Palacký University Olomouc

Faculty of Arts

Department of English and American Studies

Hesitation Markers and Fillers

in the Speech of Interpreters

(Bachelor's Thesis)

2022   Karolína Řeháková

Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

Hezitační zvuky a výplňková slova
v řeči tlumočníků

(Bakalářská práce)

Autor: Karolína Řeháková

Studijní obor: Angličtina se zaměřením na komunitní tlumočení a překlad

Vedoucí práce: Mgr. Markéta Janebová, Ph.D.

Olomouc 2022

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně a uvedla seznam veškeré použité literatury.

V Olomouci dne ....   Podpis: ....

List of Abbreviations

HP	Hesitation phenomena
HM	Hesitation markers
SD	Speech disfluency
FS	False start
S-R	Self-repair
OU	Original utterance
FP	Filled pause
SI	Simultaneous interpreting
AIIC	Association Internationale des Interprètes de Conférence (International Association of Conference Interpreters)
SL	Source language
TL	Target language
EVS	Ear-Voice Span
H/I	Hesitation per interpretation
F/I	Filler per interpretation

Table of contents

Introduction

Speech disfluencies, referred to in this thesis as hesitation phenomena, such as filled and unfilled pauses, false starts, repetitions or filler words among others, are natural occurrences in spontaneous speech. It is impossible to fully omit such phenomena without sounding artificial or rehearsed. In basic interpreter training, students are taught to actively avoid such utterings for the sake of fluency and accurate information transfer into the target language. The act of simultaneous interpreting encompasses a wide range of processes, such as listening, understanding, reformulating and self-monitoring. All of these have to be performed within a limited time frame, oftentimes even overlapping, and thus can be highly mentally taxing. Due to this, a complete deletion of hesitations is difficult to achieve.

In the theoretical part of this thesis, I pay attention to past research on hesitation phenomena in spontaneous speech. Many scholars from different fields have scrutinized them and thus adopted a slightly different terminology. I provide categories for the most commonly occurring phenomena as suggested by Rose (1998) and other authors, working with the data acquired from the speeches of interpreters in the European Parliament, which is the main subject of the practical part. A chapter of the theoretical part describes the process of simultaneous interpreting, looking closely at the underlying mental processes with the help of Gile's Efforts Models and various interpreting strategies, looking at how those might tie into the production of hesitations. I also comment on the possible pragmatic role of hesitation markers in signalling an upcoming delay in speech as suggested by Clark and Fox Tree (2002), and whether they should be considered non-linguistic elements or words on their own.

In the practical part, I examine speeches of Czech simultaneous interpreters in the European Parliament. I inspect individual hesitation phenomena, mainly focusing on fillers, with close attention to their place of occurrence within a sentence as well as looking at the surrounding speech environment and giving factors that might be involved in their production (for example dealing with the tempo of the original speech, difficult terminology, reformulation or omissions of certain pieces of information and others). Influence of the original speech will also be considered.

The goal of this thesis is to put together a comprehensive overview of past research done on the topic of hesitation phenomena. It works with the premise that speech of

simultaneous interpreters is mostly spontaneous, with the interpreting itself taking place "online", without prior preparation. It aims to show that some underlying factors for the dispersion of hesitations can be found.

1   Hesitation phenomena

Every communication aims to be as fluent as possible – to retain the greatest amount of information and to be clearly understandable to the listener. Flawless conversations are often presented to us by the media, done by professional public speakers. In this setting, we scarcely encounter imperfect speech that would be comparable to how we talk every day, and elements such as *um* and *uh* are often considered undesirable and with a "disagreeable status" (Fox Tree 2001, 320).

We are aware of a so-called "ideal delivery" when speaking, defined by Clark (2006, 245) as "a single action with no suspensions – no silent pauses, no fillers, no repeats, no self-corrections, no delays except for those required by the syntax of the sentence." Kosmala and Morgenstern (2018, 2), however, note that "speech disfluency is an inherent human phenomenon as speakers typically do not know in advance what they are going to say and how they are going to say it." O'Connell and Kowal (2005, 557) agree with this notion and add that speaking cannot occur in an ideal continuous flow due to these reasons:

> (1) Every speaker must breathe, and breathing inevitably disrupts the flow of speech. (2) The capacity of listeners to understand is limited by the density of speech per time unit; intelligibility is diminished by failure to interrupt speech. (3) Language is reductively dialogical; listeners turn into speakers and speakers in turn into listeners. Turn-taking disallows continuity. (…)

This implies that disfluent features, although generally perceived as redundant elements, are a natural part of spontaneous speech. Many authors have thus decided to investigate them through both descriptive (e.g. Maclay and Osgood 1959) as well as experimental (e.g. Goldman-Eisler 1961) types of research.

What does it mean to hesitate in speech? Lickley (2015, 21) answers this question in simple terms:

> Hesitation usually involves the temporary suspension of flowing speech. It may be achieved by stopping altogether and remaining silent for a moment, by prolonging a syllable, by producing a filled pause or a lexical filler, or by repeating the onset of the current phrase. It may also be achieved overtly in a

phrase openly expressing the speaker's uncertainty. Combinations of these phenomena are normal.

In this description, we can already recognize some individual categories of hesitation phenomena (HP) that will be described in the Typology section. Rose (1998) also remarks that HP slow down the transmission of lexicalized information, and Gósy (2007, 93) adds that these elements "do not add propositional content to an utterance." Fox Tree (1995) gives a figure of approximately 5 of every 100 words being influenced by some type of disfluency, and Volín (2016, 54) recorded that in the Czech language, these sounds account for 20 to 30% of the overall speaking time.

The topic of speech disfluencies (SD) is relevant to a wide range of disciplines, for example medicine (studies on aphasiac patients or children with developmental issues, e.g. Quirting 2019), psychology and psycholinguistics (e.g. Mahl 1956) or speech recognition software and artificial intelligence (e.g. O'Shaughnessy 1993). Naturally, the more authors pay attention to SD, the more varied the terminology, descriptions, definitions and typology are, as Boughaba (2021, 16) states: "there is a lack of consensus over the definition of speech disfluencies since scholars have examined the phenomenon in different disciplines and from different perspectives." Additionally, Shriberg (1994, 11) mentions a number of terms that have been used for the same phenomenon: "disfluencies, (self)repairs, (self)corrections, reformulations, restarts, edits, and hesitations," with some of those terms used as cover terms and others only as names for subordinate categories. In this thesis, I will be using the term "hesitation phenomena" (HP) as an umbrella term encompassing different categories that are elaborated on in Chapter 1.1, and this term will be used interchangeably with "speech disfluencies," (SD) describing the same hypernym.

1.1 Typology

As has already been mentioned, different authors use different categorizations for individual disfluencies and hesitations. Shriberg (1994, 9) goes as far as to say that "it is probably not an exaggeration to say that there are as many different classification systems as there are studies involving classification." Most authors deal with a narrow set of HMs according to the data they are examining, and those are put into general categories. To give an example of such categories, let me mention several authors; Boonsuk et al. (2019) work with filled pauses, small words, and repeats. Wiese (1984), aside from filled pauses, also works with repetitions and corrections. Maclay and Osgood (1959) define four types of hesitations: repeats, false starts, filled pauses and unfilled pauses. Mahl (1956), instead of a filled pause, marks an "ah", and his other categories include sentence correction and incompletion, repetition, stutter, intruding incoherent sound, which is also similar to a filled pause, tongue-slip, and omission, which could be called an unfilled pause (silence). Jean E. Fox Tree, a prominent name in the study of collateral signals in spontaneous speech, works with several types of HMs throughout her work: "uh" and "um" (Fox Tree 2001; Clark and Fox Tree 2002), false starts and repetitions (1995) or discourse markers such as *well*, *I mean* and *oh* (1999). And the last example is Hieke's (1981) unique view on taxonomy of hesitations, listing two superordinate groups: "stalls," which include silent pauses, filled pauses, prospective repeats and syllabic prolongations, and "repair," which encompasses false starts, retrospective repeats or bridging.

Some of the above-mentioned categories overlap and are commonly encountered, some are unique and used more sparsely. In my opinion, the most comprehensive yet easily understandable system of typology has been summarized by Rose (1998) in his Master's thesis *The Communicative Value of Filled Pauses in Spontaneous Speech*. This classification most closely reflects the findings in my own dataset as well, therefore I have decided to introduce these groups because I will be working with them further in the practical part. Rose identifies the following categories:

1.1.1 *False starts*

In beginning their utterance, if the speaker says a few words and then stops themselves mid-sentence, it is considered a false start. Maclay and Osgood (1959, 24) consider false

starts to be "all incomplete or self-interrupted utterances." Rose (1998, 9) gives this example (note that the transcriptions used by Rose are simplified in this work):

(1)   (…) lightning has not struck me yet // as far as what I finally want
       to do or what I'm really capable of doing … // er … we'll see //
       [I still have] I'm twenty seven now // so I still have a few years
       to figure out some things (…)

The first attempt at lexicalizing the thought is discarded: *I still have.* This is then followed by *I'm twenty seven now* which could be called a "fresh start" (Fox Tree 1995). Maclay and Osgood (1959) also differentiate between a "retraced false start" and "non-retraced false start," according to whether the speaker "backed up in an attempt to correct one of the words he had already used," as in *I saw a very big // a very small boy.* In (1), the false start is retraced to some degree in the next sentence: *so I still have a few years (...),* but this retraction does not immediately follow the false start. Rose (1998) also notes that FS can be followed either by a revised attempt at lexicalizing them correctly, such is the case in the example (1), or abandoned altogether, thus, if it occurs in a dialogue, releasing the conversational turn.

An example of a false start from my dataset is as follows:

(2)   (…) vyznat v aktivech které drží jiné než americké institute //
       [myslím že naš*] můžeme říci že naštěstí jsme neviděli až
       takovou destrukci tady v Evropě (…)[1] (_18282000)

The interpreter starts the segment *myslím že naš-* with the intended *naštěstí* half-articulated. This word is then used in the next part: *můžeme říci že naštěstí.* False starts, as the name suggests, occur only at the start of utterances, and by that characteristic are different from self-corrections.

---

[1] Interpreted from the source speech: "One thing we can be thankful for in Europe is we have not seen the same scale of destruction as has happened in the United States." No variation of "I think that" (*myslím, že*) can be found in the original.

1.1.2 Repeats

Maclay and Osgood (1959, 24) define repeats as "all repetitions, of any length, that were judged to be non-significant semantically" and that they "can vary from a single phoneme to an extended stretch that could, theoretically, be of any length." They note an important distinction between two types of repetition; the first one changes the meaning of the sentence, as in *very very big boy*, with the repeated *very* serving as an intensifier, thus describing the boy as huge. The second case of repetition does not have any impact on the meaning, as in *I I saw a very big boy*, which can be taken for a marker of hesitation. Lickley (2015) notes that repetitions also occur normally in fluent, non-hesitant speech. What distinguishes hesitant and "non-hesitant" repetitions is their prosody and subsequent presence of another hesitation device, most often silent pauses or prolongations (lenghtenings).

Boonsuk et al. (2019, 138) come with a similar and simple definition of "immediate repetition of a sequence of one or more words," and Wiese (1984, 18) mentions repetitions as "the unchanged re-occurrence of some substring of an utterance." Rose describes repeats as lexical items that are said again, but also that occur mid-sentence, distinguishing them from restarts. This is the example Rose gives (1998, 10):

(3) (…) I just think of always getting the **[best possible] best possible** results with my students (…)

In my example of a repetition, only one word is repeated in each instance:

(4) (…) především musím zdůraznit skutečnost že tento orgán je nástrojem **nikoli= nikoli** cílem protože ten orgán není ničím jiným než **nástrojem= nástrojem** který má zlepšit regulatorní konzistentnost (…)[2] (_16061000)

Hieke (1981) gives two types of repetitions: *prospective* and *retrospective. Prospective* repeats are anticipatory. They are linked to the planning process and serve as means to

---

[2] Interpreted from the source speech: "But above all, I must stress the fact that the body is a means to an end and not the end in itself. The body is nothing more than an instrument in order to improve regulatory consistency."

gain time in the search for proper lexical items. *Retrospective* repeats "perform primarily a bridging function to prior speech segments which have become separated from the rest through intervening time (due to pauses and other hesitations)" (Hieke 1981, 152). In the case of my example, the first repetition *nikoli= nikoli* would be an instance of a retrospective repeat, since there is a short silent pause in-between them. This also conforms to the need of the interpreter to pause and listen to the original speaker, in order to successfully continue the interpretation.

1.1.3 Restarts

Oftentimes, authors group restarts under more general categories, such as repetitions or aforementioned false starts, as is the case of O'Shaughnessy (1993), who uses the terms fresh start and restart interchangeably. Rose (1998) lists it separately but his entry on this category is short. Restart is a simple case of hesitation in which the speaker begins their utterance, then stops themselves in the middle of it, returns to the beginning and iterates the same words again. Restarts are similar to fresh starts since both of these phenomena occur at the beginning of utterances, and another similarity can be drawn between restarts and repetitions because both use the same information again.

Rose (1998, 10) provides this example:

(5) (…) but yeah **[my first r-] my first reaction** to that erm was a reaction to myself (…)

Observing restarts in my dataset proved to be a difficult task. Most of such phenomena are a cross between self-corrections, false starts and repetitions. In the case of interpreting, the beginning of utterances and sentence boundaries are not always clear due to the rapid simultaneous mode. Because of the rarity of this category, I have decided not to record it, instead grouping such instances into the three above-mentioned categories according to their similar characteristics.

1.1.4 Self-corrections

Self-corrections, also called self-repairs (Levelt 1983), happen when the speaker identifies a mistake in one word or a stretch of words they have uttered, and goes on the repair the utterance. Rose (1998, 10) describes self-corrections as "utter[ing] one word,

and then a replacement which is to be understood to constitute a retraction of that word."
In concordance with this description, he provides this example:

(6)　(…) I teach only the **[fifth] five**-year olds (…)

I believe such definition to be too narrow; it should not involve only whole word retractions or replacements. Speakers can stop themselves mid-word, realizing their mistake and self-correcting it. Fox et al. (2009, 59) use a broader description of self-repair as "the process by which speakers stop an utterance in progress and then abort, recast or redo that utterance." Boughaba (2021, 17) reflects this statement, describing repairs in similar words as "occur[ing] when the speaker notices that there is an error in his utterance, and he tries to repair himself before finishing the words or phrases." An example of a self-correction from my dataset that is not a retraction of a whole word, but rather self-correcting while continually speaking, would be this:

(7)　(…) evropská ekonomika bude potřebovat obrovské **instit- @eh**
　　in- @eh: investice v budoucnosti takže bez uvedených fondů i
　　soukromých fondů by: se: Evropa ještě pomaleji vzpamatovala
　　(…)[3] (_18282000)

The interpreter started the utterance with *instit-* followed by a short filled pause @eh, attempted to self-correct with *in-* that was left unfinished, followed by a long filled pause @eh: (the system of labelling of HMs in this thesis is explained in Chapter 4.1). Finally, the interpreter arrived at the intended *investice*. This misplacing likely happened due to the similarly sounding *instituce/investice* pair, each, however, having a different meaning. In the context of European Parliament where both expressions are commonly used, this substitution is understandable, and this error is likely tied to the word retrieval process.

Levelt (1983, 45) delves deeper into the underlying processes that tie into the production of self-corrections, as described here:

---

[3] Interpreted from the source speech: "Let me be clear, the EU economy is going to need massive investment in the time ahead. Without sovereign wealth funds, private equity and alike, Europe's recovery=recovery from today's turmoil would be all the slower." Note that the speaker makes a repetition, which could be tied to the interpreter's lengthening, because she is waiting for the next words to come. The original speaker also has a strong accent, adding to the difficulty of the interpretation.

Self-correction in speech results from a complicated interplay of perceptual and productive processes. In order to make a repair, the speaker must, firstly, notice some trouble and interrupt his or her flow of speech, and, secondly, create a new utterance, which takes care of the trouble and its potential consequences for the listener.

He uses the term "self-repairs" for this hesitation phenomena, and he pays attention to the creation process of a S-R, dividing it into three phases. In the first phase, the speaker monitors their own speech and is able to interrupt it when they detect trouble. The second phase involves a type of hesitation: most commonly pausing, but the speaker can also use a so-called "editing term." An editing term is a means of signalling this trouble in speech production to the listener; it can be a filled pause (e.g. "uh") or a word (e.g. "sorry," "well," "rather"). The third phase then consists of the repair proper, arriving at the intended "correct" linguistic element(s).

Apart from the creation process, Levelt (1983, 44) also describes the composition of a self-repair. As a model sentence, he uses this:

(8) (…) Go from **left** again to, **uh** … from **pink** again to blue (…)

The first part of this segment is called the "original utterance" (OU). In this example, it is *Go from left again to*. In this original utterance, a "trouble spot," or more technically, a "reparandum" can be found, in this case, it is the word *left*: the speaker produced this word erroneously and will shortly attempt to correct it. A repair can start directly at the reparandum, which is the case in the example (7) from my dataset, or the speaker can backtrack to an earlier point. After three more syllables (*again to*), the speaker has realized this mistake and arrives at a "moment of interruption." The natural flow of speech is halted in this instance, and the speaker produces a hesitation, specifically a filled pause "uh". This can also be called the aforementioned "editing term." What comes after this "editing phase" is the repair proper. In this case, the speaker retraces in the span of one syllable (*from*), which they repeat, and replaces the incorrect *left* with the intended *pink*, which Levelt calls "alteration." The speaker then finishes with *pink again to blue*, and the

repair is thus completed. This structure, along with the individual phases of a self-repair, are best described visually in (9):
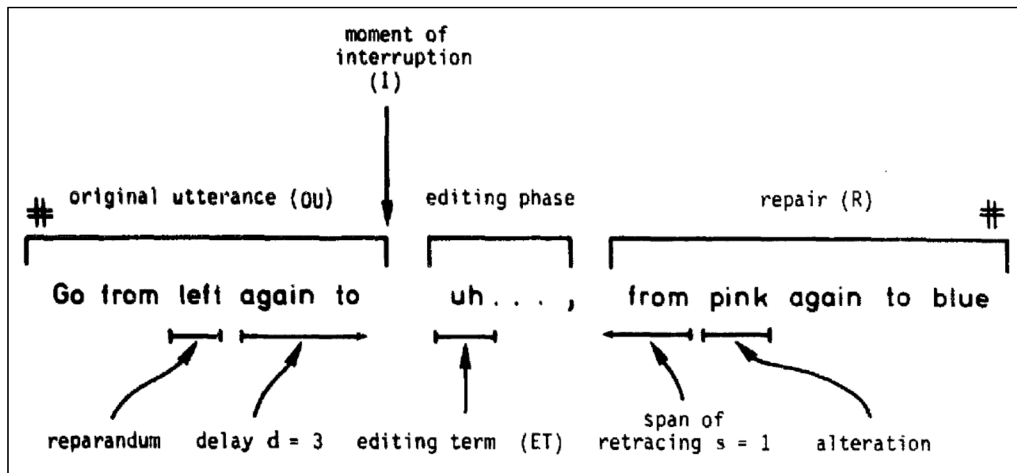
(9)   The structure of a self-repair as described by Levelt (1983).

Shriberg (1994) suggests some alterations to this system. According to her, the term "reparandum" should be used for the entire segment of the OU that is meant to be deleted, rather than just the one element perceived as "incorrect." She renames the "moment of interruption" to "interruption point" which is equivalent to the "cutoff" used by Blackmer and Mitton (1991). She also notes that the place of the interruption point is only a surface feature; the actual detection of an erroneous element in the speaker's mind might have happened earlier than at this specific point. Shriberg (1994) introduces a novel term, the "interregnum," which is the equivalent to the "editing phase" Levelt (1983) uses. She clarifies that interregnum is a more neutral term, since it "can be used to specify the temporal region from the end of the reparandum to the onset of the repair even if this region contains no editing term, and it does not imply an editing function for the speaker" (Shriberg 1994, 8).

To conclude, Rose (1998) notes that the three preceding categories (repeats, restarts, self-corrections), if grouped together, can also be generally referred to as only "repairs," since the initial "wrong" information is replaced with the "correct" one.

1.1.5   Lengthenings

A lengthening, also called a prolongation (Lickley 2015) or a drawl (Wiese 1984), happens when the speakers draws out the pronunciation of a word past its normal length.

This usually occurs at the end of words. Betz and Wagner (2016, 1) point out that lengthening is a feature that occurs naturally in spoken speech as well, and is "in its default form a cue for perceiving phrase boundaries." Therefore, they distinguish this type from "disfluent lengthening," which is what this thesis is concerned with, and define it as "a marked prolongation of one or more phones, resulting in above-average syllable and word duration" (Betz and Wagner 2016, 1) This is accompanied by an unexpected halting of speech rate, which in turn evokes disfluency and hesitation. Rose (1998, 10) uses this example:

(10) (…) well it goes back **to:** always wanting to be a missionary (…)

To is drawn out past its usual length. A colon **[:]** is placed next to the vowel *o* to indicate such instance; this sign is a simplified version of a length mark **[ː]** that is used in phonetics to indicate vowel length. Another way of marking this phenomenon would be to transcribe it as *tooooo,* but the precise length of the prolongation is not the subject under study in this thesis.

An example of a lengthening from my dataset is this:

(11) (…) jsme omezeni v tom co můžeme navrhnout a nebo **se:**
musíme více zaměřit na spolupráci s ostatními **a: to:** vytváří
rámec našeho veškerého konání (…)[4] (_10324000)

A common instance of lengthening which occurs in English is when *the* is pronounced as *thee* – the reduced vowel schwa (ə) is replaced with non-reduced vowel, in this case (e) (Fox Tree 1997).

1.1.6 Pauses

Pauses are perhaps the most prominent category of hesitation phenomena and the simplest in execution. Without being a linguist or a researcher, even an uninformed listener notices silent, prolonged pauses or their vocalized and lexicalized variants. "The modern pioneer

---

[4] Interpreted from the source speech: "(…) We are at least limited in what we can propose or we have to work together with the other institutions. And that is what sort of what=what is- what forms the framework for whatever we do." The lengthening might be influenced by the original speaker's false starts and repetitions of the word "what".

of the science of pausology" (Dechert and Raupach 2011) is often credited to be Frieda Goldman-Eisler with her experimental research on the duration and distribution of pauses in speech (e.g. 1958).

Pausing occurs in fluent speech as well, most commonly at significant grammatical points (e.g. after sentences). Goldman-Eisler (1958) set a precedent regarding the length of a pause that is not hesitation-related. According to her, such pause is shorter than 250 ms and is usually tied to articulatory adjustments. Many authors adhere to this set length and use it in their work, recording only pauses longer than 250 ms in relation to HP (e.g. Boomer 1965, Quirting 2019).

Two groups of pauses are commonly recognized: unfilled/silent pauses and filled pauses, and this distinction is used by a majority of authors (e.g. Maclay and Osgood 1959, Boomer 1965, Gósy 2007, Lickley 2015). Let me examine both categories separately.

1.1.6.1 Unfilled pauses

Unfilled pauses, as the name suggests, are periods of silence at points where silence is not expected to be present, given the prosody of the sentence. The speaker, in their brief moment of uncertainty in the speech planning process, stops themselves and ponders silently. Rose (1998) does not state any examples of silent pauses, so I will demonstrate with two of my own:

(12) (…) konzi- po @eh konzistenci o upozorňování těch národních @ehm regulátorů podle článku sedm @_ nový mechanismus arbitráže v článku sedm a parlamentu ukazuje že (…)[5] (_16061000)

---

[5] Interpreted from the source speech: (…) to reinforce the article seven consistency procedure for notifying national market, in which, by the way, the body will play its part. Parliament's new arbitration mechanism in the article seven A shows that the Commission and the Parliament (…)" Here, it is clear to see that the pause the interpreter produces is to mark sentence boundaries, mimicking the original speaker's pause as well.

(13) (…) jaksi novou překážkou pro konkurenční boj do budoucna a
máme spoustu důkazů které naznačují že přechod @_ tento @eh
@eh př- přechod @eh m:ůže být problematickým pro nové @_
pro @ehm nové investory (…)[6] (_16061000)

I chose two utterances from the same interpreter in order to see a clear difference. Taking into account the nature of simultaneous interpreting, the silent pause in (12) could be considered an "initial pause," because it occurs in-between grammatical boundaries (sentences). It should therefore not be considered a hesitation, as Tissi (2000, 113) explains: "The initial pause, namely the silence before starting with the linguistic task, was left out [of her analysis], considering that SI requires by definition a certain time-lag which cannot be considered a non-fluency." The two silent pauses in (13), however, are hesitation pauses, given the fact that the following segments are repetitions of what the speaker said before he resorted to pausing. This is a distinction closely tied to the matter under scrutiny in this thesis. In normal spontaneous speech, silent hesitant pauses would be more obvious to the listener.

1.1.6.2 *Filled pauses*

In comparison to their silent counterpart, filled pauses employ a type of sound to signify a moment of hesitation, as Boughaba (2021, 17) explains: "Filled pauses occur when the speaker cannot maintain the flow of speech and introduces sounds such as 'uh' and 'umm' within his utterance instead of silence." Maclay and Osgood (1959, 24) list that filled pauses are "all occurrences of the English hesitation devices [ɛ, æ, r, ə, m]," of which schwa [ə] is the commonest instance. The devices which fill this type of pauses are generally referred to as "fillers." Clark and Fox Tree (2002) are of the opinion that *uh* signals a minor delay, whereas *um* signifies a major delay. Rose (1998, 11) gives these examples of filled pauses:

---

[6] Interpreted from the source speech: "(…) must not become a new and enduring bottleneck for competition in the future. And we have plenty of evidence that the transition to fiber will make the business case for alternative investors much more difficult." The interpreter is likely waiting for the rhematic information that appears at the end of the segment, and in doing so, resorts to omission of several words.

(14) (…) my cousin's daughter came down and said **[er]** princess
diana was in an accident (…)

(15) (…) so it's hard to say **[erm]** probably: the: blame lies with many
different people (…)

An example from my dataset that is similar to (13) also shows filled pauses:

(16) (…) tento **@eh @eh** př- přechod @eh m:ůže být problematickým
pro n:ové **@_** pro **@ehm** nové investory to znamená že
alternativní operátoři musejí investovat do svých **@eh** vlastních
kabelů anebo používat **@ehm:** bitstream **@eh** těch zavedených
hráčů (…)[7] (_16061000)

In (16), we can note two forms of filled pauses: *eh* and *ehm*. The third type which occurs
in Czech is the prolonged nasal *mm*. Similar hesitation sounds are used in different
languages as well. The most common transcriptions of FPs in English are *er* and *erm* in
British English and *uh* and *um* in American English. In German, for example, those
sounds are perceived to be *äh, ähm* or *mm*, and in Italian as *eh, ehm, mm* (Tissi 2000).
Clark and Fox Tree (2002) provide an overview of filled pauses across different
languages:

---

[7] Interpreted from the source speech: "And we have plenty of evidence that the transition to fiber will make
the business case for alternative investors much more difficult. Because unbundling of fiber is currently
neither technically nor economically possible. Which means that alternative operators must invest in their
own fiber or use a bitstream service of the incumbent." This segment is very heavy on terminology and
dense, which might be the reason the interpreter produces more fillers and chooses to omit some parts
altogether.

Language	Fillers	References
German	äh, ähm	Fischer (1999)
Dutch	uh, um	Swerts (1998)
Swedish	eh, äh, ääh, m, mm, hmm, ööh, a, öh	Allwood et al. (1990) and Eklund (1999)
Norwegian	e, e=, e==, eh, eh=, m, m=, m==, hm, mm, em=, øhø, aj	Svennevig (1999)
Spanish	eh, em, este, pues	Brody (1987) and van der Vlugt (1987)
French	eu, euh, em, eh, oe, n, hein	Duez (1982, 1991, 1993)
Hebrew	eh, e–h, em, e–m, ah, a–m	Maschler (1997)
Japanese	eeto, etto, ano, anoo, uun, uunto, konoo, sonoo, jaa	Cook (1993), Emmett (1996, 1998) and Hinds (1975)

(17)　　Fillers in several languages as summarized by Clark and Fox Tree (2002, 92).

1.1.7　Filler words

Rose (1998, 11) does not list filler words as a separate category; he views them as lexicalized variants of filled pauses, in the sense that the "paralinguistic" *uh* and *um* is replaced with words, as seen in these examples:

(18)　(…) a:nd this bandstand also had **[like]** a kitchen area underneath

　　　 so it was a fairly high bandstand (…)

(19)　(…) when people are very old **[you know]** the cars that they like

　　　 the cars that they rode in (…) everything starts to disappear (…)

Rose (1998) also lists other expressions: "well," "so," "okay" or "let's see." Other frequent filler words in English might include *I think, I mean, you see.* Filler words have also been given different names, for example "performance additions" (Clark and Fox Tree 2002), "small words" (Boonsuk et al. 2019), "explicit editing terms" (Boughaba 2021) or "editing expressions" (Clark and Wasow 1998).

In Czech, these expressions are colloquially called "vycpávková slova", "parazitní slova" or "slovní vata". Some examples of Czech filler words include *jakoby, jako, takže, prostě, no, vlastně* or the expletive *vole* (Mikuláštík 2003, 122). An example of the filler word *jaksi* in my dataset is as follows:

(20) (…) komise chce tento zmatek jasně: **jaksi** @eh: vyřešit určitě nechceme soukromý belgický orgán který nemá co dočinění s komunitárním přístupem aby se: **jaksi** zapojoval do evropského rozhodování (…)[8] (_16061000)

In both cases in (20), *jaksi* is accompanied by another hesitation device: prolongation of *jasně:* in front of the first *jaksi*, which is then followed by a prolonged filled pause @eh:.. The second *jaksi* appears likewise in tandem with a prolongation *se:*. This would conform to the definition of hesitation phenomena I use in the beginning of this thesis, in which Lickley (2015) states that "combinations of these phenomena are normal."

---

[8] Interpreted from the source speech: "The Commission wanted to end this confusion by establishing a clearly-defined and accountable authority. We certainly do not want a Belgium [sic] private body, alien to the community approach and the guarantees it provides, to mix into European decision-making."

1.2 Hesitations in discourse

Hesitation phenomena have not been viewed only in negative light and as redundant elements. Questions about their role in communication have also been raised. Gilquin (2008, 2) states that "hesitation markers, by signalling a small delay, ensure that the speaker can keep his/her turn in the conversation and is not interrupted by the other participants." She even claims that "the function of hesitation is crucial as a conversational strategy" (Gilquin 2008, 3). Her research is concerned with hesitation markers among learners of English as a foreign language, and within this context, she says that "in their search for a formulation which is acceptable in the foreign language, they [the learners] are likely to experience many planning problems and, therefore, need techniques that enable them to gain time while they are trying to solve these problems" (Gilquin 2008, 3). In this view, hesitation fillers would be the perfect markers for the speakers to use when having word retrieval problems.

In his early research, Clark (1994, 247) states that "whenever speakers foresee a delay or interruption they cannot prevent, they can help their addressees prepare for it by warning them about it" with the use of the fillers "uh" and "um". He also mentions that "uh" is used to signal short interruptions, whereas "um" signals longer interruptions. In later work, together with Fox Tree (Clark and Fox Tree 2002), they claim that hesitation fillers are not just automatic by-products, but that speakers have control over their production and might use them to indicate for example "that they are searching for a word, are deciding what to say next, want to keep the floor, or want to cede the floor" (Clark and Fox Tree 2002, 73). They propose a "filler-as-a-word hypothesis," with fillers serving as cues to the upcoming delay (2002, 79):

> *Filler-as-a-word hypothesis. Uh* and *um* are interjections whose basic meanings are these:
> (a) *Uh*: "Used to announce the initiation, at *t*('uh'), of what is expected to be a minor delay in speaking."
> (b) *Um*: "Used to announce the initiation, at *t*('um'), of what is expected to be a major delay in speaking."

Apart from fillers, Clark and Fox Tree (2002, 80) extend a similar hypothesis to prolongations:

Prolongation hypothesis. Speakers prolong a syllable or its parts to signal that they are continuing a delay that is on-going at t(syllable).

They consider the fillers *uh* and *um* to be words, and as such, they explain that they conform to the phonology, syntax, semantics and pragmatics of English (Clark and Fox Tree 2002, 104). I believe such stance to be a little extreme, since most authors do not consider hesitation fillers to be words, but rather signals or "paralinguistic cues" (Brennan and Williams, 1995).

To conclude, hesitation markers do have their pragmatic roles in utterances. They are most commonly used to signal that the speaker is preoccupied with retrieving certain words or planning their utterance, and that the speaker does not want to lose their turn and will continue in their speech momentarily.

2   Simultaneous interpreting

So far, I have described various disfluencies in the context of spontaneous, every day speech. Such conversations are characterised by their unpreparedness, informal status and turn-taking in case of dialogues. While using HP in normal speech might be acceptable to a certain degree, interpreting takes place in a formal setting and clear delivery is of great importance since participants rely on the interpreter to render the information correctly and with little to no disturbance. Since this thesis is concerned with hesitation phenomena in the speech of simultaneous interpreters in the context of conference interpreting, I supply a definition of SI taken from the AIIC website[9] (International Association of Conference Interpreters):

> In standard simultaneous mode, the interpreter sits in a booth with a clear view of the meeting room and the speaker. He or she listens to and simultaneously interprets the speech into a target language. Standard simultaneous interpreting requires a booth (fixed or mobile) that meets ISO/IEC standards for sound insulation, dimensions, air quality and accessibility as well as for the appropriate equipment (headphones, microphones).

SI is often described in comparison to consecutive interpreting, in which the interpreter is taking notes while the speaker delivers a part of their speech, and the interpretation takes place after the speaker has finished. This gap between the source text and target text is not present in SI – the interpretation happens in real-time with very small to no delay.

SI is considered to be a highly mentally taxing activity due to the simultaneity of listening to the speech in the source language and speaking in the target language. In order to better understand this "fundamental difficulty" (Gile 1995, 159), I mention Gile's Efforts Model and coping tactics and strategies interpreters can use to deal with this difficulty.

---

[9] https://aiic.org/site/world/conference/glossary. Accessed April 5, 2022.

2.1 The Efforts Model in SI

In observing trainee interpreters and from his personal interpreting experience, Daniel Gile noticed that difficulties arise not only in dense or highly technical speeches as one would predict, but also in slow, drawn-out deliveries where no obvious obstacles can be identified. This led to the creation of the Efforts Model which dissects the individual stages of SI that can happen one after another, but very often overlap. This Model can also be used in consecutive interpreting and translation, but for the purposes of this thesis, I will mention it only in connection to SI. Gile (1995, 161) lists two underlying ideas that lie behind the Efforts Model:

1. Interpretation requires some sort of mental "energy" that is only available in limited supply.
2. Interpretation takes up almost all of this mental energy, and sometimes requires more than is available, at which times performance deteriorates.

The four individual Efforts that can be used to describe the underlying processes are Listening and Analysis Effort, Production Effort, Memory Effort and Coordination Effort.

2.1.1 Listening and Analysis Effort

The Listening and Analysis Effort, also called Comprehension Effort, is closely tied to actively listening for the source text and understanding it. Gile (1995, 162) describes this Effort as "consisting of all comprehension-oriented operations, from the analysis of the sound waves carrying the source-language speech which reach the interpreter's ears, through the identification of words, to the final decision about the "meaning" of the utterance." Gile (1995) also mentions that it is not clear how far the spoken information needs to be understood in order for interpretation to begin, but he suspects that this comprehension "goes at least as far as understanding the underlying logic of each sentence" (Gile 1995, 162).

2.1.2 Memory Effort

The Memory Effort, as the name suggests, is concerned with the interpreter's ability to remember the information they have heard in their short-term memory before they can interpret it. Long-term memory is also important for word retrieval and understanding of the topic at hand. In the simultaneous mode, memory is seen as temporary storage since there is no note-taking and limited options for looking up additional information.

2.1.3 Production Effort

The Production Effort plays an important role in the output phase of interpretation. It is seen as "the set of operations extending from the mental representation of the message to be delivered, to speech planning and the performance of the speech plan" (Gile 1995, 165). The Production Effort is connected to the interpretation act proper.

2.1.4 Coordination Effort

The three preceding Efforts can happen one after another, but the fourth Effort seems to be above this sequence. The Coordination Effort is vital in maintaining balance and self-monitoring, and "corresponds to resources required to coordinate the three other efforts" (Gile 2009, 168).

2.1.5 Equations

With the individual Efforts explained, the process of simultaneous interpreting can be summarized in a simple equation:

(21) **SI = L + M + P + C,**

which can be explained as "simultaneous interpreting equals Listening and Analysis Effort (L) plus Memory Effort (M) plus Production Effort (P) plus Coordination Effort (C)."

In order for this process to be successful, certain conditions have to be met. Gile (1995, 170) states that "at any point in time, one, two, or three of the Efforts are active simultaneously;" it has been proven that interpreters are capable of listening and speaking

at the same time (e.g. Lambert 1992). Thus, the total requirements (TR) for this process are all requirements for each Effort combined together, as follows:

(22) **TR = LR + MR + PR + CR,**

LR = capacity requirements for L,

MR = capacity requirements for M,

PR = capacity requirements for P,

CR = capacity requirements for C.

As such, the total requirements for the interpreting task at hand cannot exceed the available processing capacity the interpreter currently has, as expressed by this equation:

(23) **TR ≤ TA.**

The same equation can be applied for individual Efforts as well, in the way that the requirements for an Effort should not exceed the available capacity for that Effort, for example:

(24) **LR ≤ LA,**

LR = Listening and Analysis Effort requirements,

LA = capacity available for L.

If equations (22) and (23) are not adhered to, saturation occurs; the interpreter cannot meet the total requirements which results in mental exhaustion and inadequate interpreting performance. That is why the Coordination Effort should not be overlooked because it serves as a "spirit level" for the interpreting process, prompting the interpreter to constantly self-monitor their execution. Gile also mentions saturation in connection with his Tightrope Hypothesis, which says that oftentimes interpreters work very close to saturation. This might result in interpreting failures, "not because they [interpreters] do not have the necessary knowledge at their disposal, but because speeches are "too fast" or "too dense", in other words because they do not have the capacity to process them rapidly enough" (Gile 2009, 182).

2.2 Coping tactics and strategies in SI

Apart from understanding the underlying Efforts and distributing the available "mental energy" between them accordingly, interpreters can also use a handful of coping tactics and strategies to ensure a smooth transfer of information and clear delivery. I list three strategies that I consider to be useful but at the same time possibly harmful when it comes to the production of various hesitations.

2.2.1 Anticipation

Anticipation is a very useful tool for interpreters, as it can save time and aforementioned mental capacity. Simply explained, anticipation is an instance in which the interpreter renders the source speech before the speaker actually says it. This cannot be done to its entirety and with longer strings of words, but to a certain degree, anticipation of the incoming information is possible. Gile (2009, 173) differentiates two types of anticipation: linguistic anticipation and extralinguistic anticipation.

Linguistic anticipation is based on "transitional probabilities" with which words follow one another within a sentence. Gile (2009, 173) gives this example from English: "the probability that an article will be followed by a noun or an adjective is high and the probability that it will be followed by another article or a verb is low." Other easily-anticipated constructions are collocations or standard phrases. This can differ in other languages. In Czech, given the declension system, anticipation is possible due to gender and case. If the speaker, for example, says *velká*, we would anticipate a feminine noun to follow: *konference, analýza, instituce, anticipace;* or a plural neuter noun: *velká města.* Declension rules would not allow a masculine or a singular neuter noun to follow: **velká orgán, *velká kuře.*

Extralinguistic anticipation is concerned with the knowledge of the conference situation, its aims or topics, the participants or the unfolding statements. In this context, Gile (2009, 174) defines anticipation as "some knowledge of the probability of the speaker reacting or speaking in a particular way in the context of the situation at hand, not necessarily as the exact prediction of the speaker's words." For example, at the start of a conference, a speaker might begin with *I would like to…,* and it can be anticipated that they would very likely want to say *…thank you all for coming today*, or *…welcome you all on today's event.* Similar anticipations can be made if speakers begin their

sentence with *no*; the interpreter can then expect a disagreement or a different opinion, working with the situational context. Anticipation can also be wrongful, which results in incorrect interpretation and possible hesitation devices, for example false starts and self-corrections.

2.2.2 Manipulating the Ear-Voice Span

The Ear-Voice Span, or *decalage*, is the time lag between the original speaker's utterance and the interpreter's rendition in the target language (TL). The typical length of the Ear-Voice Span that most researchers work with is approximately 2 to 4 seconds (Lederer 1978), but longer EVS has also been recorded, up to 10 seconds (Oléron and Nanpon 1965). Gile (2009, 204) lists Lengthening or Shortening the Ear-Voice Span under his Preventive tactics, which can be used when interpreters are under processing capacity pressure and sense that problems may arise.

According to Gile (2009, 204), by lengthening the EVS, the interpreter puts more space between the spoken information and their reformulation in TL. This puts pressure on the short-term memory, since there is more to remember before it is interpreted, but it may increase comprehension, because the interpreter is capable of understanding the source speech more thoroughly.

On the other hand, Gile (2009, 204) explains that by shortening the EVS, short-term memory requirements are alleviated, but the comprehension of the source speech may be corrupted. The interpreter can be left with sentences which are difficult to render correctly in the TL given different sentence structures and grammatical rules. Additionally, by shortening the EVS too much, interpreters can be bereft of their anticipation potential and may arrive at content-related misunderstandings which then have to be corrected. This can give rise to self-repairs.

2.2.3 Reformulation

Reformulation is an obvious choice because no two languages will render the same information in the exact same way. Reformulation can happen in many forms: Gile (2009, 206) lists many individual Reformulation tactics. Among the easiest ones to execute are using the boothmate's help or consulting documents in the booth. Due to limited time in the simultaneous mode, the boothmate will often only indicate certain reformulations

rather than explain in length, and looking up additional information, most often in glossaries and dictionaries, has to be done quickly as well.

Another reformulation tactic Gile suggests is replacing a segment with a superordinate term or a more general speech segment. This can be done when the interpreter cannot find the proper words at the moment, and has to compensate for that incapability. Gile (2009, 206) gives an example of describing "DEC, IBM, Hewlett Packard et Texas Instruments" more generally as "computer vendors". Using this tactic comes with the loss of the information in its entirety, but interpreters should be careful to still retain the core meaning of the information. Similarly, interpreters can resort to explaining or paraphrasing. This tactic can be highly efficient, but Gile (2009, 207) lists two drawbacks: the time spent on the reformulation, and the possibility of attracting the listeners' attention to the fact that the interpreter is having problems, thus lowering their credibility.

The last tactic that I want to mention in connection to reformulation is omitting the content of a speech segment. This tactic is fairly easy in execution, but choosing which information is vital and has to stay in the rendition and which information can be left out is difficult. In spite of the difficulty, Gile states that "situations (…) where the only possibility of keeping interpreting and serving best the interests of the participants requires forced choices are rife in daily practice, especially with read speeches with dense passages (…)" (Gile 2009, 210). It is up to the interpreter to recognize which information has more value than some other segments; it is usually the core meaning or message. Omission can also happen due to possible jeopardy of the desired outcome, usually when something inappropriate or culturally insensitive has been said. Here, the interpreters' knowledge about the cultural background of their work languages is crucial.

2.3 Hesitations in SI

Research on hesitation phenomena in simultaneous interpreting remains sparse, since the majority of works on the topic of HP are concerned with every-day spontaneous speech. The speech of interpreters is also unprepared, therefore to a certain degree spontaneous, but interpreters usually have information about the conference they are a part of, the topics which will be discussed and possible opinions of the speakers that will take the floor;

preparation is a fundamental part of the interpreting profession. I mention three authors that have described HP within the context of simultaneous interpretation.

I have already mentioned Benedetta Tissi (2000) in the section 1.1.6.1 describing unfilled pauses, and that is also the focus of her work *Silent Pauses and Disfluencies in Simultaneous Interpretation: A Descriptive Analysis*. She comes up with a specific taxonomy with respect to SI consisting of two groups: those are silent pauses, with two subordinate categories being grammatical and/or communicative pauses and non-grammatical pauses, and disfluencies, which include filled pauses and interruptions. The aim of her study, as she describes, is "to analyse whether and to what extent the presence of such occurrences in the ST affects the interpreter's comprehension and delivery" (Tissi 2000, 103). The interpretation is from German to Italian. From her results, she concludes that the effect on the interpreters' delivery "is not as direct as one could assume" (Tissi 2000, 120) and that no clear trends can be identified (Tissi 2000, 122). However, she makes an interesting note on "the communicative, sometimes even strategic use of some non-fluencies" (Tissi 2000, 121), mainly silent and filled pauses before a correction or lenghtenings of the tonic vowel which draw attention to it.

A study from Maria Bakti (2008) compares disfluencies in the output of trainee and professional simultaneous interpreters. This interpretation is done from English to Hungarian. Bakti works with the classification of error-type disfluencies which include restarts, grammatical errors, or false word activations. The results showed that both groups of interpreters produced similar disfluencies, with restarts being the most common, and that "this analysis signals problems at the stages of lexical access and grammatical planning (...)" (Bakti 2008, 12).

The most recent study comes from Boughaba (2021). It investigates speech disfluencies in simultaneous interpretations of spontaneous and non-spontaneous speech from English to Arabic in order to find out whether the degree of spontaneity affects the interpreters' performance. The results showed that the interpretations of spontaneous speech contained longer pauses than the source speeches (Boughaba 2021, 20), which Boughaba accounts to the spontaneity of the delivery. They also showed that the most prominent disfluencies were silent pauses, followed by prolongations and filled pauses (Boughaba 2021, 21).

The research into hesitation-related topics within SI continues to this day and not many overall conclusions can be stated yet, as Tissi says: "further experiments with larger samples will have to be carried out to draw significant conclusions" (Tissi 2000, 122).

3 Methodology

In the practical part of this bachelor's thesis, I analysed several speeches of interpreters in the European Parliament with focus on their usage of hesitations; I have already used examples from those speeches when describing individual HP in the Typology chapter. The debates which were consequently interpreted were held in the European Parliament in the periods between September and October 2008, with one held in February 2009 and one in October 2009, all of them in Brussels.

I was only interested in the English-Czech language combination, with Czech as the interpreters' A language (mother tongue) and English their B language ("a language in which the interpreter is perfectly fluent in, but which is not a mother tongue"[10]). The dataset was kindly provided to my supervisor and me by Prof. Dr. Volker Gast of the Department of English and American Studies at the Friedrich Schiller University Jena in Germany. I would like to take this opportunity to thank him.

In a .zip file, each speech had its own folder. Inside every folder, audio files of twenty-three working languages of the European parliament were present, along with a video file of the original speaker and an .eaf file. To access this file, I worked with the ELAN software, which is a tool for annotation and transcription of audio and video recordings, and is freely available. Through this interface, I could listen to the original English recording as well as the Czech interpretation; these audios could be played separately, and an option for overlap was also available through the settings. I could also listen to another language if I needed a reference (Slovak was available as well). On the left side, as can be seen in (25), a video of the original speaker delivering their speech was visible. Therefore, I could listen to the whole speech and its interpretation with a complete context of the situation. Transcripts of the debates in multiple languages are available on the website of the European Parliament,[11] and I used those when I struggled to understand the speaker, most often when it came to names and abbreviations or because of strong accent.

---

[10] Definition taken from the AIIC website, https://aiic.org/site/world/about/profession/abc. Accessed May 19, 2022.
[11] https://www.europarl.europa.eu/plenary/en/debates-video.html.

(25) ELAN interface.

3.1 Data description

Eleven interpretations of different lengths and numbers of hesitation phenomena were scrutinized. The total time analysed was almost two hours, more specifically 1 hour, 45 minutes and 28 seconds. The number of original speakers was seven, four women and three men, and the number of interpreters was twelve, six women and six men. In the first speech, two interpreters took turns, the exchange happening approximately at the 16:06 mark; I included this speech because these two interpreters did not speak in any other of the examined files. The transcribed Czech interpretations with markings of hesitation phenomena can be found in the PDF Appendix that is attached to this thesis. The inspected ELAN files are available on the physical appendix in the form of a CD-ROM, or upon request from either the author or the supervisor of this thesis.

File name	Original speaker	Interpreter	Length
_16061000	F1	MI1 + MI2	18:21
_22481300	F1	FI1	6:36
_18282000	M1	FI2	13:03
_21373400	F2	FI3	4:34
_10324000	F3	MI3	8:56
_23435000	M1	MI4	8:41
_10223800	F4	FI4	11:15
_16543500	M2	FI5	5:02
_18572300	M1	FI6	8:20
_19024300	M3	MI5	7:50
_09194200	F1	MI6	12:50
Total: 11	3 males 4 females	6 male interpreters 6 female interpreters	01:45:28

(26)   Description of analysed speeches.

3.2   Research scope

The available dataset offers many possibilities for various analyses. Due to the quantity of the data, it was necessary to narrow the research scope for the purposes of this thesis.

As was already mentioned in the introduction, my interest was in the placement of hesitations within utterances (*where?*) and in the underlying factors that were tied to their production (*what are they influenced by?*). From the described categories in section 1.1, the subject under investigation was chosen to be filled pauses with their corresponding fillers. The theory was that these fillers would be the hesitation device that occurs the most, despite the general tendency to avoid their production. The speculated reason for this was because of the need to temporarily alleviate one's mind in order to think of the next segments, but at the same time indicate that the speech has not ended and will continue within seconds.

In the analysis chapter that is to follow, I describe the findings in all eleven speeches. However, in order to answer the questions of where and what might be the

influence of the fillers' production, the scope needed to be narrowed once again. Three interpretations out of the eleven were chosen for these tasks. All three speeches were delivered by the same original speaker. The interpreters therefore had similar conditions when it came to accent, rate of delivery, density and general vocabulary of the speaker.

4 Analysis

Using the ELAN software, I analysed all eleven interpretations. The *transcription_cze* tier contained automatic transcription of whole utterances, but this transcription was very flawed, with some words misspelled, some replaced with a similar-sounding word or some segments missing entirely. It was therefore necessary to correct these annotations which meant closely listening to the whole timespan of one hour and forty-five minutes. Hesitation markers were also not included in this automatic transcription. Marking of all specified HMs and corrections of the interpretations were done manually.

The goal was to transcribe the interpretations in a way that would most closely reflect the reality of what was spoken. That included false pronunciation of intended words, resulting in non-words, for example **opravdovoeá* or **jednotliveí*. Capital letters at the beginning of sentences and in proper nouns were disregarded and written in lowercase. Given personal names were put into square brackets, as in *[pascalina napoletano]*, but geographical names were not put into brackets, as in *basilej*. The letters of abbreviations were kept in uppercase and also in brackets, as in *[ACHT]*. Any numerals were deleted and transcribed in words, including years. Punctuation, such as commas, periods or question marks, were not marked in any way.

nějakou dobu pociťovat začalo to vlastně krytý hypoték a · banky a dění se nestarali o · standarty půjčování protože Mohli ty svoje použít nás akorát sekuritizaci a pak vlastně si získala Respekt abilita pro ty vysoce rizikové produkt · a vlastně je to teď na celém světě že tam jsou tyto produkty aniž by někdo byl udělal pro to za · opravdová hodnocení rizika jak Minulý rok jsme viděli že bylo přímo neuvěřitelné jak málo rozuměli vlastně vedoucí představitelé finančních institucí Jaká rizika podstupují · a to jsem nevybrala potřebovala pořádnou analýzu řetěz Rezek aby měli · představu co představu co jsou tyto velice komplexní produkty ale vždycky se tak · balíčky znovu přebalovala a nikdo pořádně nevěděl · jsou rizika a před měsícem jsem vlastně hovořil o tom že to je jako kdyby letadlo pomalu padalo pomalu havarovalo ale teď už se to dostalo do opravdových obrátek teď a Spojené státy už oznámili že

@(inbr) a důsledky krize budeme ještě nějakou dobu pociťovat @(inbr) @eh začalo to vlastně @eh s krizí hypoték @(inbr) a: · @m banky a jiní se nestaraly o: · s:tandardy půjčování protože mohly ty svoje @ehm poz- @eh použít na sekura- sekuritizaci @(inbr) a: pak vlastně si získala respektabilita i: pro ty: vysoce rizikové produkty · a vlastně je to teď na celém světě že tam jsou tyto produkty @(inbr) aniž by někdo byl udělal pro to zav- · opravdovoeá hodnocení rizika @(inbr) @eh minulý rok @eh jsme viděli že bylo přímo neuvěřitelné jak málo rozuměli @(inbr) vlastně vedoucí představitelé finančních institucí jaká rizika podstupují · a to jsem @eh by bylo potřebovalo pořádnou analýzu rizisk- rizik @(inbr) aby měli: @eh · představu co představu- co jsou tyto velice @(inbr) komplexní: produkty ale vždycky se tak · ty balíčky znovu přebalovaly a nikdo pořádně nevěděl · jaká jsou rizika @(inbr) před @eh měsícem jsem vlastně hovořil o tom

(27)   Comparison: Speech _18282000 with automatic annotation (above) and
corrected annotation with clear markings of HP (below) in the ELAN
interface.

Working within the ELAN interface was very useful when it came to listening to the spoken information. Once I was finished with marking the hesitations and done any needed corrections, I exported the interpretations into a .txt file that could be subsequently opened in MS Word or other document editor of choice. This way, it was much easier for me to read the whole text, and what is more important, effectively search within it and clearly highlight fillers, thus making the job of statistically counting the individual HMs much easier. ELAN also has a search function, but I have found it to be not effective for my purposes.

In order to be able to analyse the place of occurrence of fillers, I added sentence boundaries, marked with the forward slash symbol (/), where the interpreters lowered their voice and a new segment began. With this done, I could see whether a filler occurred at the start of a sentence or somewhere during it. I could have done this in ELAN as well, but in that interface, I focused mainly on the sound and needed corrections.

From the different HP described in the Typology section in the theoretical part of this thesis, false starts, repeats, self-corrections, lengthenings, filled and unfilled pauses were chosen to be recorded. As I already stated when describing restarts, this category was not included in the analysis because of its very similar characteristics to false starts, self-corrections and repetitions. Filler words were also not examined because they are oftentimes regarded as discourse markers which is not the topic of this thesis.

4.1 System of labelling

Before I describe the findings, I will first explain how individual HP within the transcriptions are marked. It was important to decide upon a clear and uniform system of labelling in order to further work with the data effectively and to quickly identify hesitant elements from words in vast strings of text.

In order to highlight the "paralinguistic" filled pauses with fillers and distinguish them from words, the at-sign (@) was chosen, followed by the corresponding form. The at-sign was also used when referring to unfilled pauses, together with the underscore (_) indicating a brief moment of silence. To mark lengthenings, a colon (:) was used because of its similarity to the length mark (ː) that is commonly used in phonetics. For repetitions, the equals sign (=) was chosen for obvious reasons, because the same segment or information is said again. For false starts, which occurred only a handful of times, an

asterisk (*) was used, indicating that a word is started and then aborted. Self-corrections were marked with the help of a dash (-), similarly as a false start, because an initial sound was cut off and a corrected segment was said right after.

The fillers that were used when producing a filled pause are realizations of the monophthong schwa, marked /ə/. The phonological description of schwa is a "lax central mid vowel" or "reduced vowel" (Volín et al. 2013, 32). In simple terms, as found in an online dictionary[12], schwa is described as a "mid-central, neutral vowel sound typically occurring in unstressed syllables in English, however spelled, as the sound of *a* in *alone* and *sofa*, *e* in *system*, *i* in *easily*, *o* in *gallop*, *u* in *circus*." In Czech, schwa is described as "vokál neutrální" (*neutral vocal)* or "vokál redukovaný"[13] (*reduced vocal*).

The markings, together with the name of the hesitation and an example in usage is shown in table (28):

Name	Label	Description	Example
Filled pause	@eh	Short schwa	*a @eh situace se mění rychlostí světla*
Filled pause	@eh:	Long schwa	*cíl parlamentu @eh: zapojit se:*
Filled pause	@ehm	Starts with schwa, ends in bilabial nasal	*tak musejí @ehm si uvědomovat*
Filled pause	@ehm:	Lengthened variant of @ehm	*nebo jiné @ehm: neotevřenosti*
Filled pause	@m	Bilabial nasal without any vowels	*máme tady: @m málo odborníků*
Filled pause	@m:	Lengthened variant of @m	*k těmto otázkám @m: rámcová pravidla*
Unfilled pause	@_	Period of silence	*já se velice těším na naší diskuzi @_ která proběhne*

---

[12] Definition taken from https://www.dictionary.com/browse/schwa. Accessed June 17, 2022.
[13] Both descriptions taken from https://www.czechency.org/slovnik/MONOFTONG?bib=true. Accessed June 17, 2022.

Lengthening	:	Prolonged pronunciation usually at the end of a word	*ztrácíme: důchody (ztrácímeee důchody)*
Repetition	=	Repetition of one sound or a part of word	*un=u=u v Evropské unii*
False start	*	When a word/phrase is started, but aborted	*myslím že naš* můžeme říci že naštěstí*
Self-correction	-	When a word/phrase is cut off and a corrected segment follows	*nová směrnice odkazuje na podpo- na potřebu mediálních orgánů*
Glottal stop	ʔ	Abnormal obstruction of airflow within a word	*začala s hypoteʔční krizi*
Inbreath	@(inbr)	Very noticeable intake of breath in an unusual place	*musí být tady dozor nad přeshraničními @(inbr) finančními institucemi*

(28)   Labelling of individual hesitation markers.

There are two labels in the table (28) that are not considered hesitation devices in this thesis, but are still signalized in the transcriptions. The first one is a glottal stop, marked with its corresponding phonetic symbol ʔ. This marking appeared with a very low incidence, and it exclusively occurred mid-word, five times after a vowel and once after a consonant, in this case the sibilant [s]: *jaʔk, poʔtřebujeme, hypoteʔční, sʔpirále, jednaʔk, přiʔspět.* This mark was used to signalize a "hitch" in the interpreter's voice, when the airflow in the glottis was obstructed, and the result was the word being seemingly "cut off" in the middle. With the goal of transcribing the utterances to most closely reflect the spoken reality, the glottal stop was therefore recorded for the overall clarity of the speeches.

The second label thus far not discussed is inbreath, marked as @(inbr). The abbreviation was put into brackets to be easily identifiable from the various fillers which were also marked with the at-sign. Breathing is a natural part of speaking and is therefore

not looked at as a hesitation in this thesis. Intakes of breath occurring at the start of sentences were disregarded; only those that occurred in an unusual place within a sentence and were loud enough were marked in the transcriptions, possibly having a disruptive effect to the listener. Such inbreaths could also be considered by-products of the interpreters' delivery in the simultaneous mode which can be hasty.

4.2 Overview

Across all eleven interpretations, 1230 hesitation markers were identified. 702 of those were filled pauses with fillers, confirming the theory that they would be the most frequently used hesitation device. The second group with the highest number of instances were lengthenings, occurring 256 times, and the third group were self-corrections, appearing 157 times. The hesitation that occurred the least amount of times was a false start with only 2 instances across two different speeches. From now henceforth, the interpretations will be referred to as "Speeches" with a number according to the order in which they appear in table (29):

File	Filled pause	Lengthening	Self-correction	Repeat	False start	Unfilled pause	H/I
_16061000	122	61	31	20	1	32	267
_22481300	11	1	4	1	0	21	38
_18282000	90	41	40	0	1	0	172
_21373400	89	5	10	3	0	1	108
_10324000	27	44	5	0	0	6	82
_23435000	22	17	8	0	0	8	55
_10223800	84	19	7	0	0	8	118
_16543500	39	6	10	0	0	0	55
_18572300	13	15	17	6	0	0	51
_19024300	57	30	7	0	0	1	95
_09194200	148	17	18	1	0	5	189
Total:	**702**	**256**	**157**	**31**	**2**	**82**	**1230**

(29)   Numbers of hesitations in individual interpretations.

The last column named H/I lists the number of hesitations per given interpretation. The interpretation with the highest number of hesitations is Speech 1 with 267 of them, but it is also the longest utterance, spanning over 18 minutes. The interpretation with the second highest number of HP is Speech 11 which is almost 13 minutes long and has 78 less hesitations than Speech 1. Because of the different lengths of the interpretations (which can be found in table (26)), I needed an instrument with which I could compare the number of hesitations accounting for the different times. For this, I calculated the rate of a hesitation per second as number of hesitations divided by seconds. With Speeches 1 and 11, I arrived at the same rate of 0.24 hesitation per second, meaning that the amount of hesitations accounting for the different time spans of the interpretations is comparable.

Looking at the other end of the ranking, the interpretation with the least amount of HP is Speech 2, having 38 hesitations. This interpretation is 6 minutes and 36 seconds long with the rate of hesitation per second being 0.09, very low incidence. This interpreter used more unfilled pauses (21) than filled pauses (11), indicating that the interpreter preferred momentarily halting their speech and pondering in silence rather than using filler sounds.

Interestingly, Speech 2 is not the shortest one – that is Speech 4 with 4 minutes and 34 seconds, and it contains 108 hesitations, ranking in fifth place when it comes to the number of HP. Inspecting once again the rate, I found out that it is 0.39, meaning that despite the shorter time span, a hesitation occurred more often in Speech 4 in comparison to Speech 3.

What I would also like the point out is that the interpreter in Speech 3 used the largest amount of self-corrections (40), and also a significant amount of filled pauses (90), the third highest number. Due to this, her delivery appeared to be "chopped up" and gave an impression that she is unsure of herself. This might have in turn lowered her credibility as an interpreter and possibly had a disruptive effect on the listener, since a clearer delivery is easier to follow and understand.

As a last comment, I would like to point out that the probable incidence of HP is of course tied to the original speech. If it, for example, contains more terminology or idiomatic expressions, the interpreter is likely going to have to exert more mental energy and in turn resort to hesitation devices in order to, as discussed in chapter 1.2, signal that a short or long delay in their speech is coming. The usage of HP, more specifically fillers that I will be paying closer attention to, also likely conforms to some personal preference, as can be seen with Speech 5, where the interpreter produces 27 fillers but 44

lengthenings, suggesting that he would rather prolong a word at its end than use a filler sound. Such theories might be an interesting topic for future studies.

4.3 Fillers

The analysis further focuses on fillers used in filled pauses. Three main phonologic realizations of fillers were identified in the eleven interpretations, with three lengthened subtypes: @eh (@eh:), which is realized as the phoneme schwa /ə/, @ehm (@ehm:), starting also with schwa and ending with the bilabial /m/, and @m on its own, with no vowel present (and its lengthened variant @m:). Out of these three types, the first short type @eh was used most often, with 518 occurrences. The second most-used type of filler was its lengthened variant @eh: with 75 instances, and third ranked the @ehm variant, appearing 54 times. The least recorded phonologic realization of a filler in this sample of speeches was the lengthened bilabial @m: with only 2 instances.

File	@eh	@eh:	@ehm	@ehm:	@m	@m:	F/I
_16061000	85	7	23	5	1	1	122
_22481300	10	0	1	0	0	0	11
_18282000	65	12	9	0	4	0	90
_21373400	45	0	11	0	33	0	89
_10324000	18	5	2	0	1	1	27
_23435000	18	4	0	0	0	0	22
_10223800	75	4	3	0	2	0	84
_16543500	28	8	0	0	3	0	39
_18572300	8	1	2	0	2	0	13
_19024300	41	16	0	0	0	0	57
_09194200	124	18	3	0	3	0	148
Total:	**518**	**75**	**54**	**5**	**49**	**2**	**702**

(30)   Numbers of fillers across inspected speeches.

The three highlighted interpretations, Speech 3, Speech 6 and Speech 9 were the interpretations chosen for further analysis of placement and associated factors with the production of fillers. These three speeches were conducted by the same original speaker

and they were interpreted by three different interpreters (two women and one man), one per each speech. Although the interpreters have the same conditions when it comes to the Irish accent of the original speaker and his rate and delivery of speech, the original speeches have different lengths: Speech 3 is 13 minutes and 3 seconds long, Speech 6 is 8 minutes and 41 seconds long, and Speech 9 is 8 minutes and 20 seconds long, comparable to Speech 6.

4.3.1 *Place of occurrence*

Let me turn the attention to the placement. Different authors theorized about the place of occurrence of fillers within utterances. Boomer (1965, 148), for example, presents one general hypothesis relating to all forms of HP which says that "hesitations in spontaneous speech occur at points where decisions and choices are being made." This possibly suggests that this position might be at the start of utterances since a speaker has to decide how they are going to start their utterance and have to choose the words in their mind before they produce them out loud.

The analysis of the placement of fillers undertaken in this thesis is similar to Boonsuk et al.'s (2019) research. They examined not only fillers but also small words and repeats. Their results showed that within their sample of 15 conversational extracts, the specified hesitations occurred the most in the middle position.

In my analysis, a "beginning position" was chosen to recorded be when a filler occurred within a range of three spaces from the onset of a sentence, after a sentence boundary marked with a forward slash (/). If the filler occurred later in the sentence, it was deemed to be "middle position." A possible "end position" would be accounted for at the very end of the sentence, but no such instance was recorded. To be clear, let me demonstrate the differences with two examples taken from Speech 3:

(31) / před **@eh** měsícem jsem vlastně hovořil o tom (…)

(32) / **@eh** v minulý **@ehm** roku jsem vás tady informoval na plenárním zasedání (…)

The filler in example (31) occurs in the second place within the sentence and is therefore accounted for as appearing in the beginning position. In example (32), the very first item

after the sentence boundary is a filler @eh, and could be described as appearing in the "initial position," that is the very beginning of the sentence, but as described above, in this thesis, this is recorded as the beginning position. Continuing in the sentence, there is another filler, @ehm, but this one appears in the fourth position. It is therefore viewed as appearing in the middle position. Behind the production of this filler is probably the incorrect Czech declination of the adjective *minulý* if the interpreter wanted to use *roku* next.

With these definitions of beginning and middle position, an analysis of all 125 fillers was conducted. These were the results I arrived at:

File	Beginning position	Middle position
Speech 3	13	77
Speech 6	1	21
Speech 9	0	13
Total: 125	**14**	**111**

(33)   Placement of fillers within utterances.

Similarly to Boonsuk et al.'s (2019) results, the analysis showed that the majority of fillers occurred in the middle position, that is within the sentence, not at its beginning. Within the first three "spaces" in the sentences, a filler occurred only in 14 instances out of the total of 125, which accounts for 11.2%. Notably, the overwhelming majority of fillers in the beginning position were produced by the interpreter of Speech 3. In Speech 6, only one instance of the beginning position was recorded, and no fillers appeared at the beginning of sentences in Speech 9. 111 instances were recorded to appear in the middle position; in the summary of their findings, Boonsuk et al. (2019, 139) similarly conclude that "considering they [the speakers] think and speak at the same time, there is bound to be an increase in the use of hesitation markers in the middle of their sentences."

As I already mentioned in the overview section, the female interpreter of Speech 3 gave the biggest impression of being "unsure" of herself, employing significant amounts of hesitations and also using lots of inbreaths in unusual places. This "uncertainty" is reflected in her usage of fillers in the beginning position as can be seen in the table (33).

4.3.2 Associated factors

To answer the question of what the production of fillers might be influenced by, I once again inspected each filler, this time comparing the interpretations with the original speeches in order to see how the information was related into the target language. I listened to the original speaker through ELAN, but also used the transcriptions on the website of the European Parliament. After this inspection, four categories were discovered to be associated with the production of fillers.

	Temporal constraints	Lexical retrieval	Structural changes	Co-occurrence	Not identified
Speech 3	12	20	20	26	12
Speech 6	4	3	9	5	1
Speech 9	0	3	3	7	0
Total: 125	16	26	32	38	13

(34) Description of associated factors with the production of fillers, in numbers.

The most represented category was a co-occurrence with 38 instances. The least number of fillers, perhaps surprisingly, happened because of temporal constraints, in 16 cases. No apparent factor associated with the production of fillers could be uncovered with 13 fillers. In some cases, the factors overlapped and more than one influence on the production could be observed. Let me now explain what is meant by the individual factors and how they are associated with the fillers, along with examples.

4.3.2.1 Temporal constraints

In simultaneous interpreting, time is of the essence. Interpreters have a limited span in which they have to listen to a speech in the source language, decode it and render it in the target language. As described in chapter 2.2.2, interpreters can to a certain degree work with this time span by either lengthening the delay between them and the original speaker, or shortening it, speaking almost in perfect simultaneity, if the nature of the source speech allows it. Similarly, when a speaker talks fast and uses many terms, the interpreter tries to keep up with the pace as to not lose any vital information, and when a speaker talks slow, the interpreter has to wait for more information to be spoken in order to grasp the overall meaning.

Temporal reasons such as these mentioned were one of the factors that were uncovered in the analysis. I accounted not only for the hurried (or slow) pace of the original speaker and therefore fast (or slow) delivery by the interpreter, but also for waiting for information that appears later in the speech, and which would be more appropriate to say first, according to the thematic and rhematic elements in Czech sentence structure. I provide two examples of such instances:

I also welcome that the United States authorities have shown recognition of the need to address in their proposals similar assets held by some non-US financial institutions.	(…) příslušné orgány ve spojených státech @(inbr) se @eh snaží @eh také teď @eh: se vyznat v aktivech které drží jiné než americké @(inbr) instituce (…)[14]

(35)   Example of fillers associated with temporal constraints,
taken from Speech 3. Left: original speech, right: interpretation.

Example (35) is of an instance where the delivery of the original speaker is fast and the interpreter has to keep up. With that, certain omissions are tied, but the primary associated factor in this case was deemed to be temporal constraints. The interpreter produces three fillers in three places and also noticeable inbreaths which could be also accounted for the rapid delivery.

The closing date for actually making= making our submission has actually passed by a few days.	(…) my jsme předložili náš návrh @eh: @(inbr) @_ v podstatě až pozdě několik dní po termínu (…)[15]

(36)   Example of a filler occurring due to waiting for additional
information, taken from Speech 6. Left: original speech, right: interpretation.

In example (36), the interpreter is waiting for additional information to be said before he commits to an interpretation. First, he mentions the submission (*předložili náš návrh*) which is a topic of the debate discernible from the previous context, so this piece of

---

[14] Back translation: "(…) the relevant authorities in the United States @(inbr) are @eh now also trying @eh: to get to grips with assets held by non-US @(inbr) institutions (…)"
[15] Back translation: "(…) we submitted our proposal @eh: @(inbr) @_ basically late a few days after the deadline (…)"

information is known. Then he halts, which is evident from the production of the prolonged filler @eh:, also an inbreath and an unfilled pause @_, during which he is listening. After that, he continues with the information that the due date for the submission has passed (*až pozdě několik dní po termínu*) which is the new, so far not mentioned, piece of information. Therefore, the production of this filler is marked to the temporal constraints, waiting for a sentence to continue in order to ensure clear interpretation.

4.3.2.2 Lexical retrieval

The second factor that influenced the production of fillers was discovered to be lexical retrieval. Within this thesis, this is perceived to happen in the cases of recalling difficult terminology, names of institutions or individuals, set abbreviations, various idiomatic expressions or overall difficult constructions. The workload on the interpreter is high in these moments when they have to remember or recall larger amount of information. The most amount of fillers tied to lexical retrieval happened when names and abbreviations were mentioned, since they usually have a set equivalent in Czech and the interpreter has to recall (or quickly search for) the right expression. The name of an institution can be seen in example (37):

The Committee of European Securities Regulators will play a strong coordination role (…)	(…) no a rovněž je to evropský= evropský výbor **@eh:** orgánu pro dohled který bude hrát velice důležitou roli (…)[16]

(37)   Example of fillers associated with terminology,
taken from Speech 9. Left: original speech, right: interpretation.

Here, the interpreter produces a repetition of *evropský* and a filler @eh: before *orgánu*, thinking about the correct equivalent in Czech.

Not only names were considered to be a factor influencing the production of a filler, but also, as I mentioned, difficult or unusual expressions, as demonstrated by this example:

---

[16] Back translation: "(...) so and it's also the European=European Committee @eh: of the body for supervision which will play a very important role (...)" Note that the name of the Committee in this back translation is purposely nonsensical.

50

| We need to continue to work closely with other regulatory authorities and to the extent possible **dovetail** our responses. | (…) a pokud to je vůbec možné @(inbr) musíme **@eh @eh** o- **@eh** seřizovat dohromady naši odezvu (…)[17] |

(38)   Example of fillers associated with an idiomatic expression, taken from Speech 3. Left: original speech, right: interpretation.

In example (38), the original speaker uses the verb *dovetail*, which can be defined as "to fit together well, or to cause something to fit together well with something else."[18] In this particular context, the verb's meaning can be taken as "to cooperate" or "to manage the response together." The interpreter was likely not expecting this verb and struggled with the idiomatic meaning expressed by it. In Czech, the idiomaticity had to be left behind, and instead, the interpreter arrived at *seřizovat dohromady* (*to coordinate, to work together on*). Up to three fillers @eh were produced, also with an unfinished word evident by the scraped *o-*. The fillers were therefore accounted for the difficult idiomatic expression, thus falling into the category of lexical retrieval.

4.3.2.3   *Structural changes*

The nature of the interpreting profession, that is to ensure communication between two languages, naturally includes some changes between sentences in the source language and their transformations in the target language. As discussed in chapter 2.2.3, interpreters can use different tactics and strategies for reformulations, omissions, additions or specifications. Some of these tactics were discovered to be tied to the production of fillers within my analysis sample, and I would like to provide two examples of such changes in the target deliveries.

---

[17] Back translation: "(…) and if it is even possible @(inbr), we have to @eh @eh o- @eh coordinate our response (…)"

[18] Definition from Cambridge Dictionary, https://dictionary.cambridge.org/dictionary/english/dovetail.

| One observer referred some months ago to this unfolding crisis as like watching a train crash in slow motion. | (…) / před **@eh** měsícem jsem vlastně hovořil o tom že to je jako kdyby letadlo pomalu padalo (…)[19] |

(39)   Example of a filler tied to omission and reformulation, taken from Speech 3. Left: original speech, right: interpretation.

In example **Chyba! Nenalezen zdroj odkazů.**, rather significant omission and reformulation happen concurrently. The interpreter replaces an outside observer with the actual speaker (*před měsícem jsem – já*), and also reduces *some months ago* to only one month (*před měsícem*). After that, the mention of *this unfolding crisis* is dropped, instead only *this* is used (*že to je*), but I think this is acceptable because the fact that the speaker is talking about a crisis is known from the preceding context. At last, instead of a hurrying train, the interpreter uses a metaphor with a falling plane (*jako kdyby letadlo pomalu padalo*). Considering all these factors, the filler @eh that occurs at the beginning of the sentence is therefore accounted for as tied to the interpreter's changes within the interpretation.

Let me show a second example of a little different case of a structural change:

| Madam President, at my stage of life, I do not get very surprised too often, so I am not the least bit surprised at what goes on here in the European Parliament and the views of some people. | (…) děkuji paní předsedkyně / mě: @eh takováto reakce příliš nepřekvapuje / já jsem v evropském parlamentu a @eh není to poprvé (…)[20] |

(40)   Reformulations and additions tied to the production of fillers, taken from Speech 6. Left: original speech, right: interpretation.

Significant omissions can be seen in example (40) – *I am not the least bit surprised at what goes on here in the European Parliament* is expressed only by *mě @eh takováto*

---

[19] Back translation: "(…) a month @eh ago, I basically talked about that it's as if a plane was slowly falling (…)"

[20] Back translation: "(…) thank you, Madam President. I @eh am not too much surprised by such a reaction. I have been in the European Parliament and @eh it is not the first time (…)"

reakce příliš nepřekvapuje; no mention of the Parliament or what is happening within it is rendered in the interpretation. Or with *views of some people*, the interpreter does not talk about any *názory některých lidí*. That does not mean the interpretation is wrong, it was simply a choice of the interpreter to omit these pieces of information and state them differently.

But what I want to turn the attention to in this example are additions which are also present - *takováto reakce* (*this reaction*) and *není to poprvé* (*it is not the first time*). The surprise is still kept in the target delivery (*příliš nepřekvapuje*), but these two pieces of information are added in order to simplify and also specify the transmitted message. The two fillers @eh, each appearing before one addition, are therefore attributed to this structural change made by the interpreter.

4.3.2.4 Co-occurrence

Many times, fillers occurred together with other types of hesitation, and this factor was the one that was identified most often, 38 times. The prevalent associated hesitation was a self-correction; the interpreter identified a mistake in the previously said segment, signalled this with the use of a filler, and then added the corrected information right after. Let me illustrate this with an example from Speech 9:

| The proposal will aim to introduce a legally binding authorisation and robust external oversight regime, whereby European regulators will have to supervise the policies and procedures followed by the credit-rating agencies. | (…) abychom měli určitý externí systém dohledu tak aby **eur-** evropské **dohled- @m** dozorčí orgá:ny měly možnost dohledu nad těmito ratingovými společnostmi (…)[21] |

(41)   Example of a filler associated with a self-correction,

taken from Speech 9. Left: original speech, right: interpretation.

---

[21] Back translation: "(…) so that we have a certain external system of supervision so that the Eur- European superv- @m supervisory bodies could have the option of supervision over these credit-rating agencies (…)"

In this example, two self-corrections are visible before the filler is produced: *eur-* and *dohled-*. The interpreter then corrects themselves with the right term *dozorčí orgány*. Before the correction, the interpreter produces a "halting" filler @m.

Another hesitation device that occurred together with a filler was a repetition:

I will be bringing to the college next week amendments to the deposit guarantee scheme which will increase the minimum level of protection, as well as requiring Member States to put in place procedures for rapid pay-out.	(…) budeme mluvit o: depozitních garancích které by **snížily** ty **@ehm jaksi snížily** jednotlivé požadavky v této oblasti tak aby jednotlivé členské státy mohly velice rychle začít vyplácet (…)[22]

(42)   Example of a filler associated with a repetition, taken from Speech 9. Left: original speech, right: interpretation.

In this example, the word *snížily* is said, after that a short *ty*, then a filler @ehm is produced, during which the interpreter listens to the speech and ponders for a moment as to how to continue. After the filler, a filler word *jaksi* is said, and then a repetition of *snížily* occurs. Two types of hesitations can therefore be seen here: repetition and filler word. That is why the filler is taken to be associated with a co-occurrence of other HP.

4.4 Summary of the Analysis

The analysis of hesitation devices in the eleven interpretations was not an easy task and took a long time due to the quantity. I approached this investigation with a theory that fillers would be the most common phenomenon occurring in the interpretations because of their function of signalling delay and also a relatively easy production. This was confirmed by the results as fillers occurred 702 times, out of the total of 1230 hesitation devices. This was the reason fillers were chosen for further analyses of placement within a sentence and associated factors with their production.

---

[22] Back translation: "(…) we will be talking about deposit guarantees which would in a way lower the @ehm in a way lower the individual requirements in this sector so that individual member states could very quickly start to pay out (…)" In the place of the Czech filler word *jaksi*, the English *in a way* is used.

Further, due to the large amount, the analysis was narrowed to three chosen interpretations of speeches that were conducted by the same original speaker, with the number of fillers reduced to 125. Before investigating the place of occurrence, my theory was that a filler would more likely occur at the beginning of a sentence due to the fact that the interpreter has to decide with which words they are going to commit to the interpretation. This theory was not proven to be true in the sample of three chosen speeches, and instead, the overwhelming majority of fillers appeared in the middle position, that is later on within a sentence.

The next step was comparing the interpretations with the original speeches and looking at the surrounding speech environment in order to try uncover possible factors associated with the production of the fillers. Four categories were discovered: temporal constraints, that is if the original speaker was too fast, or if vital information came later in the sentence; lexical retrieval, that is difficult terminology, names or abbreviations; structural changes done by the interpreter, for example reformulations, omissions and additions; and co-occurrence with other types of hesitation phenomena, which was a factor that occurred the most, 38 times.

As a last comment, it is important to say that the results of the placements and the associated factors are strictly tied to the three chosen speeches, and they might not hold true for other speeches out of the total eleven that were investigated. Due to the quantity of the data, but simultaneously the limited scope of this bachelor's thesis, the research objectives needed to be adequately narrowed. As such, there is room left for future investigations on this topic.

Conclusion

The goal of this bachelor's thesis was to introduce hesitations in speech, often referred to as hesitation phenomena, and look at them in the context of simultaneous interpreting.

In the theoretical part, hesitations are described as a natural part of speech that on one hand interrupt the continuous flow, but on the other hand give clues as to what might be coming later in the segment or that the speaker is momentarily preoccupied with the speech planning process. Categories of hesitation phenomena are listed, according to the typology summarized by Rose (1998). His examples are used in the descriptions of the types as well as examples from a sample of interpretations from the European Parliament. Different descriptions are mentioned and points of view of various authors are provided. A chapter on the function of hesitations in discourse is also provided.

The second part of the theoretical background is dedicated to simultaneous interpreting. In comparison to spontaneous speech, the time span in which interpreters have to transfer the information is limited. In order to uncover the internal processes of the interpreters' minds, Gile's Efforts Model is introduced and his coping tactics and strategies are described, some of which might tie into the production of hesitations. Past research on hesitations in simultaneous interpreting is also mentioned.

The practical part is an analysis of hesitation markers across eleven interpretations which were scrutinized through the ELAN software. The categories which were described in theory (minus restarts and filler words) are observed in the speeches and statistical overview is given. The first hypothesis that fillers would be the most commonly occurring hesitation is proven by the data. The scope is then narrowed to the filled pauses with fillers, and subsequent analysis of three speeches shows that they occur more likely within a sentence rather than at its beginning, disproving a previously stated hypothesis that they would occur at the start. Factors associated with the production of fillers are also scrutinized and four are discovered, with the factor occurring most often being co-occurrence with other hesitations.

In conclusion, some trends in the placement of fillers and factors were uncovered, suggesting that they might not be randomly dispersed, and that speakers do use them as a means to signal speech planning problems. The available dataset also provides different areas worth researching, for example discourse markers which were only briefly mentioned when talking about filler words, or a cross-linguistic analysis of the usage of different types of hesitation devices.

Resumé

Tato bakalářská práce se zabývá hezitacemi v řeči tlumočníků z angličtiny do češtiny. Jejím cílem je stručně popsat fenomén hezitace v mluvené spontánní řeči, představit různé způsoby, jakými řečník může hezitovat (či váhat), popsat možné funkce těchto výrazů v diskurzu a zohlednit i faktory simultánního tlumočení v produkci těchto zvuků v rámci blíže vymezeného tématu práce.

První kapitola stručně definuje pojem hezitace jakožto neočekávané přerušení v plynulé řeči. Ideálně by komunikace měla probíhat nepřerušeně a bez jakýchkoliv zaváhání, což často slyšíme v médiích a u řečnických profesí, mezi které patří i tlumočníci, ale realita bývá jiná. Mluvčí přirozeně dělají pauzy ve svých promluvách, prodlužují slova na jejich koncích, začnou promluvu, ale pak ji nedokončí, opakují některé segmenty či značí zaváhání různými zvuky nebo výplňkovými slovy. Tato problematika je relevantní pro několik vědeckých oblastí, jako například medicína, kdy jsou zkoumáni pacienti s afázií či řečovými vadami, psychologie a psycholingvistika nebo vývoj automatizovaných softwarů pro rozpoznávání řeči.

V rámci první kapitoly jsou hezitace rozděleny do několika skupin. Zde je použita kategorizace podle Rosea (1998). Celkově je popsáno sedm kategorií: falešné začátky, opakování, restarty, autokorekce, koncové prodloužení, pauzy a výplňková slova. Pauzy jsou nadále rozděleny do dvou podkategorií, a to vyplněné pauzy a nevyplněné (tiché) pauzy. Každá kategorie je stručně definována, zohledněny jsou úhly pohledů několika různých autorů a přiloženy jsou také příklady jednotlivých hezitací v promluvách, které jsou převzaty z Roseovy práce (1998), ale také jsou použity příklady z vlastního souboru dat tlumočnických promluv.

V třetí části první kapitoly je stručně pohovořeno o možných funkcích hezitací v diskurzu. Hezitace jsou používány k tomu, aby řečník dal najevo, že se musí vypořádat s nějakým problémem, jež souvisí s plánováním nadcházející řeči či s opravením řeči předcházející. Zároveň řečník nechce svou promluvu ukončit a bude pokračovat v rámci několika málo chvil. Hezitace mohou taktéž posluchače upozornit na nadcházející prodlevu, která může podle Clarka a Fox Treeové (2002) být krátká, potom je použit výplňkový zvuk (filler) *uh*, anebo dlouhá za použití *um*. V různých jazycích se mohou tyto zvuky lišit, českými ekvivalenty jsou pak *eh* a *ehm*. Clark a Fox Treeová (2002) dokonce považují výplňkové zvuky *uh* a *um* za samostatná slova, která podléhají anglické

fonologii, syntaxi, sémantice a pragmatice. Převažující názor je však ten, že výplňkové zvuky jsou spíše „paralingvistickými" elementy, jež je často obtížné definovat.

Druhá kapitola je věnována simultánnímu tlumočení. Nejvýraznějším znakem tohoto druhu tlumočení je souběžnost několika procesů – tlumočník musí zároveň poslouchat, dekódovat zprávu, pochopit ji a přeformulovat do cílového jazyka. Také musí být vybaven širokou škálou vědomostí a slovní zásoby, od kulturních souvislostí pracovních jazyků až po úzce vymezené téma dané konference. Pro nastínění těchto vnitřních procesů je v rámci této kapitoly popsán Model úsilí, jehož autorem je Daniel Gile (1995). Tento model říká, že tlumočníci musí rozdělit svou mentální kapacitu mezi aktivní poslouchání a analýzu, ukládaní informací do krátkodobé paměti, produkování cílové promluvy a vše musí být zkoordinováno a vyváženo. Když nedojde k adekvátnímu rozdělení mentální kapacity mezi tyto procesy, může dle Gila (1995) dojít k saturaci – tlumočník je přehlcen a to má negativní dopad na jeho tlumočnický výkon. Dále Gile představuje taktiky a strategie, které tlumočníkům mohou napomoci dosáhnout nejlepšího možného transferu. Zde je zmíněna anticipace neboli předvídání, manipulace s časovým posunem a různé způsoby reformulace. Krátce je také zmíněn předešlý výzkum hezitací v rámci simultánního tlumočení – tato oblast výzkumu je stále poměrně málo probádaná, protože se většina autorů soustředí na hezitace ve spontánní mluvené řeči v běžných každodenních situacích.

V praktické části byla provedena analýza jedenácti promluv tlumočníků z Evropského parlamentu s důrazem na jejich užívání hezitací. Na transkripci těchto prvků byl použit software s názvem ELAN. Nejprve bylo potřeba hezitační zvuky v promluvách ručně vyznačit, protože automatická transkripce tyto prvky nijak nezapisuje. Poté byly hezitace statisticky zhodnoceny. V užší části analýzy byly dále prozkoumány pouze vyplněné pauzy s výplňkovými zvuky (fillers) v rámci tří zvolených tlumočnických promluv. Zvláštní pozornost byla věnována umístění výplňkových zvuků v rámci věty, zda-li se vyskytují spíše na začátku nebo na pozdějších pozicích uvnitř věty, a faktory související s produkcí těchto zvuků. Zde byly vytyčeny čtyři kategorie, které ovlivnily produkci vyplněných pauz: časová omezení (příliš rychlé či pomalé tempo), vybavení si lexikálních elementů (nejčastěji terminologie a určitých názvů) strukturální změny ze strany tlumočníků (vynechávky, reformulace či doplnění výpovědi) a společný výskyt s ostatními typy hezitací (nejčastěji s autokorekcí). Na základě výsledků dat bylo zjištěno, že výplňkové zvuky se vyskytují spíše v průběhů vět, protože řečník zároveň

mluví a přemýšlí nad dalšími segmenty, a nejčastějším přidruženým faktorem byl výskyt společně s ostatními hezitacemi.

Spojením teoretické a praktické části poskytuje práce stručný vhled do problematiky hezitací v promluvách a v rámci blíže specifikovaného tématu se věnuje výskytu hezitací v kontextu simultánního tlumočení. Využité tlumočnické promluvy jsou velmi obsáhlé a poskytují i další možnosti pro různé druhy analýz, které by mohly být v budoucnu probádány, jako například umístění jiných druhů hezitací, ne pouze vyplněných pauz, nebo větší soustředění se na výplňková slova či mezijazykové porovnání co se týče užívání různých hezitačních prostředků.

Annotation

Autor práce: Karolína Řeháková

Studijní obor: Angličtina se zaměřením na komunitní tlumočení a překlad

Fakulta: Filozofická fakulta

Katedra: Katedra anglistiky a amerikanistiky

Název bakalářské práce: Hesitation Markers and Fillers in the Speech of Interpreters

Název česky: Hezitační zvuky a výplňková slova v řeči tlumočníků

Vedoucí práce: Mgr. Markéta Janebová, Ph.D.

Rok obhajoby: 2022

Počet znaků: 88 975

Počet stran: 56

Anotace: Cílem této bakalářské práce je stručně představit fenomén hezitace neboli zaváhání v řeči a blíže se věnuje výskytu těchto prvků v promluvách simultánních tlumočníku v Evropském parlamentu. Představeny jsou různé typy hezitací a jejich možné funkce v diskurzu. Specifika simultánního tlumočení jsou taktéž zmíněna. V rámci praktické části je provedena analýza těchto prvků v tlumočnických promluvách se zaměřením na místo výskytu těchto elementů ve větě a přidružených faktorů, které ovlivňují jejich produkci.

Klíčová slova: hezitace, řečová dysfluence, simultánní tlumočení, Evropský parlament

Annotation: The goal of this bachelor's thesis is to introduce the hesitation phenomena in speech and it pays closer attention to the occurrence of these elements in the speeches of simultaneous interpreters in the European Parliament. Different types of hesitations are introduced and their possible discourse functions are presented. Specifics of simultaneous interpreting are also mentioned. In the practical part, an analysis of hesitation devices in interpretations is conducted with focus on their place of occurrence and associated factors of their production.

Key words: hesitation, speech disfluency, simultaneous interpreting, European Parliament

Works cited

Betz, Simon, and Petra Wagner. 2016. "Disfluent Lenghtening in Spontaneous Speech." In *Elektronische Sprachsignalverarbeitung*.

Blackmer, Elizabeth R., and Janet L. Mitton. 1991. "Theories of monitoring and the timing of repairs in spontaneous speech." *Cognition* 39: 173-194.

Boomer, Donald S. 1965. "Hesitation and Grammatical Encoding." *Language and Speech* 8(3): 148-158.

Boonsuk, Yusop, Eric A. Ambele, and Chamaiporn Buddharat. 2019. "Position of Hesitation Markers in Everyday, Informal Conversation in English." *ABAC Journal* 39: 129-140. Accessed 29th May, 2021.

Boughaba, H. 2021. "Speech Disfluencies in Simultaneous Interpretations of Spontaneous and Non-Spontaneous Speeches." *International Journal of Linguistics, Literature and Translation* 4(7): 15-21.

Brennan, S. E., and M. Williams. 1995. "The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers." *Journal of Memory and Language* 34: 383-398.

Clark, Herbert H. 1994. "Managing problems in speaking." *Speech Communication* 15: 243-250.

Clark, Herbert H. 2006. Pauses and Hesitations: Psycholinguistic Approach. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*. Elsevier. 244-8.

Clark, Herbert H., and J. E. Fox Tree. 2002. "Using uh and um in spontaneous speaking." *Cognition* 84: 73-111.

Clark, Herbert H., and Thomas Wasow. 1998. "Repeating Words in Spontaneous Speech." *Cognitive Psychology* 37: 201-242.

Dechert, Hans W, and Manfred Raupach. 2011. *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler.* Berlin, New York: De Gruyer Mouton.

Fox, B., F. Wouk, M. Hayashi, S. Fincke, L. Tao, M. L. Sorjonen, M. Laakso and W. F. Hernandez. 2009. "A cross-linguistic investigation of the site of initiation in same-turn self-repair. In Conversation Analysis: Comparative Perspectives, 60-103. Cambridge University Press.

Fox Tree, J. E. 1995. "The Effect of False Starts and Repetitions on the Processing of Subsequent Words in Spontaneous Speech." *Journal of Memory and Language* 34: 709-738.

Fox Tree, J. E, and Herbert H. Clark. 1997. "Pronouncing "the" as "thee" to signal problems in speaking." *Cognition* 62: 151-167.

Fox Tree, J. E. 1999. "Discourse Markers in Spontaneous Speech: Oh What a Difference an Oh Makes." *Journal of Memory and Language* 40: 280-295.

Fox Tree, J. E. 1999. "Listening in on monologues and dialogues." *Discourse Processes* 27(1): 35-53.

Fox Tree, J. E. 2001. "Listeners' uses of *um* and *uh* in speech comprehension." *Memory & Cognition* 29: 320–326.

Fox Tree, J. E. 2002. "Interpreting Pauses and Ums at Turn Exchanges." *Discourse Processes* 34(1): 37-55.

Gile, Daniel. 1995. *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam: John Benjamins Pub. Co.

Gile, Daniel. 2009. *Basic Concepts and Models for Interpreter and Translator Training. Revised Edition.* Amsterdam: John Benjamins Pub. Co.

Gilquin, Gaëtanelle. 2008. "Hesitation markers among EFL learners: pragmatic deficiency or difference?" *Pragmatics and Corpus Linguistics: A Mutualistic Entente* 119-150. Berlin, New York: De Gruyter Mounton.

Goldman-Eisler, Frieda. 1958. "Speech Production and the Predictability of Words in Context." *Quarterly Journal of Experimental Psychology* 10, no. 2: 96-106.

Goldman-Eisler, Frieda. 1961. "The Distribution of Pause Durations in Speech." *Language and Speech* 4(4): 232-237.

Goldman-Eisler, Frieda. 1968. *Psycholinguistics: experiments in spontaneous speech.* London: Academic Press.

Gósy, Mária. 2007. "Disfluencies and Self-Monitoring." *Govor* 24: 91-110.

Hieke, Adolf E. 1981. "A Content-Processing View of Hesitation Phenomena." *Language and Speech* 24, no 2.: 147-160.

Kosmala, Loulou, and Aliyah Morgenstern. 2018. "Should 'uh' and 'um' be categorized as markers of disfluency? The use of fillers in a challenging conversational context." *Fluency and Disfluency across Languages and Language Varieties.*

Lambert, Sylvie. 1992. "Shadowing." *Meta* 37(2): 263-273.

Lederer, M. 1978. "Simultaneous Interpretation. Units of Meaning and other Features." In *Language Interpretation and Communication,* edited by D. Gerver and H. W. Sinaiko, 323-332. New York: Plenum.

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

Lickley, R. J. 2015. "Fluency and Disfluency." In *The Handbook of Speech Production,* M.A. Redford (Ed.).

Maclay, Howard S., and Charles E. Osgood. 1959. "Hesitation Phenomena in Spontaneous English Speech." *WORD* 15: 19-44.

Mahl, George F. 1956. "Disturbances and Silences in the Patient's Speech in Psychotherapy." *The Journal of Abnormal and Social Psychology* 53(1): 1-15.

Mikuláštík, Milan. 2003. *Komunikační dovednosti v praxi.* Grada Publishing a.s.

O'Connell, Daniel C., and Sabine Kowal. 2005. "*Uh* and *Um* Revisited: Are They Interjections for Signaling Delay?" *Journal of Psycholinguistic Research* 34(6): 555-576.

O'Shaughnessy, Douglas. 1993. "Analysis and Automatic Recognition of False Start in Spontaneous Speech." In *Proceedings of the 1993 IEEE international conference on Acoustics, Speech, and Signal Processing: Speech Processing - Volume II (ICASSP'93).* IEEE Computer Society, USA, 724–727.

Oléron, P., and H. Nanpon. 1965. "Research into simultaneous translation." In *The interpreting studies reader,* edited by F. Pöchhacker and M. Shlesinger, 69-76. London: Routledge.

Quirting, Gerd. *Hesitation phenomena in adult aphasic and normal speech.* Berlin, Boston: De Gruyter Mounton, 2019.

Rose, R. L. 1998. "The Communicative Value of Filled Pauses in Spontaneous Speech." Master's thesis. University of Birmingham.

Shriberg, Elizabeth Ellen. 1994. "Preliminaries to a Theory of Speech Disfluencies." PhD thesis. Berkeley, University of California.

Tissi, Benedetta. 2000. "Silent Pauses and Disfluencies in Simultaneous Interpretation: A Descriptive Analysis." *The Interpreters' Newsletter* 10.4: 103-127.

Volín, Jan, Lenka Weingartová, and Radek Skarnitzl. 2013. "Spectral Characteristics of Schwa in Czech Accented English." *Research in Language* 11(1), 31-39.

Volín, Jan. 2016. "Kolik řeči je vlastně v řeči." *Naše řeč* 1: 52-55.

Wiese, R. 1984. "Language Production in Foreign and Native Languages: Same or Different?" In *Second Language Productions,* edited by H. W. Dechert, D. Möhle, & M. Raupach, 11-25. Gunter Narr Verlag.

Used software:

Max Planck Institute for Psycholinguistics. 2020. *ELAN* (version 6.0). Nijmegen: The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

Inspected speeches are available on the enclosed CD ROM.