitions will not be considered in the present study.

According to the research conducted by Maclay & Osgood (1959), it is rather function words² 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* (see subsection 2.1.2.1). 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 with 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 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 we will deal only with those. As we can see, overt repairs can be accompanied with different types of discourse markers such as ‘or

² E.g. subject personal pronouns, possessive pronouns, articles, prepositions, or numbers.

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, in Levelt's corpus they represent only 1%. 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 thesis, 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 an evidence of the fact that false starts and repairs are basically the same thing (see Maclay & Osgood, 1959; Levelt, 1983, 1989). However, the study by Maclay & Osgood (1959) implicitly proves that there is a certain difference between these two terms. For them, 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 that we already talked about earlier in this chapter. Therefore, in the present study, retracted false starts will be included under the repairs, and the term false starts will be used only for non-retracted false starts.

2.1.2.3 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 would be pronounced as [tu:] and [ði:] (Clark, 2002). However, lengthening can affect any speech sound and in the present study, we will therefore focus our attention on lengthening of vowels and consonants in any position. The criterion we 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.

2.2 Production of dysfluencies

In the previous subchapter, we described the main characteristics of seven types of dysfluencies, as well as their typical positions within the utterance. This subchapter will deal with some important issues regarding the actual production of dysfluencies. First, we will consider some of the main factors that might have a potentiality to influence dysfluent behaviour of speakers. After general factors we will focus on the main functions that dysfluencies may have and at the end we will consider the issue of dysfluencies from the perspective of the listener, i.e. what are the effects of dysfluencies on listeners’ judgements of speakers, listeners’ attention and processing of the utterance.

2.2.1 General factors influencing dysfluent behaviour

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?

Psychological factors

Finding ourselves in situations that make us stressed and tense is always unpleasant but often unavoidable. These situations may include those challenging ones such as speaking in public or being interviewed. We tend to care a lot about our performance and the impression we leave. This type of stress can affect our speech in many ways and speech full of dysfluencies may be one of the by-products. Moreover, different types of speeches also put on speakers different cognitive effort. For example, it was suggested that public speaking is correlated with high cognitive effort (Ascher, 2002), which can influence our speech in many ways.

In their study, Buchanan et al. (2014) tested the influence of psychological stress on the speech production by using the most effective laboratory tasks to study the participants' reaction on psychological stress: the TSST (Trier Social Stress Test) to induce stressful situation and the placebo TSST used for creating less stressful situations.³ Their findings show that the speech fluency of the participants was significantly reduced when put in stressful situations. However, the most significant increase was in the production of silent pauses, while the amount of filled pauses was much higher in the non-stressful speech. We can assume then that this particular type of dysfluency is either not related to psychological stress or regarded by speakers as unsuitable. Nevertheless, Buchanan et al. (2014) still proved that psychological stress can affect speakers both verbally and non-verbally (e.g. higher heart rate). The effect of stress has been observed also by psychotherapists who found out that a sudden onset of different types of dysfluencies indicated activation of an anxiety-provoking area of the patient's brain⁴ (Maclay & Osgood, 1959).

Social factors

Another important group of factors that are known to have a potential influence on the shape of language in general is social factors. These factors include regional background, socio-economic background, educational background, age, and gender. However, since the influence of the social factors, specifically age and gender, is the main topic of this thesis, we devoted an entire subchapter for the discussion of social factors and their influence on language variation (see subchapter 2.3).

³ For in detail description of the TSST task and the placebo TSST task, see Buchanan et al. (2014).

⁴ By anxiety we do not refer to the clinical disorder, but rather nervousness or uneasiness caused by a temporal situation.

Other factors

Besides psychological factors and social factors, we will also include a number of other factors that are believed to have a potentiality to influence dysfluent behaviour of speakers, but since they come from various domains, we will unify them under a general term *other factors*.

Bortfeld et al. (2001) introduce in their study a group of five different factors and examined their effects on the production of dysfluencies: age, gender, familiarity between the speakers, topic of the conversation, and conversational roles. As it was previously mentioned, age and gender are social factors and will be dealt with in detail in subchapter 2.3.

Concerning familiarity between speakers, Bortfeld et al. (2001) suggest that we might produce more dysfluencies when talking to a stranger, since we are more prone to become anxious. However, they also admit that there is a possibility for us to be more dysfluent while talking to an intimate. This is based on their general assumption that talking to an intimate makes us feel more confident and thus we can rely on their help if we experience some planning difficulty. To test this, they used a corpus which provides speech data of 48 speakers: 24 pairs of male and female strangers and 24 married couples. Their speech was recorded while participating in a referential communication study. They were asked to describe two sets of pictures. First set contained 12 pictures of children and the second set contained 12 pictures of abstract geometric figures. The results show that speakers were more dysfluent when talking to a familiar addressee. The same result was found in the study by Branigan et al. (1999). He used a corpus consisting of dialogues between 16 pairs of friends and 16 pairs of unfamiliar speakers. However, the difference between dysfluency rates was in both studies regarded as not significant, i.e. the probability value was higher than 0.05 ($p > 0.05$).

Another two factors considered in the study by Bortfeld et al. (2001) are topic of the conversation and conversational roles. To test the influence of the topic, they used the already-mentioned *referential communication task*, i.e. a picture description task where the participants were asked to describe pictures of children and pictures of abstract geometric figures. The results show that the participants were producing more dysfluencies when describing pictures of children. The explanation of this can be found in the study by Schachter et al. (1994) who studied the occurrence of dysfluencies in social science lectures, natural science lectures and humanities lectures. They found out that it is the humanities lectures that contained the highest rate of dysfluencies. These results prove their assumption that a topic that is characterized by a richness of vocabulary and thus a greater variety of options at a

choice point (same case with the pictures of children), will naturally contain more dysfluencies.

The influence of conversational roles on the production dysfluencies was tested by using the same *referential communication task* while one of the speakers was a so-called *director* and the other one was a *matcher*. The director was supposed to navigate the matcher to line up a set of pictures in a particular order. They found out that directors were producing more dysfluencies than matchers, which is probably due to their production of longer utterances. It has been suggested that they are generally associated with more planning difficulties and thus with higher dysfluency rates (Oviatt, 1995; Shriberg, 1996; cited in Bortfeld et al., 2001).

All these factors were also examined by Branigan et al. (1999) who call them *non-linguistic factors*. However, they also focused on another potential factor affecting fluency of speech and that is eye-contact. They examined whether the ability to see the conversational partner would influence production of dysfluencies. The results show that dysfluency rate is higher for the no eye-contact situations which might suggest that seeing our conversational partner can result in a more effective turn-taking and no need for repetitions to get the opportunity to talk. The effective turn-taking and the strategies to achieve it will be discussed in the next section dealing with the functions of dysfluencies.

2.2.2 Functions of dysfluencies

Dysfluencies as cues for listeners

In section 2.2.1 focusing on the description of different types of dysfluencies, we already discussed the most important positions in which these dysfluencies tend to occur. However, these positions can be also seen from a different perspective and that is as cues for listeners about the nature of the upcoming utterance.

One of the main goals of speakers is to be understood by their listeners. On the other hand, the goal of listeners is to extract meaning from the linguistic input presented by the speakers and realize what or who the speaker refers to (Arnold et al., 2003). This is referred to as *reference-resolution* and to achieve it, listeners need to use the combination of the lexical meaning of the referent and several discourse constraints which can make some referents more accessible than others (Arnold et al., 2003); for example, given information is much more accessible than new information. Arnold et al. (2003) suggest that this can be explained by so-called *expectancy hypothesis*: even though changing of topic is quite common, we

usually expect speakers to keep talking about the same thing and previously given information thus have a much higher expectancy than new information. Moreover, reference to new information causes more production difficulties and thus tends to appear together with some dysfluency. This suggests that dysfluencies can function as cues for listeners – they can expect from speakers to refer to a discourse-new entity which was not previously mentioned (Arnold et al., 2003).

Moreover, study by Watanabe et al. (2008) also proved that dysfluencies (specifically filled pauses) can function as indicators of long or complex utterances. In their experiment, the participants were presented with a pair of shapes appearing on a computer screen; one was of a simple shape and the other one of a complex shape. One second after the visual stimulus, they listened to a speech referring to one of these shapes. As soon as they realized which shape was referred to in the utterance they were supposed to press a button. The target phrases describing the shape were preceded by a filled pause, silent pause, or no pause. The results show that their response times were much shorter for the utterances preceded by filled pauses. We can assume then that certain types of dysfluencies can function as general indicators of difficult to process information which does not include only long and complex utterances but also low-frequency words or less predictable words, another typical position of filled and silent pauses.

Dysfluencies as communicative acts

As it was previously mentioned, one of our main goals when addressing our utterance to someone else is to be clear and well understood. To achieve this, we need to know how to effectively coordinate speech actions, i.e. *turn-taking* with our addressees. According to Clark (2002), there are four different strategies that are used in these situations.

The first strategy is to signal our desire to initiate speaking and this is often done by using so-called *orienting expressions*, such as conjunctions ‘and’, ‘but’, ‘so’ or discourse markers ‘well’, ‘now’, ‘so’, ‘anyway’. However, there is another possibility to signal the onset of the utterance and that is the usage of pre-utterance fillers such as ‘uh’ and ‘um’ or the simple repetition of the first word of the utterance.

The second strategy is to pursue an ideal delivery of the utterance. In order to meet listeners’ expectations, we often prefer to restart the utterance than use a repair, something that Clark (2002; pg. 8) calls *continuity principle*: "a preference for producing constituents fluently."

However, while speaking, we often encounter some planning problems and the ideal delivery is thus very rarely achieved. Therefore, the next strategy is to signal the intention to suspend speaking in order to warn the addressee. One of the most common signals is the lengthening of reduced vowels in function words, e.g. *the* pronounced as [ðɪ:], or the lengthening of any syllable that precedes the suspension point of the utterance. The vowel and consonant lengthening was already discussed in subsection 2.1.2.3.

The last strategy is to signal the intention and the duration of the delay after the suspension of the utterance. As already discussed, the best signals for delay is a usage of filler ‘uh’ for a short delay and ‘um’ when we expect a longer delay. Another possibility is to use so-called *mid-word cut-offs* which in our terminology would be just a repetition of a part of the word; for example *th- there is a potential problem* (Clark, 2002).

Individual patterns of dysfluent behaviour

The occurrence of dysfluencies is a result of speakers’ planning difficulties during language production and their dysfluent behaviour thus cannot be consciously controlled (Braun & Rosin, 2015). Since the processes of language production are considered to be speaker-specific, there are many studies suggesting and proving that also the production of dysfluencies is 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). This is of a great importance mostly for forensic phoneticians who use individual speech patterns to detect a criminal, but also for neurolinguists who can use dysfluencies as parameters to distinguish between individual planning processes (Braun & Rosin, 2015). However, we need to bear in mind that the individual patterns of dysfluent behaviour are results of many other factors, not just individual planning processes. As suggested by McDougall et al. (2015), it is mostly social factors that can play a big role in the language variation, and apart from fluency, they can also affect speech rate.

2.2.3 Dysfluencies and their effects on listeners

In the last section of the discussion of dysfluencies production we will consider the issue from a different perspective and list some of the effects dysfluencies can have on listeners, specifically effects on listeners’ judgements, attention, and processing.

Dysfluencies affecting listeners' judgements

As it is suggested, listeners have a general tendency to monitor not only what speakers say but also how they say it (Brennan & Williams, 1995). The form of the delivery thus proves itself to be a very important aspect for listeners and influences the way they perceive speakers. As suggested by Fox Tree (2002), dysfluent speech can make speakers look as less honest or as less comfortable with the topic. In the study by Brennan & Williams (1995), they focused on listeners' interpretations of speakers' utterances and their effects on listeners' interpretation of speakers' metacognitive state, i.e. the awareness of their own knowledge. For the first experiment, they used a simple question-answer task where they tested participants' *feeling of knowing (FOK)*, which is basically "people's ability to assess and monitor their own knowledge" (Hart, 1965; cited in Brennan & Williams, 1995; pg. 384). The answers for this experiment were used for the next two experiments in which they studied listeners' *feeling of another's knowing (FOAK)*. The results show that listeners can be truly sensitive to the occurrence of dysfluencies. The FOAK answers were much lower, proving that listeners' judgements of speakers' knowledge are based on their display of confidence about the topic or their commitment to the topic.

Dysfluencies affecting listeners' attention

As previously mentioned, certain dysfluencies can function as cues or indicators of speakers' upcoming delay (Clark, 2002). However, they also have a short-term consequence on listeners and that is increasing their attention to an upcoming utterance (Corley et al., 2007). This was evidenced in the study conducted by Fox Tree (2001), in which he showed that listeners were much faster in recognising the word when it was preceded by a dysfluency. However, it needs to be pointed out that this is based on the study of filled pauses, specifically 'uh'. As it was suggested, this particular filler signals short delay and thus has a potentiality to increase listeners' attention. 'Um' signals a long delay and might not have any effect on listeners' attention. This might be due to listeners' incapability to maintain heightened attention for so long (Fox Tree, 2001). Moreover, it is suggested that there might be also a shift from listeners' heightened attention and anticipation of the upcoming speech to listeners' help to complete speakers' ideas (Fox Tree, 2001); for example:

Ken: I like driving. I really do. I enjoy it very much.

Louise: I used to like it until I became the complete sl-um,

Ken: 'Slave'? Yeah. (Jefferson, 1974; cited in Fox Tree, 2001)

Dysfluencies affecting processing and comprehension

Since dysfluencies are considered to be phenomena which interrupt the flow of speech, there is a general assumption that dysfluencies can inhibit and slow down on-line processing and comprehension (Fox Tree, 1995). However, it has been proven that listeners' heightened attention to an upcoming utterance can be beneficial for the overall processing and comprehension of the utterance, specifically when it comes to the occurrence of fillers. We can also see the beneficial nature of the fillers when they represent an editing phase of a repair dysfluency. According to Levelt (1989), fillers, as well as silent pauses, can warn listeners that the articulated message was wrong and it will be substituted. This facilitates their processing of the utterance. As the study by Fox Tree (1995) proves, also repetitions can help listeners to recognize the word faster – target words were identified about 84 ms faster when appearing in the presence of repetitions. However, false starts seem to have a completely different effect. When a false start was absent, the target word was identified about 22 ms faster, suggesting that false start is a type of dysfluency that can truly hinder comprehension and is not in any way beneficial.

2.3 Dysfluencies in a sociophonetic context

2.3.1 What is sociophonetics?

Before we get to the discussion about sociophonetics, it seems important to discuss the notion of *sociolinguistics* from which the study of sociophonetics is derived. The main focus of *sociolinguistics* as a discipline is the influence of different social factors on the language variation, change and use. It developed several subfields and one of the most influential subfields is the *variationist tradition* which was established by an American linguist called William Labov (Baranowski, 2013) and his research conducted on the island called Martha's Vineyard in Massachusetts in 1961. What defines variationists is their approach to *variability*. They believe that "a language system that did not display variability would not only be imaginary but [also] dysfunctional, since structured variability is the essential property of language that fulfils important social functions and permits orderly linguistic change" (Milroy and Gordon, 2003; p. 4). But what does *variability* actually refer to?

As mentioned above, in our specific context, the term *variability* basically refers to a principal characteristic of any language: a disposition for variation and change which exists at any level of linguistic representation. However, the study of a socially conditioned variation is

much more focused on phonetics than on any other language domains (Hay & Drager, 2007), which brings us to the discussion of *sociophonetics*.

Sociophonetics is basically a term that refers to an interface of sociolinguistics and phonetics (Baranowski, 2013) and focuses on phonetic variation, i.e. the pronunciation differences among individuals, and social factors conditioning the variation (Hay & Drager, 2007), which will be discussed in section 2.3.2. A specific feature of pronunciation that differs (varies) among speakers is so called *variable* and its actual realizations in speech are so called *variants* (Meyerhoff, 2006).

As Meyerhoff (2006) points out, there are also variations in the speech within an individual speaker, called *intraspeaker variation*, as opposed to variation between individual speakers, i.e. *interspeaker variation*. This proves that speakers can alternate between the ways they speak when speaking in different situations or to different interlocutors. These variations can be also conditioned by a personal mood or intentions of a single speaker.

In the next section, we will offer a discussion of the main social factors influencing phonetic variation and language variation in general.

2.3.2 Social factors conditioning language variation

Regional background

Speakers' origin is one of the main social factors affecting the way people speak. If a person from England meets a person from North America, even though they would be still speaking English, their 'Englishes' would be notably different. This is because of the usage of different *dialects*, i.e. varieties of a specific language which are characterized by sets of distinctive features at the level of pronunciation, vocabulary, and sentence structure (Meyerhoff, 2006). For example, regarding the pronunciation, one of the most noticeable differences between these two dialects is *rhoticity*. North American English is known to be a rhotic dialect, which means that pronunciation of [r] occurring before consonants and at the end of words is retained, for example in words like 'farm' and 'far'. Standard British English, on the other hand, is not rhotic, and loses [r] in all these environments (Brinton & Arnovick, 2000).

As Brinton & Arnovick (2000) point out, we can distinguish between two major groups of dialects. The former example is the case of so-called *national dialects* or *national varieties*. The existence of different national dialects of English is a result of a gradual spread of English through British Empire during a massive colonization happening between the

seventeenth and nineteenth century. The national varieties that we recognize today include English of North America, Canada, Australia, New Zealand, South Africa, and more.

The second group of dialects is called *regional dialects* and they exist within each national dialect. For example, in Britain we can distinguish between dialects such as Cockney, Birmingham, Devonshire, Cornwall, or Yorkshire, to name just a few. Moreover, in Britain there are also dialects which are commonly called *dialects with semi-national status*, and those are Scottish, Welsh, and Irish (Brinton & Arnovick, 2000).

A very important feature that is attributed to a dialect of an individual is a so-called *accent*. It is a set of phonological characteristics that are typical of a specific variety. Phonological differences, together with morphological differences, are the most significant and noticeable features among the dialects as well as the most reliable cues to speakers' origin.

In regard to the influence of an individual's origin and production of dysfluencies, the results of the study conducted by McDougall et al. (2015) suggest that the range of occurrence of dysfluencies is not dependent on speakers' accent. They tested 20 male speakers of Standard Southern British English (SSBE) and 20 male speakers of York English. However, when looking at the results in detail, the SSBE speakers produced more filled pauses and repetitions than York speakers, but produced less prolongations (vowel and consonant lengthening) and interruptions (false starts). As we know, speakers' variation can be also dependent on other factors, such as socio-economic background or educational background. In their study, SSBE speakers were students of University of Cambridge, while York speakers were recruited from outside of a job centre which suggests that the results might be due to the combination of different social factors, not only accent.

Socio-economic background

The combination of different social factors also includes speakers' socio-economic background, i.e. affiliation into a specific *social class*. The notion of social class has a long history with the latest theories being associated with Karl Marx and Max Weber. For Marx, the main distinction was between those who produced capital, i.e. working class, and those who controlled the capital produced by others, i.e. middle class. However, according to Weber's theory, the society can be divided into many more social classes, not just two. A person's status is primarily defined by their economic situation, which is, however, often influenced by their life style and life choices (Meyerhoff, 2006). In general, social class can be then defined as a group of people sharing the same status within the society, which is based

mostly on their occupation, income and wealth, but also their aspirations and mobility (Meyerhoff, 2006).

Socio-economic background affects speech of an individual to such an extent that speech can be the most important factor in revealing speakers' affiliation to a specific social class. There are several linguistic markers which are typical of speech of people of lower-status classes and those of higher-status classes. Kerswill (2007) suggests that sociolinguists have been pointing out social differences in the way talk is organized already in the late 1950's. One of the most prominent figures, Basil Bernstein, points out the differences of the talk organisation between working-class children and middle-class children. In his study conducted in 1971, he found out that the working-class children spoke in a *restricted code*, characterised by "unfinished and short sentences, simple clauses, limited usage of adjectives and adverbs, and also different types of dysfluencies such as repetitions and some hesitation phenomena" (Bernstein, 1971; cited in Stockwell, 2002, pg. 56). Middle-class children, on the other hand, spoke in an *elaborated code*, a discourse featuring "accurate grammatical order, complex sentences, impersonal pronouns, passive constructions and unusual adjectives and adverbs" (Bernstein, 1971; cited in Stockwell, 2002, pg. 56). The usage of a specific code also suggests a close connection with an educational failure, an issue that will be discussed later in this section (Bernstein, 1958, 1971; cited in Kerswill, 2007).

However, the most important works showing the influence of socio-economic background on the speech are those of William Labov. We already mentioned his dialectal research conducted on the island Martha's Vineyard, whose methods and principles established the field of sociolinguistics (Meyerhoff, 2006). However, his most influential work discussed in the relation to the study of social classes and their influence on the speech, is the study of the social stratification of /r/ in department stores in New York City, conducted in 1966. The most typical feature that sets New York City variety apart from the Standard American variety is its r-less pronunciation, which means that unless it occurs before a vowel, /r/ is not pronounced (Meyerhoff, 2006). However, the r-less pronunciation is a typical social marker of people from lower-status classes and this variable is thus, as Labov (2006) points out, the best social differentiator of New York City speech. He selected three different department stores with the highest ranking, middle ranking and lowest ranking on the price and fashion scale, which predicts the socio-economic status and stratification of their customers, as well as the sales people. His hypothesis that "sales people in the highest ranked store will have the highest values of (r), those in the middle ranked store will have intermediate values of (r), and those in the lowest ranked store will show the lowest values"

was confirmed and his results thus clearly show the effects of social class on the people's speech.⁵

However, as Meyerhoff (2006) points out, some speakers can experience a feeling that the variety they use is in some way inferior or even ugly. This is usually referred to as *linguistic insecurity* and the r-less pronunciation of some New Yorkers is one of the best examples. Allegedly, some of them have such negative feelings about this stereotyped feature that they disapprove of the r-less pronunciation even when it comes to their families or friends (Meyerhoff, 2006). Linguistic insecurity is manifested in speakers' attempt to reach a more prestigious pronunciation which often results in so-called *hypercorrection*, i.e. the production of a specific variant that does not actually occur in their social class.

In the UK, measure of social class developed in the 1970's and places people in one of seven different social classes, based on their occupation: elite, established middle class, technical middle class, new affluent workers, traditional working class, emergent service workers, and precariat (see Figure 4). It provides and compares the results of "the largest [web] survey of social class ever conducted in the UK, the BBC's Great British Class Survey (GBCS)" (Savage et al., 2013; p. 220), as well as "nationally representative survey made by a survey firm GfK" (Savage et al., 2013; p. 231).

	% GfK	% GBCS	Description
Elite	6	22	Very high economic capital (especially savings), high social capital, very high highbrow cultural capital
Established middle class	25	43	High economic capital, high status of mean contacts, high highbrow and emerging cultural capital
Technical middle class	6	10	High economic capital, very high mean social contacts, but relatively few contacts reported, moderate cultural capital
New affluent workers	15	6	Moderately good economic capital, moderately poor mean score of social contacts, though high range, moderate highbrow but good emerging cultural capital
Traditional working class	14	2	Moderately poor economic capital, though with reasonable house price, few social contacts, low highbrow and emerging cultural capital
Emergent service workers	19	17	Moderately poor economic capital, though with reasonable household income, moderate social contacts, high emerging (but low highbrow) cultural capital
Precariat	15	<1	Poor economic capital, and the lowest scores on every other criterion

Figure 4. Summary of seven social classes in the United Kingdom (Savage et al., 2013; p. 230)

⁵ For the entire method and results of the study, see Labov (2006).

As previously mentioned, in our study, we focus our attention only on two social factors: age and gender. Regional background, socio-economic background, educational background or the effects of other factors such as religion and ethnicity are not considered. However, almost all the participants come from South East region or London region, with just few individuals who are originally from South West, North West, Midlands or East Anglia, but have lived in the South East region or the Greater London for several years. The majority of the participants are either current or former students, researchers, or workers at the University of Reading, Berkshire, and on the basis of their stated educational background (the majority of the participants have a university education), we also assume an affiliation to similar socio-economic classes.

Educational background

Performance in linguistic skills, as Stockwell (2002) points out, is considered to be an important indicator of the level of education as well as intelligence. As mentioned above, the study conducted by Basil Bernstein proves that the level and the type of education as a social factor influencing speech is closely related to socio-economic background of speakers by being actually treated as one of the indicators of socio-economic background. The family conditions in which they are raised and the amount of socialization they experience affect speakers' communication skills (Bernstein, 1971; cited in Young, 2002). The theory of restricted and elaborated code usage was tested by Wodak (1996; cited in Kerswill, 2007) by using the technique of oral retelling of stories. The results show that the middle-class people were focused on providing very accurate and backgrounded stances, while working-class people would use their own point of view in retelling the story, using comments like 'You can't do anything about it, anyway.' Wodak suggests that this is due to the years of socialization of the middle-class people, received mostly through education, which leads to "oversophisticated and fact-oriented summaries" (Wodak, 1996; cited in Kerswill, 2007; pg. 58).

From a general point of view, it can be assumed that if speakers come from well-educated families of a good social status and good income, the possibility to obtain a better education is undoubtedly higher, especially when it comes to the higher education. In all the parts of the UK, except Scotland, students are required to pay a tuition fee for both levels of higher education, i.e. undergraduate and postgraduate level, and good socio-economic background is thus beyond any doubt a very important factor in obtaining decent education. However, we have to bear in mind that personal endowment and motivation are also very

crucial in the process of obtaining education, but as Lawton (1968) points out, without an access to the particular educational institutions, it is difficult to proceed in the intellectual growth.

Other factors

For the sake of completeness, we briefly mention here three other factors that many studies about social factors and their relation to language include in their discussions: ethnicity, race and religious affiliation of speakers.

Ethnic groups, as Milroy and Gordon (2003) point out, are best described as minority groups which are formed on the basis of some shared cultural characteristics such as sense of place, common history, social ideology, often religion and even communicative conventions. Some of the examples of ethnic groups that Milroy and Gordon mention are African American and Latino communicates in the United States.

Race and religion are both very closely associated with ethnicity. Race is seen as "physical variations singled out by members of a community and treated as ethnically significant" (Giddens, 1989; cited in Milroy and Gordon, 2003; p. 109). Religion, as mentioned above, is often considered to be one of the characteristics of a specific ethnic group, or "a culturally accepted indicator of ethnicity" (Milroy and Gordon, 2003; p. 114).

The discussion of these three factors as sociolinguistic factors will be left out and we move directly to the next section dealing with the influence of two factors considered in the empirical part of the present study: gender and age.

2.3.3 Influence of gender and age as social factors

Gender

Gender and its relationship to language has been of a great interest to researchers from various fields, including sociolinguistics and many studies prove that there is a great amount of noticeable differences between the speech of men and women. Smith (2002) points out that they can be observed in the pronunciation, where women prove to produce more standard and correct pronunciation than men, but also in grammatical forms, choice of vocabulary, choice of speech styles, and in the prosody, such as intonation, pitch, or rhythm. However, in some of the studies researchers refer to the influence of sex, while other studies use the term gender, suggesting the synonymity between the two, even though there is a relevant difference. Sex refers to "biologically and physiologically based distinction between males and females"

(Meyerhoff, 2006; p. 201). Gender, on the other hand, is socially constructed and learnt. During our life, we acquire certain characteristics that the society perceives either as masculine or feminine (Talbot, 2010). However, when conducting a research, Milroy and Gordon (2003) suggest that "at the data collection stage, it makes sense to talk of sampling speakers according to sex, but when interpreting the social meaning of sex-related variation, we should think of gender as the relevant social category."

It has been argued that it is becoming more and more common to talk about language and gender, rather than language and sex (Meyerhoff, 2006), and therefore, in the present study, we will be using only the term gender. However, it should be pointed out that none of the male participants were perceived as more feminine and none of the female participants were perceived as more masculine. Furthermore, none of the participants claimed to be of a different gender than they would be expected to be based on their biological sex.

When considering gender in relation to the production of dysfluencies, there are several studies that identified gender as a potential factor influencing dysfluent behaviour. Branigan (1999), whose study and method was already mentioned earlier in section 2.2.1 dealing with the factors influencing dysfluent behaviour of speaker, found out that men were more dysfluent than women, with the average dysfluency rate 4.35 per 100 words as opposed to 3.76 per 100 words for women.

Also Bortfeld et al. (2001), whose study and method were mentioned in the same section, considered gender as a potential influence. According to their results, men had higher overall dysfluency rate than women, with the rate of 6.80 to 5.12 per 100 words. The results show this difference is mostly due to the production of fillers and repetitions. A higher dysfluency rate for fillers for men was also found in the study conducted by Shriberg (1996) who used the Switchboard corpus of informal telephone conversations on several prescribed topics. However, there is no specific reason that would make us assume that men are more dysfluent than women. Shriberg suggests this difference might be due to their desire to hold the floor of the conversation, i.e. to speak to a person or a group of people for a long time without allowing them to take turn. However, she also suggests that the effects of this specific variable should be considered in relation to other social or cognitive factors and therefore we now turn to the other social factor, age.

Age

Similarly as gender, age is a social factor which is found to be well-reflected in the actual speech. As Helfrich (2002) suggests, there are several speech cues, including phonological,

syntactic, semantic, extralinguistic, and paralinguistic, that have a potentiality to differentiate between people of different age groups. The older age groups cues emerge during the process of ageing and ageing-related changes that, apart from the way of talking, influence also "cognitive, motor, and perceptual functioning" (Bortfeld et al., 2001; pg. 128). Some of the studies point out that ageing can actually improve people's speech. Over time, according to some of the studies, people develop:

- greater ability to define words (Obler & Albert, 1984; Sandson et al., 1987, cited in Bortfeld et al., 2001),
- richer vocabulary and increased conceptual development (Harwood, 2006; cited in Meyerhoff, 2006), and
- the usage of more elaborate syntactic forms (Obler & Albert, 1984; cited in Bortfeld et al., 2001).

However, on the other hand, ageing can make older people to experience more difficulties to retrieve the words (e.g. Rastle & Burke, 1996; cited in Bortfeld et al., 2001), which can cause higher dysfluency rates in their speech. Indeed, the study conducted by Bortfeld et al. (2001) shows that older speakers (ranged from 63 to 72 years old) had an average dysfluency rate of 6.65 per 100 words (for fillers, repetitions, and repairs) as compared to younger speakers whose dysfluency rate was 5.55 per 100 words. Bortfeld et al. also mention earlier studies focused on the influence of age on dysfluent behaviour of speakers (e.g. or Albert, 1980 or Schow et al., 1978) which also found higher dysfluency rates for older participants.

2.4 Hypotheses

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: Male speakers are more dysfluent than female speakers.

H2: Older speakers are more dysfluent than younger speakers.

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

3. Method and Material

3.1 The process of recording and the participants

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-07mkII. In both cases we used the sampling rate of 48-kHz. 32 native English speakers from different areas of England, however, 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.

3.2 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. We used the following coding:

	Code	Explanation
Filled pause – vowel + nasal	FIL-VN	any pause filled by insertion of vowel and nasal, i.e. um
Filled pause – vowel + vowel	FIL-VV	any pause filled by insertion of vowel, i.e. uh
Silent non-grammatical pause	PAU	any silent pauses occurring in non-grammatical places
False start	INT	any non-retracted false start, i.e. incomplete utterances
Repair	COR	any type of an overt repair, i.e. an error already articulated
Repetition	REP-P, REP-W1, REP-W2, REP-W3	repetition of a part of a word, one word, two words, or three words and more
Vowel lengthening	EXT-V	any vowel whose duration is ≥ 200 ms
Consonant lengthening	EXT-C	Any consonant of duration ≥ 200 ms

Table 1. Coding of the examined types of dysfluencies

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:

$$\text{total amount of dysfluencies} / \text{total amount of words} \times 100$$

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

3.3 Statistical analyses

The present thesis 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.

4. 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.

4.1 Overall results

All dysfluency rates we consider in the present study are per 100 words, including filled pauses, repeated words, and fragmented words.

Considering the influence of age, 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: $F(1,30) = 1.37$; $p > 0.2$.

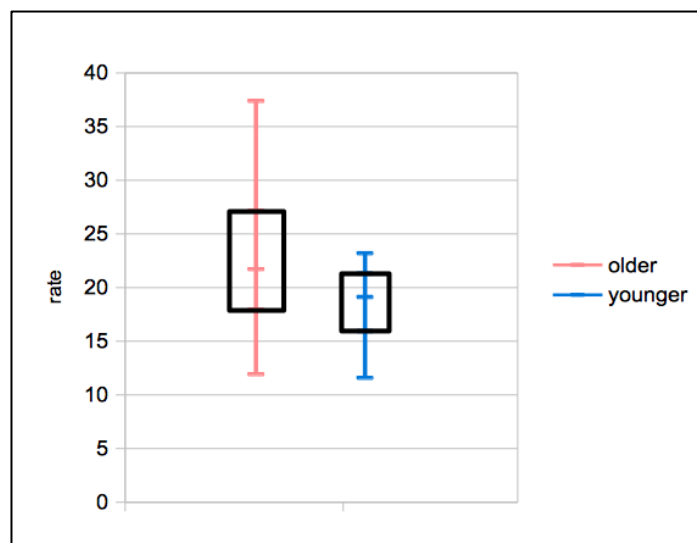


Figure 5. Boxplot reporting dysfluency rates by age group

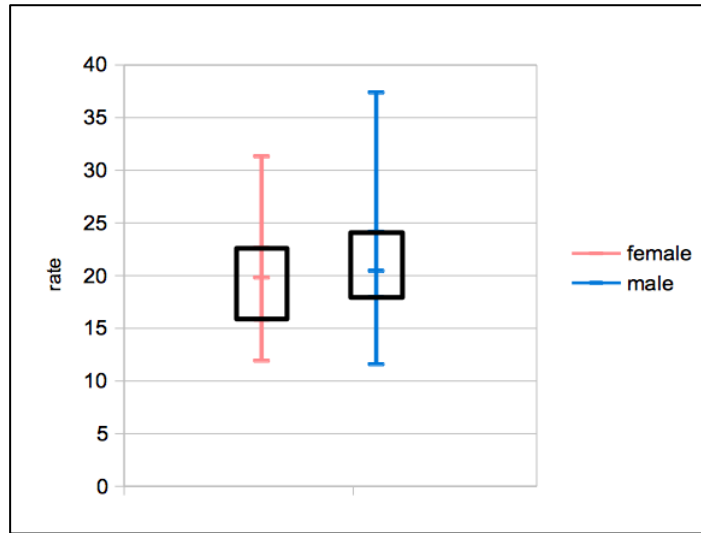


Figure 6. Boxplot reporting dysfluency rates by gender group

For the purpose of demonstration, we include a table that shows dysfluency rates per 100 words of two age groups and two gender groups.

	Average dysfluency rates per 100 words
Older speakers	22.59
Younger speakers	18.60
Male speakers	21.70
Female speakers	19.49

Table 2. The average dysfluency rates per 100 words of all four groups of speakers

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: $t(14) = - 1.22; p = 0.24$.

The table below shows the groups as they were compared in the post hoc t-tests and their respective average dysfluency rates per 100 words. These values are also represented in Figure 5.

	Average dysfluency rates per 100 words
Older male speakers	24.63
Older female speakers	20.56
Younger male speakers	18.78
Younger female speakers	18.43

Table 3. The average dysfluency rates per 100 words for all four groups of speakers

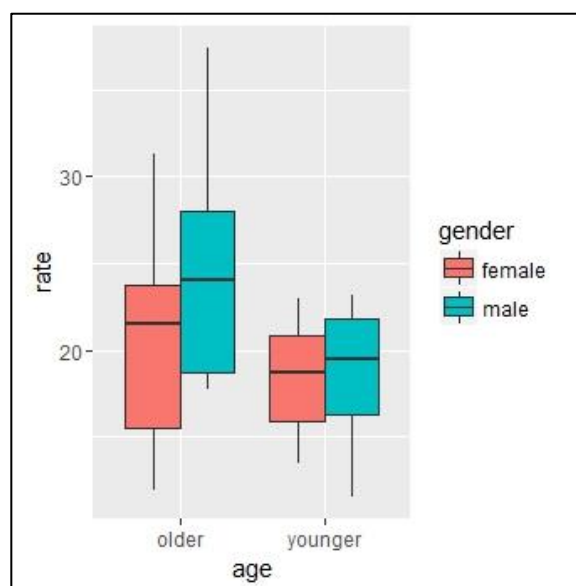


Figure 5. Boxplot reporting dysfluency rates for all four groups of speakers

4.2 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, we 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 we 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: $F(1,28) = 4.77$; $p < 0.05$.

We 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: $F(1,28) = 3.75$; $p < 0.1$.

In the table below, we offer average dysfluency rates of all four groups of speakers for all significant types of dysfluencies.

	Older speakers	Younger speakers	Male speakers	Female speakers
False starts	0.36	0.60	0.50	0.46
Vowel lengthenings	3.76	3.30	3.01	4.04
Consonant lengthenings	4.26	2.60	3.81	3.05
Repetitions	2.19	1.32	2.12	1.39

Table 4. The average dysfluency rates of all four groups of speakers for all significant types of dysfluencies

4.3 Individual preferences

As it was previously stated in section 2.2.2, 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, we tried to even out all the factors (regional, socio-economic, and educational) as much as possible so our results would not be compromised in any way (see section 2.3.2).

In this subchapter, we 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 we will consider the female speakers and in the second part we will look at the preferences of the male speakers. Older speakers will be marked with (o), younger speakers with (y).

Regarding female speakers, from Figure 6 we can see that 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’. Only exceptions were speakers F05 (y), F06 (y), and F16 (y), who produced a slightly higher amount of ‘uh’.

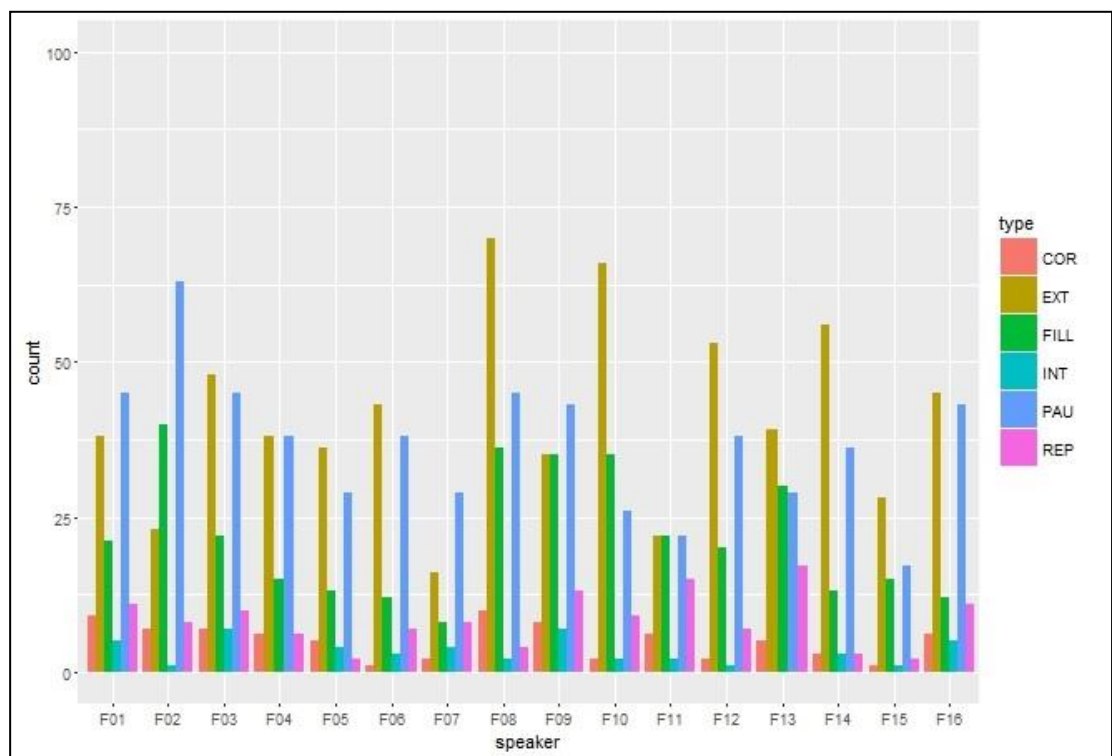


Figure 7. Total amount of dysfluencies per five-minute sections produced by each female speaker

From the figure below we can see that 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.

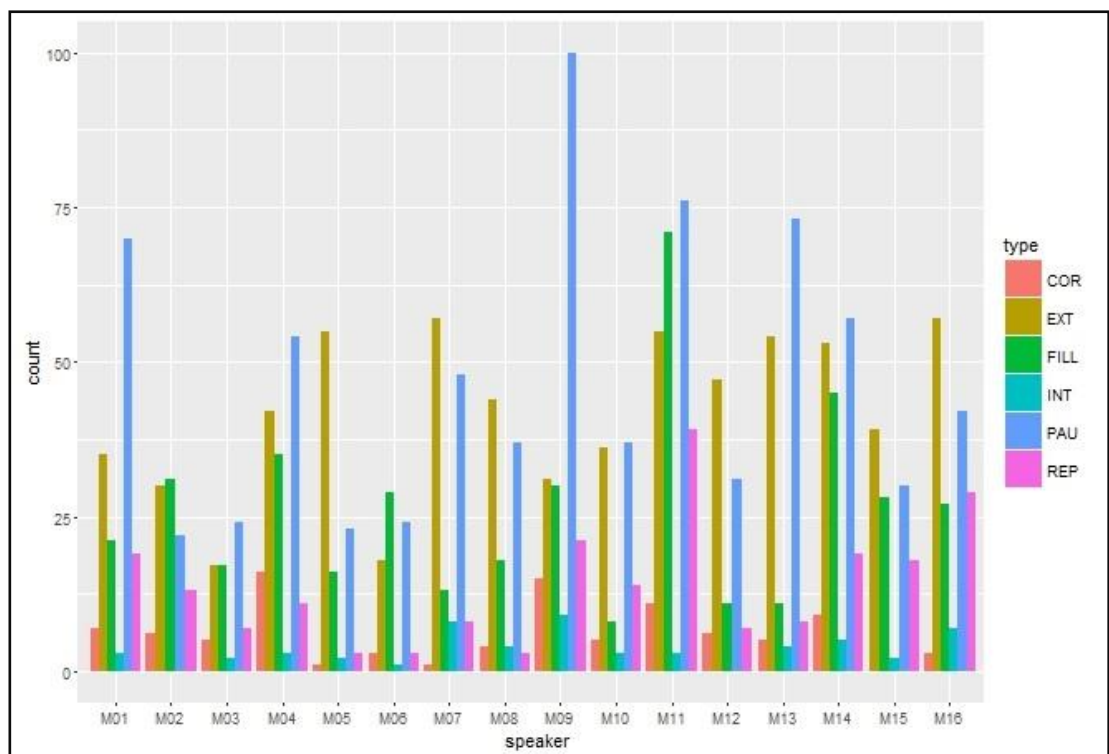


Figure 8. Total amounts of dysfluencies per five-minute sections produced by each male speaker

Below, for the purpose of demonstration, we offer two more figures: Figure 9 shows the total amount of vowel and consonant lengthenings produced by every single speaker. Figure 10 shows the total amount of two types of filled pauses, ‘uh’ and ‘um’.

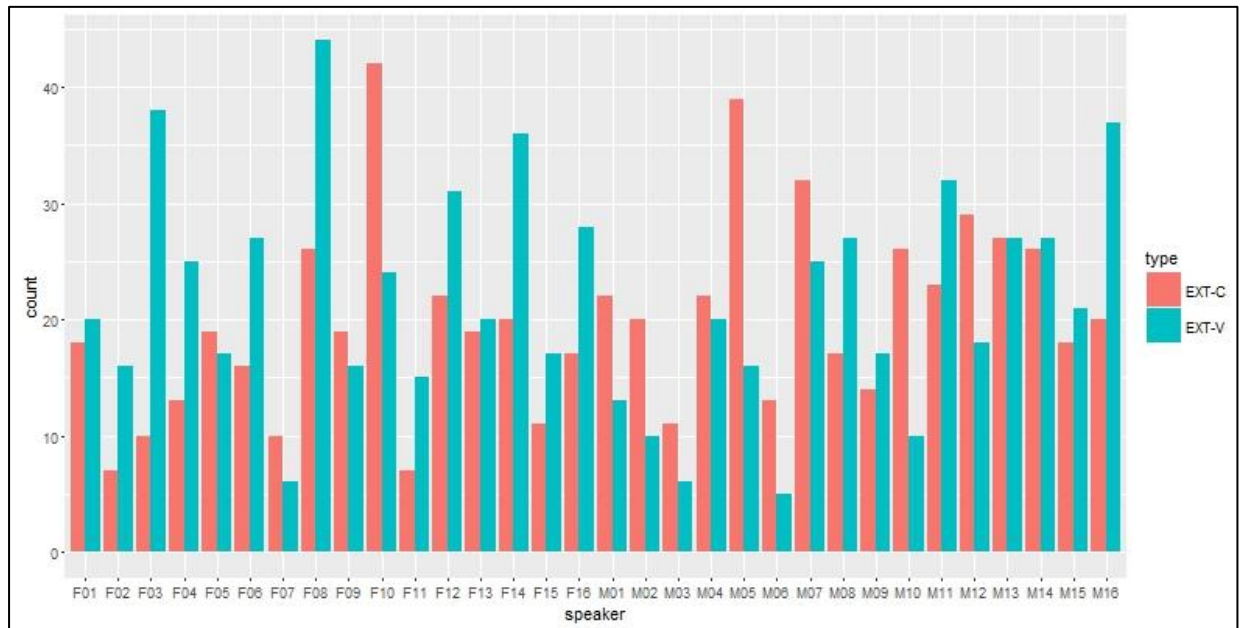


Figure 8. Total amounts of vowel and consonant lengthenings produced by each speaker in five-minute sections

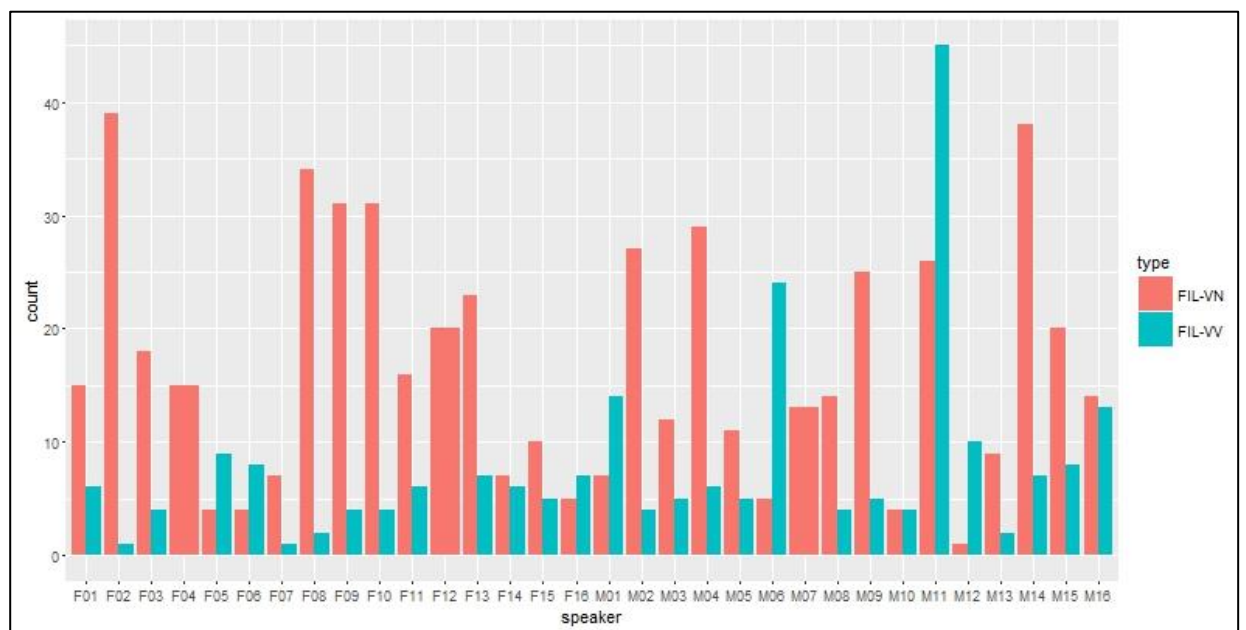


Figure 9. Total amounts of both types of filled pauses produced by each speakers in five-minutes sections

5. Discussion

In the first part of the research, we looked at the influence of age, gender and the interaction between these two factors. We compared younger and older speakers, including both male speakers and female speakers, and then we 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 our 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, our last hypothesis was confirmed. We 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 a bit difficult to determine. As we mentioned in the theoretical part, there is no particular reason to assume why male speakers should be more dysfluent than female speakers. We offered a suggestion made by Shriberg (1996), who says that it 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 for everyday conversations, where both sides get a chance to ask

and also 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. They 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 we 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 actually driven by our male speakers. The same results was 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 (subsection 2.1.2.2), 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. Also Bortfeld et al. (2001) studied the influence of age and their results, such as our results, proved that older speakers were more dysfluent than younger speakers. However, we need to point out that the age groups they worked with differ from ours. For the group of older speakers, we included also speakers who in general might be classified as 'middle-aged.' The reason for their inclusion was the lack of participants we could use to actually create additional groups of middle-aged female speakers and middle-aged male speakers. In future research, the study

of the influence of age on the production of dysfluencies should undoubtedly include a group of middle-aged speakers in order to see to what extent the results would change.

However, when considering age effects, we also need to point out that we found an exception: false starts. The results show that it is actually younger speakers who produce more 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 actually proven to improve people's speech and since their vocabulary is actually 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 in general, the production of dysfluencies was found to be affected by age but not gender. However, there are specific types of dysfluencies for which this statement proves to be incorrect. Future research could take into consideration more factors and could also create additional age groups.

6. Conclusion

The aim of the present thesis 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.

In the theoretical part, the main issues regarding the nature of dysfluencies as well as their production were discussed. In the beginning, a general overview of language production and trouble detection was provided and then we focused on the main characteristics of the seven types of dysfluencies together with their typical positions within the utterance. Afterwards, we discussed the main factors influencing the production of dysfluencies, different functions of dysfluencies as well as effects their occurrence have on listeners. Lastly, we provided an overview of different social factors that are known to have an influence on language variation and thus might have an influence on speakers' dysfluent behaviour as well. Age and gender were the factors we examined in the present thesis and thus were discussed in a separate section.

In the first chapter of the empirical part, we presented the process of recording and recruitment of the participants. We recorded speech of 32 native English speakers from England and the recordings were obtained in two ways: in a sound-proof studio or by using a handheld recorder. 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 questions about their hobbies and interests.

The participants were selected according to their age and gender in order to form 4 groups of 8 participants: younger male speakers, younger female speakers, older male speakers, and older female speakers. All the acquired recordings were orthographically transcribed, analysed in software designed for phonetic analysis, Praat and then segmented into words and phonemes. Then we identified the presence of dysfluencies by using specific codes. Afterwards, we ran two different scripts: one to extract the total amount of dysfluencies for every single speaker and one to calculate the amount of words pronounced by every single speaker. By using a specific formula, we then calculated the amount of dysfluency rates per 100 words.

The empirical part of the present thesis was built on three different hypotheses. To test them and to see whether the differences between the dysfluency rates of our four social groups were significant, we ran several statistical tests. The first hypothesis was confirmed

with a statistically significant difference between older speakers and younger speakers, while the second was not confirmed. There was no statistical difference between male speakers and female speakers in general. The third hypothesis was confirmed only with repetitions.

Subsequently, in the discussion part of the present thesis, we offered some potential explanations for the reasons of the results and proposed several suggestions for the future research.

Bibliography

Arnold, J. E., Fagnano, M. & Tanenhaus, M. K. (2003). Disfluencies Signal Thee, Um, New Information. *Journal of Psycholinguistic Research*, 32 (1), pp. 25 – 36.

Ascher, M. L. (2002), Paradoxical Intention. In: Hersen, M. & Sledge, W. (Eds.), *The Encyclopaedia of Psychotherapy*, pp. 331 – 338. New York: Elsevier Science.

Baranowski, M. (2013). Sociophonetics. In: Bayley, R., Cameron, R. & Lucas, S. (Eds.), *The Oxford Handbook of Sociolinguistics*, pp. 403 – 424. Oxford: Oxford University Press.

Boersma, P., and 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), pp. 123 – 147.

Branigan, H., Lickley, R. & McKelvie, D. (1999). Non-linguistic influences on rates of disfluency in spontaneous speech. *Proceedings of the 14th International Congress of Phonetic Sciences. (ICPhS 1999)* (pp. 387 – 389). San Francisco: University of California.

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.

Brennan, S. E. & Williams, M. (1995). The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive State of Speakers. *Journal of Memory and Language*, 34, pp. 383 – 398.

Brinton, L. J. & Arnovick, L. K. (2000). *The English Language: A Linguistic History*. Oxford: Oxford University Press.

Buchanan, T., Laures-Gore, J. S. & Duff, M. C. (2014). Acute stress reduces speech fluency. *Biological Psychology*, 97, pp. 60 – 66.

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

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

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

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.

Collard, P. (2009). *Disfluency and listeners' attention: An investigation of the immediate and lasting effects of hesitations in speech*. (Doctoral Thesis). School of Psychology, Psychology and Language Sciences. University of Edinburgh.

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, pp. 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, pp. 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), pp. 320 – 326.

Fox Tree, J. E. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34(1), pp. 37 – 55.

- Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech*, 4, pp. 18 – 26.
- Hay, J. & Drager, K. (2007). Sociophonetics. *The Annual Review of Anthropology*, 36, pp. 89 – 103.
- Helfrich, H. (2002). Age Markers in Speech. In: Scherer, K. R. & Giles, H. (Eds.), *Social Markers in Speech*, pp. 63 – 108. Cambridge: Cambridge University Press.
- Kerswill, P. (2007). Social class. In : Llamas, C. Mullany, L. & Stockwell, P. (Eds.), *The Routledge Companion to Sociolinguistics*, pp. 51 – 61. New York: Routledge.
- King, J., Foulkes, P., French & P., Hughes, V. (2013). Hesitation markers as parameters for forensic speaker comparison. *International Association of Forensic Phonetics and Acoustics Annual Conference (IAFPA 2013)*. Tampa: University of South Florida.
- Labov, W. (2006) *The Social Stratification of English in New York*. Cambridge: Cambridge University Press.
- Lawton, D. (1968) *Social Class Language and Education*. London: Routledge and Kegan Paul Ltd.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, pp. 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, pp. 19 – 44.

- Machač, P. & Skarnitzl, R. (2009). *Principles of Phonetic Segmentation*. Praha: Nakladatelství Epocha.
- Meyerhoff, M. (2006). *Introducing Sociolinguistics*. Abington: Routledge.
- 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.
- Milroy, L. & Gordon, M. (2003). *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell Publishing Ltd.
- 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.
- Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., ... Miles, A. (2013). A New Model of Social Class? Findings from the BBC's Great British Class Survey Experiment. *Sociology*, 47 (2), pp. 219 – 250.
- Schachter, S., Rauscher, F., Christenfeld, N. & Crone, K. T. (1994). The Vocabularies of Academia. *Psychological Science*, 5 (1), pp. 37 – 41.
- 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), pp. 153 – 169.
- Smith, P. M. (2002). Sex Markers in Speech. In: Scherer, K. R. & Giles, H. (Eds.), *Social Markers in Speech*, pp. 109 – 146. Cambridge: Cambridge University Press.
- Stockwell, Peter (2002). *Sociolinguistics. A Resource Book for Students*. London: Routledge.
- Talbot, M. (2010). *Language and Gender*. Cambridge: Polity Press.

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.

Volín, J. (2007). *Statistické metody ve fonetickém výzkumu*. Praha: EPOCH.

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, pp. 81 – 94.

Wingate, M. E. (1984). Pause loci in stuttered and normal speech. *Journal of Fluency Disorders*, 9(3), pp. 227 – 235.

Young, R. (2002). Basil Bernstein's Sociolinguistic Theory of Language Codes. Retrieved from <http://zimmer.csufresno.edu/~johnca/spch100/3-3-bernstein.htm>

Zhrnutie

Táto bakalárska práca je zameraná štúdiom vplyvu dvoch sociálnych faktorov, veku a pohlavia, na častosť a obmenu rôznych typov dysfluencií v spontánnej reči rodených hovoriacich angličtiny z Anglicka. Práca je rozdelená na dve hlavné časti: teoretický základ, ktorý predstavuje zhrnutie doterajších poznatkov o danej problematike, a praktickú časť, ktorá sa venuje samotnému výskumu.

Prvá časť teoretického základu ponúka vysvetlenie pojmov ako spontánna reč, dysfluencie a taktiež ponúka popisuje vzťah medzi dysfluenciami a tvorbou jazyka. Dysfluencie boli totiž dlho považované za nežiaduce prvky spontánnej reči, ktoré nemajú pre komunikačné účely žiadny význam. Avšak posledné štúdiá dokazujú, že dysfluencie majú pre komunikáciu skutočne veľký prínos. Braun & Rosin (2015) poukazujú na to, že dysfluencie sú dôležitými indikátormi procesov plánovania reči a monitorovania vlastnej reči. V prvej časti preto ponúkame prehľad procesov, ktoré prebiehajú pri tvorbe jazyka: konceptualizácia, formulácia a artikulácia. Pri tvorbe reči sa však môže vyskytnúť niekoľko problémov. Levelt (1983, 1989) poukazuje na to, že k ich rozpoznaní vlastníme špeciálny kognitívny systém, ktorý nám pomáha vnímať našu vnútornú aj vonkajšiu reč, a tak rozpoznať chybu v akomkoľvek štádiu tvorby jazyka, čoho výsledkom sú rôzne typy dysfluencií. V tejto štúdií sa zaoberáme siedmimi typmi dysfluencií: vyplnené pauzy, tiché pauzy, opakovania, opravy, falošné začiatky, predĺženie samohlások a predĺženie spoluhlások.

V ďalšej časti sa dostávame k témam týkajúcich sa tvorby a významu dysfluencií. Je mnoho faktorov, ktoré ovplyvňujú dysfluentné chovanie u hovoriacich. Ľudská reč je často ovplyvňovaná stresujúcimi a znepokojujúcimi situáciami ako sú napríklad pohovor alebo verejný prejav. Bolo dokázané, že plynulosť ľudskej reči je pod vplyvom stresu znížená a produkcia dysfluencií sa zvyšuje (Buchanan a spol., 2001).

Existuje však i mnoho iných faktorov, ktoré ovplyvňujú dysfluentné chovanie u hovoriacich. Bortfeld a spol. (2001) poukazujú na to, že vzťah medzi hovoriacim a jeho adresátom taktiež môže hrať určitú rolu. Logicky by sme sa mohli domnievať, že ak sa rozprávame s niekým, koho nepoznáme, je veľmi pravdepodobné, že budeme nervóznejší a tým pádom budeme produkovať viac dysfluencií. Bortfeld a spol. (2001) však prišli na to, že hovoriaci produkujú viac dysfluencií počas komunikácie so známymi, čo naznačuje, že pri takýchto situáciách sa stávame sebavedomejšími a v prípade problémov s plánovaním reči, môžeme sa spoľahnúť na pomoc nášho známeho.

Dysfluentné chovanie môže byť taktiež ovplyvnené témou rozhovoru. Štúdiá dokazujú, že téma, ktorá je označovaná ako bohatá na slovnú zásob, a tým pádom dáva hovoriacim na výber viac možností, spôsobuje u hovoriacich dysfluentné chovanie. Jedným z príkladov sú napríklad prednášky humanitných odborov, ktoré v porovnaní s prednáškami sociálnych vied alebo prírodných vied obsahovali oveľa viac dysfluencií (Schachter a spol., 1994).

Jedným z faktorov ovplyvňujúcim dysfluentné chovanie, je podľa Branigan a spol. (1999) aj očný kontakt. Zistili, že dvaja hovoriaci, ktorí sa navzájom nevideli, vyprodukovali viac dysfluencií ako tí, ktorí sa videli. To naznačuje, že schopnosť vidieť nášho konverzačného partnera môže viesť k efektívnejšiemu striedaniu pri rozhovore a tým pádom ku zníženiu množstva dysfluencií.

Ako bolo spomínané, dysfluencie sa už dlho nepovažujú za nežiaduce a bezvýznamné prvky ľudskej reči. Mnoho štúdií totiž dokázalo, že dysfluencie plnia rôzne funkcie, ktoré uľahčujú komunikáciu. Arnold a spol. (2003) naznačujú, že dysfluencie plnia dôležitú úlohu pri rozpoznávaní diskurzívne novej alebo diskurzívne známej informácie. Ak je informácia diskurzívne nová, hovoriaci môže mať s jej tvorbou väčšie problémy a často sa preto pred ňou vyskytuje určitý druh dysfluencie. Watanabe a spol. (2008) svojou štúdiou ukázali, že dysfluencie môžu plniť funkciu indikátorov dlhých a zložitých úsekov reči, ako aj iných ťažšie spracovateľných informácií, napríklad menej časté alebo menej pravdepodobné slová.

Dysfluencie nám taktiež pomáhajú účinne koordinovať naše rečové prejavy s prejavmi nášho adresáta. Vyplnené pauzy sa často využívajú na signalizáciu zahájenia prejavu (Clark, 2000), pričom predlžovanie samohlások a spoluhlások naznačujú prerušenie prejavu za účelom upozornenia adresáta, že v určitom štádiu plánovania reči sa vyskytol problém. Clark (2000) taktiež zdôrazňuje, že vyplnené pauzy a opakovania časti slova slúžia ako znamenie, ktoré signalizuje zámer a dĺžku oneskorenia, ktoré nastáva po prerušení.

Ako už bolo spomínané, dysfluencie sú výsledkom problémov, ktoré sa často vyskytujú počas tvorby jazyka a dysfluentné chovanie preto nemôže byť vedome ovládané (Braun & Rosin, 2015). To naznačuje, že užívanie určitých typov dysfluencií môže byť pre každého jednotlivca špecifické. Táto skutočnosť má veľký význam hlavne pre odborníkov na forenznú fonetiku, ako aj pre odborníkov na neurolingvistiku.

Poslednou problematikou, ktorá sa zaoberá tvorbou a významom dysfluencií je ich dopad na poslucháčov. Môžu mať totiž veľký vplyv na to, ako sú hovoriaci poslucháčmi vnímaní. Dysfluentné chovanie môže vytvoriť dojem, že hovoriaci je neúprimný alebo zle oboznámený s predmetom rozhovoru (Fox Tree, 2002).

Dysfluencie majú však na poslucháčov aj priaznivé dopady. Je dokázané, že určité typy dysfluencií zvyšujú u poslucháčov pozornosť (Corley a spol., 2007) a taktiež ovplyvňujú spôsob spracovania a porozumenie prejavu. Pôvodne sa domnievalo, že dysfluencie obmedzujú a dokonca spomaľujú spracovanie a porozumenie prejavu (Fox Tree, 2001), avšak niekoľko štúdií dokázalo, že mnohé typy dysfluencií, napríklad vyplnené pauzy alebo opakovania, tieto procesy uľahčujú. Našla sa však i výnimka, a tou sú falošné začiatky. V štúdií od Fox Tree (2001) sa poukazuje na to, že práve tento typ dysfluencie procesy spracovania a porozumenia prejavu neuľahčuje, ale naopak spomaľuje.

Posledná časť teoretického základu je zameraná na popis oblasti sociolingvistiky, sociofonetiky a sociálnych faktorov, ktoré môžu ovplyvniť dysfluentné chovanie a celkove jazykovú variabilitu. Medzi tieto faktory patria miesto pôvodu, socioekonomické postavenie, úroveň vzdelania, ale aj náboženstvo, rasa a národnostná príslušnosť. V tejto práci sa však zameriavame len na vplyv pohlavia a veku, čo sú ďalšie sociálne faktory, ktoré môžu ovplyvniť výskyt a variabilitu dysfluencií.

Vzťah medzi pohlavím a jazykom je už dlho predmetom štúdií rôznych oblastí, vrátane sociolingvistiky. Mnohé z nich dokazujú, že medzi spôsobom akým hovoria ženy a akým hovoria muži existuje mnoho rozdielov. Badateľné sú vo výslovnosti, či pri výbere gramatických tvarov (Smith, 2002). Taktiež je dokázané, že existuje vzťah medzi pohlavím a dysfluentným chovaním hovoriacich. Štúdie od Branigan (1999), Bortfeld a spol. (2001), a Shriberg (1996) poukazujú na to, že práve muži sú skupina, ktorá tvorí viac dysfluencií. Napriek tomu sa však nedá určiť presný dôvod, na základe ktorého by sme mohli predpokladať, že muži budú produkovať viac dysfluencií ako ženy. Jednou z možností je, že muži využívajú dysfluencie na to, aby sa udržali v konverzácií a aby mohli rozprávať po dlhšiu dobu bez toho, aby ich niekto prerušil. (Shriberg, 1996).

Čo sa týka veku, je známe, že s pribúdajúcim vekom sa reč mení a na základe určitých znakov vieme rozlíšiť medzi rečou rôznych vekových skupín. Veľa štúdií poukazuje na to, že s pribúdajúcim vekom dochádza k zlepšeniu jazykových schopností. Zlepšuje sa schopnosť definovať slová (Obler & Albert, 1984; citované v Bortfeld a spol., 2001), slovná zásoba je bohatšia (Harwood, 2006; citované v Meyerhoff, 2006), a zvyšuje sa užívanie zložitých syntaktických foriem (Obler & Albert, 1984; citované v Bortfeld a spol., 2001). Bolo však zistené, že určité schopnosti sa vekom znižujú. Môžu sa vyskytnúť väčšie problémy s prístupom ku slovám, čo zvyšuje počet dysfluencií v reči. Bortfeld a spol. (2001) vo svojej štúdií skutočne dokázali, že starší ľudia produkujú viac dysfluencií ako mladší.

Dostávame sa k praktickej časti, ktorá skúma vplyv dvoch faktorov, pohlavia a veku, na frekvenciu výskytu a variabilitu rôznych typov dysfluencií v spontánnej reči rodených hovoriacich angličtiny z Anglicka. Je založená na troch hypotézach, ktoré sme sformulovali na základe získaných poznatkov z predchádzajúcich výskumov:

H1: Muži tvoria viac dysfluencií ako ženy.

H2: Skupina starších tvorí viac dysfluencií ako skupina mladších.

H3: Skupina tvoriaca najviac dysfluencií je skupina starších mužov.

Metóda výskumu bola založená na niekoľkých krokoch. Prvým krokom bolo získanie rečového materiálu, ktorý pozostával z 32 nahrávok rodených hovoriacich angličtiny z Anglicka, pochádzajúcich z rôznych oblastí, avšak momentálne žijúcich v regióne Juhovýchodného Anglicka alebo Veľkého Londýna. Nahrávanie prebiehalo vo forme jednoduchého rozhovoru, ktorý bol založený na otázkach týkajúcich sa uchádzačov. Uchádzači boli rozdelení do štyroch skupín na základe pohlavia a veku: muži vo veku 17-30, ženy vo veku 17-30, muži vo veku 35-72, ženy vo veku 35-68. Nahrávky boli následne ortograficky prepísané a zanalyzované pomocou programu Praat, kde sme vyhľadávali sedem typov dysfluencií. V ďalšom kroku sme pomocou dvoch skriptov vypočítali celkový počet dysfluencií a následne celkový počet slov vyslovených každým jednotlivcom. V programe Microsoft Excel sme vypočítali frekvenciu výskytu dysfluencií na 100 slov. Použili sme nasledujúci vzorec: celkový počet dysfluencií / celkový počet slov x 100. Týmto spôsobom sme získali výsledky, ktoré boli následne spracované pomocou štatistických post-hoc t-testov a testu ANOVA.

Celkové výsledky poukazujú na to, že vplyv pohlavia na výskyt dysfluencií nie je štatisticky významný. Muži síce v priemere vyprodukovali viac dysfluencií ako ženy, rozdiel medzi nimi však nebol štatisticky významný a tým pádom sa naša prvá hypotéza nepotvrdila. Druhá hypotéza však potvrdená bola. Skupina starších skutočne produkuje viac dysfluencií ako skupina mladších a ANOVA test ukázal, že rozdiel je štatisticky významný. Čo sa týka samotnej interakcie medzi vekom a pohlavím, bolo ukázané, že nie je štatisticky významné.

V ďalšej analýze sme sa zamerali na rozdiely medzi dvoma skupinami rovnakej vekovej kategórie ale opačného pohlavia a na rozdiely medzi skupinami rovnakého pohlavia ale rozdielnej vekovej kategórie. Na určenie štatisticky významných rozdielov sme previedli post hoc t-testy. Čo sa týka veku, významné rozdiely boli nájdené len u mužov. Starší muži produkovali významne viac dysfluencií ako mladší muži. Aj napriek tomu, že v priemere produkovali staršie ženy viac dysfluencií ako mladšie ženy, rozdiel medzi týmito skupinami nebol štatisticky významný.

Testy ďalej ukázali, že pohlavie nemá vplyv ani na jednu vekovú kategóriu. Medzi skupinou mladších mužov a mladších žien sa nenašiel skoro žiadny rozdiel. Rozdiel medzi staršími mužmi a staršími ženami bol badateľne vyšší, avšak stále nie štatisticky významný.

Pri analýze sme sa taktiež zamerali na vplyv veku a pohlavia na jednotlivé typy dysfluencií. Medzi siedmim typmi, na ktoré sme sa zamerali v tejto štúdií, štyri boli ovplyvnené buď vekom alebo pohlavím. Žiadne štatisticky významné rozdiely sa nenašli pri používaní vyplnených páuz, tichých páuz, a opráv, avšak boli nájdené pri používaní falošných začiatkov, opakovaní, pri predlžovaní samohlások a predlžovaní spoluhlások. Najzaujímavejším výsledkom bolo, že používanie falošných začiatkov je štatisticky významné pre skupinu mladších, a to aj napriek tomu, že predchádzajúce výsledky dokázali, že produkcia dysfluencií je významnejšia u skupiny starších. Zaujímavým výsledkom bolo aj to, že používanie predlžovania samohlások je štatisticky významné pre ženy, pričom celkové výsledky nenašli žiadny rozdiel medzi skupinami rozdielneho pohlavia. Na konci praktickej časti sme sa krátko zamerali aj na to, či sa medzi jednotlivcami nachádzajú určité individuálne a skupinové preferencie. V časti Diskusia sa k získaným výsledkom vraciame a ponúkame ich interpretácie. Taktiež predkladáme návrhy pre ďalšie štúdie, ktoré by mohli naše poznatky o danej problematike ešte viac prehĺbiť.