Univerzita Karlova

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Bakalářská práce

Filip Drábek

Corpus-based comparative analysis of filled pauses in teacher English of native and L2 speakers

Komparativní analýza vyplněných pauz v učitelské angličtině rodilých a nerodilých mluvčích na základě korpusových dat

Praha 2024

Vedoucí práce: PhDr. Tomáš Gráf, Ph.D.

Poděkování

Především děkuju doktoru Gráfovi za nesmírnou ochotu, laskavost, smysl pro humor a cenné rady během psaní práce – a vůbec během celého studia. Také děkuju všem hudebníkům, které jsem během psaní práce poslouchal, za to, že mi tento proces zpříjemňovali.

Prohlášení

Prohlašuju, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

Souhlasím se zapůjčením bakalářské práce ke studijním účelům.

V Praze dne _____

Filip Drábek

Declaration

I hereby declare that this thesis is my own work, that I have properly referenced and cited all used resources and literature, and that this thesis has not been used elsewhere in my university studies, nor was it used to acquire another or the same diploma.

I have no objections to the BA being borrowed and used for study purposes.

In Prague, on the _____

Filip Drábek

Abstract

This thesis investigates occurrences of filled pauses in teacher English and provides their typology along with quantitative data: frequency and the total number of the individual types of pauses. The analysis, typologisation and statistical output are performed individually for for native and non-native speakers, the characteristics of filled pauses for each group are compared at the end. All data are taken from the ETC corpus of teacher English. The thesis is written in English.

Key words: disfluency, filled pauses, teacher English, teacher English corpus, corpus analysis, native speakers, non-native speakers

Abstrakt

Tato práce zkoumá výskyty vyplněných pauz v učitelské angličtině a poskytuje jejich typologii spolu s kvantitativními daty: frekvencí a absolutní četností jednotlivých typů pauz. Analýza, typologizace a statistický výstup jsou vypracovány zvlášť pro rodilé a nerodilé mluvčí, charakteristiky pauz obou skupin jsou na závěr srovnány. Data jsou čerpána z korpusu učitelské angličtiny ETC. Práce je vypracována v angličtině.

Klíčová slova: disfluence, vyplněné pauzy, učitelská angličtina, korpus učitelské angličtiny, korpusová analýza, rodilí mluvčí, nerodilí mluvčí

Table of Contents

List of abbreviations

adj – adjective

adv – adverb

aux – auxiliary verb

caus – causative verb

CDM – cognitive disfluency marker

CiSi – clause-initial, sentence-initial

CiSm – clause-initial, sentence-medial

CmSm – clause-medial, sentence-medial

conj – conjunction

cop – copular verb

copinf – copular verb in the infinitive form

dadj – adjective with the definite article following a filled pause

det – determiner

DM – discourse marker

dn – noun with the definite article following a filled pause

ELT – English language teaching

ETC – English Teacher Corpus

FP – filled pause

iadj – adjective with the indefinite article following a filled pause

iadv – adverb with the indefinite article following a filled pause

L1 – native speaker

L2 – non-native speaker

lex – lexical verb

lexinf – lexical verb in the infinitive form

mod – modal verb

n – noun

NF – non-finite/verbless construction

not – negative particle “not”

num – numeral

prep – preposition

pron – pronoun

q – quantifier

SDM – semantic disfluency marker

List of tables

1. Introduction

Filled pauses (henceforth FPs) are among the more prominent features specific to unrehearsed, everyday speech. While they have been and are still sometimes viewed negatively – as evident in various rather expressive nicknames, such as "parasite words" – they have also been shown to play an important role in cognitive and pragmatic linguistic processes. Approaching them as functional (i.e., strategic, not pathological) elements, researchers have identified a number of specific functions which FPs serve. Some of the well-defined and accepted functions include maintaining one's speaking turn, buying time to retrieve a lexical item, or serving as a means of back-channeling, though others have been proposed.

Researchers in the 20th century and onwards generally agree on classifying FPs as one of the several types of disfluencies. Their aforementioned functional nature, however, has led to multiple different descriptions, ranging from treating them as mostly cognition-driven phenomena to actual lexemes, contributing information to the discourse. While there is currently no consensus on their status within the linguistic system of English, their uniquely meaningful nature – especially in comparison to most other disfluencies – is generally accepted.

This thesis views FPs accordingly and aims to describe their usage and distribution in teacher English as functional, natural phenomena. Although FPs have been fairly extensively researched in L1 speakers of various languages, L2 speaker usage in most languages is comparatively much less clear. Furthermore, with the growing interest in the language of language teachers, researching their usage of FPs can provide broader insight on teacher-specific linguistics patterns and paradigms. Understanding teacher language is essential for understanding the way these patterns and paradigms are transmitted to learners in the classroom.

Due to their capacity for communicating semantic/pragmatic meanings, FPs are perhaps the most ideal disfluency for the purpose of comparing disfluency usage in L1 and L2 speakers of English. On the basis of data extracted from the brand new English Teacher Corpus (ETC), this thesis aims to identify potential differences in FP usage for L1 and L2 English speakers respectively, describe them, and finally compare them both quantitatively and qualitatively.

2. Theoretical background

2.1 A general overview of (dis)fluency in language

FPs have traditionally been classified as one of the main types of linguistic disfluencies, along with repetitions, silent pauses, prolongations, self-corrections, and false starts (Williams, 2022, p. 1). Disfluencies are perhaps best defined in contrast to fluency: for example, de Jong (2016) views fluency (in L2 specifically) as "being able to smoothly and effortlessly translate intended messages to speech" (de Jong, 2016, p. 113). Williams (2022) also notes that "fluency has become synonymous with speed, and disfluency with forms of hesitation and repair" (Williams, 2022, p. 1), though Bergmann et al. (2015) warn that fluency should not be equated with the speed of speech production (Bergmann et al., 2015, p. 26). Rather, they stress another equally important factor – performance aspects, and thus define fluency "as stretches of speech in which disfluency markers[1] do not exceed a certain frequency" (Bergmann et al., 2015, p. 26).

Thus disfluencies can be rather broadly defined as linguistic phenomena which seemingly "interrupt" the flow of speech production, although their occurrence is triggered by various cognitive or linguistic processes, especially those facilitating speech production. The more contentious question is that of their meaning, function, and status in the linguistic system. Some of them have been described both as word-like units (Schmid and Fägersten, 2010, p. 755), and as "in most instances involuntary reflections of cognitive processes without any signalling quality" (Bergmann et al., 2015, p. 26). A reasonable middleground could be to classify the various disfluencies according to their respective properties, recognising their heterogeneous nature. To give an example, arguing for the word-status of unfilled pauses in spoken language is problematic, as they inherently have no phonological form, whereas other disfluencies may be more promising candidates.

Given these functional, or even word-like properties, it has already been proposed that disfluencies actually go "hand in hand" with fluent speech, as opposed to the older view – exemplified by the name *disfluency*[2] itself – that they disrupt speech. The next chapter addresses the recent shift in perspective on disfluency, whose direction this thesis follows.

[1] Disfluency markers = disfluencies
[2] Formerly *dysfluency*

2.1.1 Disfluencies: Friend or foe?

Outside the academic sphere, disfluencies are often viewed as detrimental to "proper" speech – how does the academic view compare? For example Kosmala (2021) distinguishes between *disfluency* and *dysfluency*, reserving the latter for speech pathology, such as stuttering (Kosmala, 2021, p. 71). Similarly, Rose (1998) distinguishes between disfluency which is pathological, and that which is normal (Rose, 1998, p. 9). Bergmann et al., (2015) note that "all spontaneous speech is charactersied by frequent occurrences of disfluencies, and these present a window into underlying planning processes" (Bergmann et al., 2015, p. 26). Williams (2022) even describes them as a counter-measure to the side-effects of fluent speech: "If fluency has its own markers, such as fast speech, long turns, accurate syntax, and syntactic complexity, it also demands a repertory of strategies for achieving them, and for dealing with their consequences: the complex syntax that loses its way, the speed that produces breakdowns in articulation, the accuracy that becomes pedantic, and the turn that never ends" (Williams, 2022, p. 2).

The formulations above acknowledge the natural occurrence of disfluencies in spontaneous speech, while hinting at possible explanations as to why this occurrence is natural. One important argument for the aforementioned "functional" or "meaningful" nature of disfluencies is that they occur not only in L2 speakers, but also in L1 speakers, albeit to a lesser extent (Williams, Korko, 2019, p. 723), even though L1 speakers are generally considered fluent (Bergmann et al., 2015, p. 25). Kosmala (2021) further notes that "while their overuse in speech may be perceived as 'disfluent' and uneducated (even in a native language), their total absence may be perceived as unnatural" (Kosmala, 2021, p. 72). This is further supported by one of the aforementioned definitions of fluency, which does not define it as devoid of disfluencies, but rather as containing disfluencies up to a certain, "moderate" frequency (Bergmann et al., 2015, p. 26).

To sum up, most contemporary research is not concerned with evaluating disfluencies as good or bad – their omnipresence in speech is widely acknowledged, and attention is instead paid to explicating reasons for this omnipresence, as well as their functions. As mentioned earlier, this approach will henceforth be followed for the purposes of this thesis – the next chapter discusses FPs specifically in respect to this approach.

2.2 FPs as one kind of disfluency

Although the descriptions of FPs generally agree on their basic formal properties, formulations of their function and nature vary. Williams (2022) defines FPs as "vocalized hesitation, including sounds as (English) eh, uh, uhm, er and the like" (Williams, 2022, p. 73). Rose (2019) defines them similarly: "Filled pauses are non-lexical vocalizations uttered by speakers which, by their occurrence, delay the transmission of the linguistic speech signal" (Rose, 2019, p. 19). What is apparent from these definitions is that FPs are actual sounds – as opposed to unfilled pauses.

Another trait being alluded to is their function as means of hesitation, or buying time. This would mean that FPs predominantly serve to facilitate cognition, or at least somehow relate to it. Schmid and Fägersten (2010) propose a different view. They distinguish between "semantic disfluency markers (SDMs)" and "cognitive disfluency markers (CDMs)" (Schmid and Fägersten, 2010, p. 756). They place FPs in particular into the first of the two categories, arguing that they "serve semantic functions linked to information structure, emphasis, or discourse organization" (Schmid and Fägersten, 2010, p. 755). These are said to also have language-specific functions (Schmid and Fägersten, 2010, p. 755). On the other hand, CDMs include unfilled pauses, false starts, and repetitions, and are characterised as facilitators of cognition, especially lexical retrieval, while not being viewed as language-specific (Schmid and Fägersten, 2010, p. 756).

It is worth noting that while these are certainly useful categories, one should not assume that all FPs in all situations function as SDMs, or that, for example, all unfilled pauses are CDMs. To exemplify the latter, unfilled pauses very often mark the end of a speaking turn, yielding the floor. Thus, it makes sense to treat the SDM – CDM opposition as a kind of a spectrum, or a scale, where e.g., false starts and repetitions can be viewed as the more "prototypical" CDMs.

Some FPs – specifically *uh* and *um* – have even been described as interjection-like words (Clark and Fox Tree, 2002, p. 79). Instead of being viewed as involuntary "byproducts" of cognition, Clark and Fox Tree (2002) suggest that they are interjections which at the moment of utterance inform the addressee/listener that a delay in speaking is coming (Clark and Fox Tree, 2002, p. 79). They even distinguish between their meanings, claiming that *uh* communicates a small delay, whereas *um* communicates a more prominent delay (Clark and

Fox Tree, 2002, p. 79). This view combines the two "competing" approaches to FPs. On the one hand, these FPs are closely related to cognition, having a specialised function of informing of an upcoming (cognitively triggered) delay in speech. On the other hand, they are still treated as semantic, or perhaps pragmatic markers – either way, they are meaningful and purposeful, not involuntary "noise". Viewing FPs as such (or not) might ultimately come down to one's perspective and answer to the question whether they mark disfluency, or constitute it. This is further evident from the fact that some authors (e.g., Bergmann et al., 2015) sometimes refer to them as "disfluency markers" instead, highlighting their communicative function.

To conclude this chapter, FPs have been described in many different, yet often overlapping ways. It is beyond the scope and interest of this thesis to identify the "best" approach. However, the (at least borderline in the more "sceptical" approaches) semantic nature of FPs is of great relevance for the comparative analysis of L1 and L2 speakers. If L1 speakers of English use FPs in a particular way – which we might call the "model" usage – how do the advanced L2 speakers compare? And do these L2 speakers, as teachers, provide a natural language model for their students? The following chapter is concerned with the findings in L1 speakers' usage of FPs, focussing especially on English L1 speakers.

2.2.1 FPs in L1 speakers

Even native speakers produce a large number of disfluencies, including FPs. Besides their consistently lower frequency in L1 spontaneous speech (Williams, Korko, 2019, p. 723), other differences have also been observed. Due to their proficiency – translating to effortless application of their L1's grammar – native speakers tend to produce entire syntactic "chunks" (de Jong, 2016, p. 114). This means that while L2 speakers need to devote more time and cognitive effort to e.g., properly inflect certain words, L1 speakers mostly encounter "problems" with discourse planning and lexical retreival or selection (which is also the case with L2 speakers in addition to grammar application). The consequence of this should be that (in L1 speech) phenomena such as FPs should tend to occur between the individual chunks within a sentence, whereas occurrence within that chunk should be far more rare. (de Jong, 2016, p. 114).

But what are those chunks exactly? De Jong (2016), analysing Dutch, suggests *clauses* while also interchangeably refering to these chunks (i.e., clauses) as *utterances* in the same study (de Jong, 2016, p. 115), thereafter using the term *ASU* (the analysis of speech unit) coined by Foster et al., (2000). They define the ASU as an utterance of one speaker "consisting of an *independent clause, or sub-clausal unit,* together with any subordinate *clause*(s) associated with either" (Foster et al., 2000, p. 365). De Jong (2016) then claims that FP occurrence within these is less typical for native speakers, and more so for non-native ones, which her study later confirms (de Jong, 2016, p. 130).

A counter-example is provided by Boonsuk, Ambele, and Buddharat (2019), who have investigated FP (among other disfluencies) usage in L1 informal English. One of the parameters investigated was the position of FPs within sentences. They found that FPs are exponentially more frequent mid-sentence, with the ratio of initial occurrence to medial to final being 14:116:8 (Boonsuk, Ambele, and Buddharat, 2019, p. 136). As their study focussed on native English, these results are difficult to reconcile with De Jong's (2016) findings discussed in the paragraph above.

Investigating and comparing French L1 and L2 FP usage, Kosmala and Crible (2022) report a strong tendency for FP occurrence in the middle of an intonation unit – a completely different approach – finding no significant difference between L1 and L2 speakers (Kosmala and Crible, 2022, p. 228). They do, however, find a difference in the clustering of FPs, where

non-native speakers cluster FPs more often, as opposed to the production of a single FP (Kosmala and Crible, 2022, p. 228).

Another, perhaps an even less understood proposed positional tendency of FPs, is that FPs occur before lower frequency lexemes to stall during the retrieval. This area is by no means well researched, although the hypothesis has successfully been tested on nouns (de Jong, 2016, p. 130). It is also generally thought that *content* words are more likely to be preceded by disfluencies (Boonsuk, Ambele, and Buddharat, 2019, p. 133). The explanation could be that *functional* (also called *grammatical*, as opposed to *content/lexical*) words are much more frequent overall, meaning retrieval should be faster, provided there indeed is a relationship between word frequency and retrieval speed. The following chapter addresses the findings in L2 FP usage in more detail.

2.2.2 FPs in L2 speakers

Since some of the parameters of L2 FP usage have already been suggested in the previous chapter, let us focus on summarising and elaborating on some of the points made there. Overall, it seems that in most positions where L1 speakers tend to produce FPs, L2 speakers do so as well, but more. As per the previous chapter, these include:

1) L2 speakers' tendency to cluster FPs, as opposed to the production of a single FP

2) L2 speakers' more frequent FP production before nouns

3) An overall higher frequency of FPs in L2 speech

The third point might of course be the consequence of the previous two, although as discussed earlier, speakers may produce FPs to buy time for processes other than lexical retrieval, e.g., to properly inflect certain words, which is more cognitively demanding for L2 speakers. If the first two points are indeed true, they can be viewed as contributing to the higher frequency observed in L2 speech, though with the caveat that other factors are most likely involved as well.

The second point also requires some further commentary. De Jong (2016) indeed shows that FPs are more likely to occur before nouns, and that this is even more likely in L2 speakers (de Jong, 2016, p. 130). However, that does not necessarily confirm the "FPs tend to occur before content words" hypothesis, hence the cautious formulation. For example, since nouns very often stand in the subject position, occurrence of FPs before such nouns could be caused by the need to buy time for morphosyntactic planning of the sentence/clause, rather than by lexical retrieval. Whether this tendency in L2 speakers exists for other content word classes as well, and moreover, whether it truly applies to nouns, as proposed here, will have to be tested.

Besides frequency, differences *might* also be found in the position of FPs relative to a particular syntactic unit. This parameter is different from the others in that it has been given more attention in research, as well as in that trying to interpret its results creates more questions than it answers[3]. As described in the previous chapter, L2 speakers have sometimes

[3] This embittered remark is mostly targetted at de Jong's (2016) inconsistent usage of terminology.

been reported to produce more FPs "inside" certain syntactic units, e.g., utterances, but there are also findings contradicting this.

In summary, the three numbered parameters at the beginning of this chapter are of interest, though their interpretation will likely not be as straightforward as has been suggested. The least problematic of these (perhaps because it is insufficiently researched) is the clustering – here, the conclusions of Kosmala and Crible (2022) will be tested. The data analysis part of this thesis will further consider the nature of the linguistic item following each FP, though the analysis will be extended beyond nouns, and will be more detailed. In this regard, L1 speakers will be analysed first to test whether de Jong's (2016) conclusions apply to them, with the L2 speakers being compared afterwards. The overall frequency of FPs in both groups will also be compared, and while the L2 speakers are (in accord with previous research) expected to produce them more frequently overall, perhaps the other parameters will help shed more light on why this is the case[4]. Regarding the position of FPs in relation to syntactic (or other) units, more research comparing L1 and L2 speakers should certainly be conducted. Before that can be done, however, the previous findings need to be understood more in depth. For this reason, the next chapter will focus on previous research in FP position, attempting to determine the best course of action for the data analysis part.

[4] Provided it, indeed, is the case.

2.3 On the position of FPs: Relative to what?

As explained in the previous chapters, the findings in FP positionality are rather conflicted. One reason why that is the case might be that researchers have perhaps chosen slightly (or completely) different syntactic units relative to which the FP positions were analysed[5]. Unsurprisingly, investigating FP positions in e.g., clauses *or* utterances will lead to completely different results. Individually, both might still be relevant, but that is not the case with comparing them. For example, an utterance can contain several clauses – if a particular set of speakers exhibits a tendency to produce FPs on either edge of a clause, the multi-clause utterances in the data might reflect this. Of course, the further trouble with this is that this could be taken as a proof of those speakers' tendency to produce FPs *utterance-medially*. While technically not false, such conclusion would overlook that those speakers pause in between clauses, instead viewing their speech production as more disfluent, due to the choice of *the utterance* as the framework for FP position analysis.

As such, choosing the appropriate unit is essential. De Jong (2016) settles on the ASU (Analysis of Speech Unit), following the choice of Riazantseva (2001), and Skehan and Foster (2007) (de Jong, 2016, p. 114). She finds that FP occurence inside the ASUs is more frequent in L2 speakers, and less so in native ones (de Jong, 2016, p. 130). To reiterate, Foster et al. define the ASU as an utterance "consisting of an *independent clause, or sub-clausal unit,* together with any subordinate *clause*(s) associated with either" (Foster et al., 2000, p. 365). It is also worth pointing out that the individual clauses are still marked and segmented. This means that ASUs can be syntactically very complex, theoretically even infinitely so. While an infinitely long ASU is not a realistic obstacle in speech analysis, there are several other problems associated with this choice.

Most importantly, the definition is too loose, too vague. In careful (e.g., written) language production, identifying superodinate and subordinate clauses, as well as grouping them into discrete sentences, is quite simple[6], save for the occasional ambiguity or a problematic case. But, of course, the ASU is intended for *speech* analysis. Let us thus

[5] While precise, this formulation is quite cumbersome and long. Thus, henceforth, "position (of FPs)" will refer to the position relative to a syntactic unit, unless stated otherwise.
[6] This is especially true for written language in great part due to capital letters and punctuation, but even careful, well-rehearsed speech is rather straightforward to analyse this way.

consider some specific problems with the ASU's definition, illustrated on the very data[7] which will later be analysed, and contrast it with the definitions of Foster et al. (2000).

Consider the following example:

EX. 1 *I thought that was what I wanted to do and I studied for that* (ETC: EN006)

This example is fluent and clearly contains four finite clauses. And yet, at least two problems with analysis are present here. The first is the coordinated clause "and I studied for that" – shall we consider this a separate ASU, or an element of the previous ASU "I thought that was what I wanted to do"? According to the ASU's definition above, the first option seems better, as coordination is not metioned. However, Foster et al. (2000) briefly address this later, suggesting that coordinated independent/main clauses be viewed as separate ASUs if there is at least a 0.5 second pause between them (Foster et al., 2000, p. 367). Further elaboration is not provided, making the pause criterion seem somewhat arbitrary. More troubling still, this is taken as a proof enough that the speaker is planning a separate ASU. If we consider the example above again, the first main clause has several dependent clauses following it – in such cases, it is not unreasonable to expect a pause merely for breathing or swallowing, which can easily surpass the 0.5 second criterion. To be fair, this is a broader issue in speech analysis, and certainly not unqiue to the ASU – though the pause criterion makes it all the more glaring.

The second problem lies in the segment "what I wanted to do". The fused relative itself could pose some problems, but it shall not be discussed here in detail – let us instead focus on the infinitive. Should this be considered a separate subordinate non-finite clause? Not according to Foster et al. (2000) – they claim that the non-finite verb must have some other clause element (object, subject, complement, or adverbial) to qualify as a non-finite clause (Foster et al., 2000, p. 366). It is once again not clear why this is the case, but even if it were clear, this particular case would not. Semantically, the object of "to do" is expressed by the fused relative pronoun, meaning that it is also present formally. It is worth noting that cases such as these are not addressed at all by the authors. Even if the analysis described in this paragraph were recommended for the ASU, one would still have to suspect that the rule is based on arbitrarily selected syntactic constructs, rather than on actual findings in cognitive linguistics.

[7] This data is spoken, see section 3.1 for a detailed description.

On a similar note, let us briefly consider this example:

EX. 2 *salary a month was pretty low in comparison with (er)* ***I don't know*** *a supermarket (eh) cashier* (ETC: CZ001)

Of interest here is the phrase "I don't know". One must indeed relate when asked whether this should be considered a part of the rest of the sentence – and therefore the ASU. A more rigid approach could view it as a discrete, merely inserted ASU – but then again the insertion implies some *association* with its surroundings. On the other hand, the phrase could be interpreted as an "irrelevance marker" of sorts, marking the following noun phrase as chosen ad hoc. As was the case with the fused relative example, Foster et al. (2000) do not address this type of constructions.

Either solution thus relies largely on interpretation, which is a broader issue in speech analysis. What makes the ASU more problematic, however, is the fact that in this particular case (and analogical ones, too), the phrase "I don't know" has to be declared either a discrete ASU, or a part of another. Thus, both analyses make a rather strong claim that can easily be argued against, whereas e.g., merely marking the phrase as a clause makes no such claims. The alternative would be to discard these cases as undecidable, but as they are fairly frequent, a substantial amount of data would be lost merely to "squeeze in" the ASU theory.

One final comment on the ASU has already been touched upon, so let us briefly reiterate and elaborate. Since the definition includes not only the main clause within the ASU, but also all subordinate clauses, particularly advanced and native speakers may produce highly complex and long ASUs. Such large units can be problematic for speech analysis for a number of reasons. For example, the length may trigger longer pauses (breathing, swallowing), which in turn may be misleading if pauses are treated as markers of ASU boundaries. Moreover, it is unlikely that speakers neatly plan the entire "large" utterances ahead, meaning that the ASUs would not correspond with their cognitive planning, making the ASU potentially quite unfit for (dis)fluency research.

Out of these issues, de Jong (2016) acknowledges the final one, recognising that disfluencies occur more the longer the ASU is, noting that native/advanced speakers will produce longer ASUs. To solve this issue, she chooses logistic regression analysis, treating word boundaries as potential places of occurrence for pauses (de Jong, 2016, p. 115). While

this is by all means a good step, the previously described problems remain, raising questions about the practical and theoretical plausibility of the ASU.

Boonsuk, Ambele, and Buddharat (2019) do not opt for the ASU, but their unit of choice is not exactly clear. Analysing disfluencies in informal conversations in L1 English, they interchangeably refer to the relevant unit as "the sentence" (Boonsuk, Ambele, and Buddharat, 2019, p. 136), and as "the utterance" (Boonsuk, Ambele, and Buddharat, 2019, p. 137). They define neither of these units – though it is likely that they mean the individual alternating turns of the speakers, since their data is conversational, and that both of these terms (are supposed to) have the same meaning. In any case, they do find a very strong tendency for FPs to occur in the medial position (Boonsuk, Ambele, and Buddharat, 2019, p. 136).

However, their conclusions still cannot be taken at face value. Firstly, unlike the data used for the present research, theirs was conversational. Despite the arguably problematic vagueness of their syntactic unit, it might have been fitting for their sample – though this is difficult to say with certainty for obvious reasons. As the sample in this study consists solely of monologues, analysing speaker turns is pointless. Secondly, their study deals only with L1 speakers, meaning that the results have no comparative implications on their own. Thirdly, although the study is quite recent, it uses data from Crystal & Davy (1975), which is some fifty years old. Despite the authors' assurances (Boonsuk, Ambele, and Buddharat, 2019, p. 135), the language has undergone many changes[8], making the comparison with today's English problematic.

As for Kosmala and Crible (2022), they choose a different approach altogether, analysing FP positions within intonation units. As stated earlier, they report a prevalence of medial position FPs, with no significant differences between L1 and L2 speakers (Kosmala and Crible, 2022, p. 228). This approach to FP position analysis is interesting, though it is not made clear why this unit was chosen, nor is it comparable with the findings discussed above, which the authors themselves acknowledge (Kosmala and Crible, 2022, p. 234).

[8] As theorised earlier, there might be a link between FP production and the nature of the following word. Because the processes of lexicalisation and grammaticalisation are constantly at play (yet nigh impossible to precisely account for), comparing the analysis of half a century old data with modern data could be very misleading,

Drawing a satisfying conclusion from the issues discussed in this chapter seems extremely difficult, as researchers have employed different methods of analysis, making their results nearly uncomparable. Furthermore, especially the first two methods discussed have serious theoretical and practical shortcomings. While these could potentially be salvaged by making some modifications, the modifications themselves would (mirroring the original issues) once again have to be fairly arbitrary. The method of Kosmala and Crible (2022) seems more consistent, but since the data used for our research is not annotated for intonation units, it cannot be implemented without lengthy reanalysis. Moreover, it is not clear whether analysing FP position in respect to intonation units actually reveals anything about the speakers' usage of FPs as, e.g., hesitation phenomena.

To end this chapter on a positive note, the previous research need not be mimicked. Because the choices of units above generally suffer from difficulties with interpretation, as well as from arbitrary parameters, let us instead devise an experimental, multi-level analysis which relies on subjective interpretation as little as possible. The positions of FPs in our data will be analysed in respect to the following:

1) The sentence

2) The finite clause

Even though choosing any unit is inherently somewhat arbitrary, the multi-level analysis better allows for comparison with other research, both future and past. Furthermore, these units can be rather easily defined and identified, making the analysis mostly objective. For a more detailed description of the analysis, see section 3.2. Let us now briefly formulate the research questions of this paper, concluding its theoretical part.

2.4 Research questions

The previous discussion has mostly focussed on contrasting possible differences in L1 and L2 FP usage with the findings of past research. In the light of these findings, the hypotheses have been adapted so as to allow for the results to be comparable with past research. The research questions of this paper thus are:

1) Do L2 speakers overall produce more FPs than L1 speakers?

2) Do L2 speakers cluster FPs more than L1s do?

3) Is there a general tendency to produce FPs more before (different kinds of) content words, as opposed to (different kinds of) functional words?

3a) And if yes, are there any differences between L1s and L2s?

4) Are L2 speakers more likely to produce FPs within sentences and finite clauses respectively?

As suggested earlier, answering these questions will be attempted on the basis of corpus data – the data and the method are discussed in the next section.

3. The data and the method

3.1 The source

Let us begin by describing the data used. All the data is taken from the English Teacher Corpus (ETC). The corpus in question is brand new, with the initial recordings having been recorded in the years 2022 and 2023. As the name suggests, the ETC is a corpus focussing on the language of actual (high school) teachers of English. This makes it both a highly specific and unique corpus, without a parallel worldwide.

The data for this corpus was gathered by teacher trainees studying in the Department of English Language and ELT Methodology at the Faculty of Arts at Charles University. One of the several benefits of this approach to data collection is that the trainees have varied educational backgrounds – meaning that the sample of teachers recorded is also varied, namely in terms of geography and the age of the respondents, as most of the trainees have contacted the teachers in their home region. Besides that, it also gives the trainees an opportunity to experience the teacher language first hand (from a new, linguistically aware perspective), which is further reinforced by the fact that the trainees partake in the transcription process. Naturally, this develops yet another core skill in linguistic analysis.

The basis of each recording was a series of five tasks, which will now be briefly characterised. The first task consisted of a monologue on a topic related to the teaching profession (e.g., “Why did I choose this profession?”), with the teachers having a choice of one of several topics suggested on a sheet of paper. The teachers were given a short time window (2-3 minutes) to think over what they will talk about – this window was not given for the other tasks. The second task was a dialogue between the interviewer/trainee and the teacher, where the trainee asked questions related to what had been said earlier in the monologue. The third task was a description of a series of pictures depicting a story. As for the fourth task, the subjects were asked to read aloud a text containing a number of low-frequency/foreign words, as well as numerals, represented by both Arabic and Roman numerals. Finally, the fifth task was identical to the first one, except in the teachers’ native language. Not all the teachers recorded had Czech as their (sole) L1, however. Bilingual subjects were asked to pick their dominant L1, while the subjects whose L1 was English did not partake in this task whatsoever.

The recordings took place in various environments, though attempts have been made to make the environment as quiet and as little distracting as possible. Due to the teachers' often tight schedules, their offices at their schools usually turned out to be the most plausible option – however, some were also recorded through an online call or in a cafeteria. As this was only the case with several English L1 speakers, these factors probably did not have a significant effect (e.g., anxiety due to non-native L2 production in public[9]), though it did make the transcription process more complicated. It goes without saying that such conditions are best avoided in future research. Finally, the recordings were made via ordinary mobile phone voice recorders whose quality is sufficient for this kind of research.

This initial series of recordings yielded fourty specimens between fifteen and twenty minutes long. Twenty-five of these were of native Czech speakers, while fifteen were of native English speakers. Both the interviewers and the interviewees also filled in a form providing metadata about themselves, such as language background, or time spent abroad.

Using the XMARALDA (Schmidt T and Wörner K [2014], "EXMARaLDA", In Handbook on Corpus Phonology, pp. 402-419. Oxford University Press. https://www.exmaralda.org/) software, the recordings were automatically transcribed and the transcribed segments were alligned with the recordings. At this stage, the transcription naturally contained errors and lacked all the annotation required for the corpus, such as FPs themselves. The interviewers then manually corrected the errors and annotated the transcription accordingly for all of their recordings, which was usually three per person.

The data thus transcribed and corrected was converted into text files – once again fourty (25 of Czech L1s and 15 of English). From these transcriptions, ten were picked randomly for each group of speakers. Everything besides the first task (monologue) was removed, including the language produced by the interviewers. In this state, the text files were uploaded into AntConc (Anthony, L. [2014]. AntConc [Version 3.4.4] [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software) where concordances for each of the six FP types (represented in the corpus as: *eh, em, er, erm, mhm, uhu*)[10] were created and then opened in Microsoft Excel, each in a separate tab. This

[9] This would also be less likely for teachers in general, as this would pose a serious complication for their ability to perform in their profession.
[10] The corpus also includes the FP represented as (mm) – though this one did not appear in the sample described above.

allowed for easy and clear navigation of the data, which was necessary for further annotation. Each FP occurrence also bore the information of the source recording, e.g., *CZ002*[11].

This initial grouping of FP occurrences yielded 451 instances of FP production. Of these, one *(erm)* of an English L1 was discarded, as it occurred at the very end of the monologue with no immediate linguistic surroundings. Furthermore, two instances of *(em)* from a single English L1 were discarded as well, as they preceded a proper noun (the name of a camp) – these types of words were not transcribed, but were replaced by a placeholder such as <*name of camp*>. The number of FP occurrences after these omissions for each FP type is represented in the table below, for both speaker groups, and in total.

Table 1.: The number of the individual types of FPs in L1s and L2

FP type	**CZ L1**	**EN L1**	**Total**
(er)	111	36	147
(eh)	59	13	72
(uhu)	2	0	2
(erm)	59	94	153
(em)	33	38	71
(mhm)	2	1	3
Total	266	182	448

This raw data had to be modified, analysed and tagged. The next section describes these processes in detail.

[11] = Czech L1 speaker number 2.

3.2 Data processing procedure

For the first research question, it was necessary to count the number of tokens produced by each speaker and by each group in total. Let us now describe this procedure, and the decisions made when deciding problematic cases, in order to determine the number of tokens produced.

Due to the nature of the annotation of the corpus, certain "tags" had to be removed. The speaker turn tags (*B:*) were removed first. As mentioned above, some proper nouns were not included in the transcription – these would usually be names of people personally known by the speaker, or names of institutions, "events" (e.g., agencies, camps). Instead, a placeholder such as <*name of language school*> would be used – these were removed from the sample (6 in total, only for English L1 speakers), as it is unclear how many tokens constitute these proper nouns. Similarly, indiscernible segments were transcribed as e.g., <*XX*> (indicating two indiscernible tokens) – these were also removed, as the indicated number of tokens is merely assumed by the transcriptionists, or even guessed (9 assumed tokens for L2s, 6 for L1s discarded). Finally, some words were annotated as e.g., *for*<*?*> – indicating uncertainty of the transcriptionist. Only two such cases appeared in the sample, and as they were both monosyllabic (meaning one can be almost certain they are not multiple tokens), they were not discarded, but treated as a single token.

A specific problem arose with certain other proper nouns, particularly with those denoting geographical and political units, as well as with some foreign words. For the proper nouns, occurrences such as *The Czech Republic* or *The United States* were treated as three distinct tokens. *Sri Lanka* appeared once in the sample, and was treated as two tokens – this decision was made due to it often being referred to as simply *Lanka*, and because the orthography influences it being perceived as a multi-word unit. Names of institutions (e.g., *Charles University*) were also treated similarly – the number of tokens corresponds to the number of orthographical words. Regarding foreign (non-English) words, they were usually annotated as <*foreign*>. These cases were also discarded, though interestingly, only a single, English L1 speaker (EN005) in the sample produced these – twice. Finally, the sample also contained the word *au pair* (CZ018) – similarly to *Sri Lanka*, this was also treated as two tokens.

Compounds presented another problem, as expected. For the sake of simplicity, compounds written as a single word were analysed as a single token (*classroom*), as were hyphenated ones (*peek-a-boo*) – compounds written as X individual words (*phrasal verbs*[12]) were analysed as X tokens. The constraints of this approach are obvious, but even though orthography is external to the linguistic system, it can still indicate a level of intergration of some compounds into the lexicon, albeit to a rather unreliable extent. Nonetheless, this still seems less arbitrary of an approach than individually evaluating the degree of "compoundness" of each compound in the sample in order to determine the number of tokens. Initialisms (*UK*) and acronyms (*TEFL*) were treated as a single token.

The sample further contained a considerable number of false starts (*about somethi= doing something* – CZ025). The procedure here was to keep all "finished" words, even if they did not "fit" the surrounding syntax. Unfinished words were removed from the sample. As for contractions (*don't*), they were treated as a single token. Complex numerals (*nineteen eighty nine*) were analysed as multiple tokens. And finally, repetitions (*I I I I*) were kept intact. This resulted in two corpora, for L1s and L2s respectively. The L1 corpus consisted of 5475 tokens. The L2 corpus was slightly larger, with 5640 tokens.

In comparison, determining the nature of the following word was a significantly more complicated process. A system of word class tags was created, which will now be briefly described, after which problematic/dubious cases will be discussed in more detail. The tags/following words were assigned – based on their nature – into "lexical", "grammatical", and "other" types of units.

LEXICAL UNITS

Nouns (*n*), adjectives (*adj*), numerals (*num*), lexical verbs (*lex*), and idioms (*idiom*) were classified as lexical, autosemantic units. Morphological nouns used as syntactic adjectives (*a language school*) presented a recurring problem. Given that FP production is more closely associated with syntactic planning, their syntactic function was given priority here. For similar reasons, gerunds were treated as nouns (unless they functioned as a premodifier), and the present participle was treated as an adjective (in the premodifier position). Some FPs were directly followed by an article, after which a noun, an adjective, an

[12] Here as a token appearing in the sample, though it is worth mentioning that phrasal verbs themselves were analysed as two or three tokens.

adverb, or a quantifier (see *GRAMMATICAL UNITS*) followed. These were considered distinct from immediate FP-word occurrences, and were glossed as e.g., *iadj* (indefinite article preceding an adjective). Cases of repeated articles (*a a new teacher*) were analysed as if a single article occurred.

Furthermore, ordinal numerals (*first*) were glossed as adjectives. Regarding idioms, only one "true" idiom was present in the sample (*go with the flow*). Since idioms are not compositional in their meaning, but rather function as a single lexeme needing to be retrieved, this case was classified as a special, discrete unit, rather than a lexical verb (*go*). Lastly, lexical verbs in the infinitive form where *to* directly followed an FP were given a special suffix (*lexinf*) to reflect the presence of the infinitive marker, but were still considered lexical units.

GRAMMATICAL UNITS

This group consisted of pronouns (*pron*), prepositions (*prep*), conjunctions (*conj*), quantifiers (*q*), the negative particle *not* (*not*), determiners (*det*); and auxiliary (*aux*), modal (*mod*), copular (*cop*), and causative (*caus*) verbs. As suggested earlier, quantifiers (*a lot*; *a couple*) were often preceded by an article – in fact, only one instance (*much*) in the sample was without an article. No difference was made between the two types of quantifiers in the classification, they were all marked as *q*. The determiner category was created to accommodate instances such as *(em) some things* – possessives and demonstratives were grouped with pronouns. Furthermore, adverbial interrogatives and deictics introducing dependent adverbial clauses were classified as pronouns. This included adverbial deictics such as *when* or *why*, but not words like *nowadays* and *sometimes* – those were grouped with adverbs.

The verbs also require some further commentary. Copular verbs are sometimes considered autosemantic units, but they are treated as grammatical units here, for all but one instance (*to become*) are of the verb *to be*. As can be seen in certain dialects/varieties of English (such as the AAVE), this copula is prone to being dropped (seen in many other languages as well), which is characteristic for grammatical items, but not for lexical ones. The verb *to become* was treated similarly for the sake of consistency. Just like with lexical verbs, some copular verbs appear in the infinitive form, with *to* directly following the FP – these were marked as *copinf*. Only a single instance of a causative verb appears in the sample (*have*

them learn certain key words), but a new category was created regardless, since it is a highly specific verb type.

OTHER UNITS

Certain units which followed an FP eluded traditional classification. These were FPs themselves (*FP*), and a varied group of discourse markers (*DM*), which included both single words (*like*; *okay*; *yeah*; *alright*), and chunks (*kind of*).

For the final research question, each FP was analysed for its position within a finite clause and the sentence. This process turned out to be more problematic than the others, particularly due to the sentence parameter.

Where does a sentence begin, and where does it end? A good indicator could be conjunctions, both subordinating and coordinating – finite clauses linked together with any conjunction should be a part of the same sentence. However, the first of many problems lies in the abundance of conjunctions (especially coordinating *and* and *but)*, which were often used to link (and coordinate) a large number of finite clauses. One of the more extreme examples is this (conjunctions connecting clauses in bold):

EX.3 *I was very much against (er) teaching* ***because*** *a lot of people in my family are teachers (em)* ***and*** *I thought (er)* ***that*** *that wasn't the right path for me (erm)* ***and*** *when I started studying (erm) (em) (er) English at university I thought* **(that)** *I would become a translator (erm)* ***and*** *(er)* ***but*** *(eh) throughout my studies I had multiple chances of (erm) teaching (er) substituting for other people (erm)* ***and*** *(erm) contrary to* ***what*** *I had planned I (er) realised* ***that*** *I really enjoyed the job* ***and*** *I really (erm) I was really good at it as well (er)* ***if*** *I may say so* ***and so*** *(er) I eventually came to terms with the fact* ***that*** *that's going to be* ***what*** *I have (eh)* ***what*** *I have to do* (ETC:CZ003)

If we follow the assumption above, this entire segment would be a single sentence. Provided that FPs indeed are often produced to buy time for syntactic planning, the relevance of the sentence (or the utterance and the ASU, for that matter) as a unit of analysis seems questionable in these cases. This particular "sentence" clearly was not planned as a whole, as is evident from the number of disfluencies, which – curiously – often precede finite clauses. It is not unreasonable to suspect that they sometimes function similarly to FPs, that they are produced to buy time. However, even if that were the case, it would still be impossible to

distinguish them from the more "connective" uses. As such, cases like the example above were analysed as entire sentences. Conversely, this approach splits asyndetically connected clauses:

EX.4 *it could be more interactive (er) it could be focused on cause and events* (ETC:EN006)

Cases like these seem more tightly connected in terms of meaning than many clauses of the previous example. An alternative to this approach would thus have to be based on semantic criteria – which unfortunately tend to be rather arbitrary. For the purposes of this paper, all finite clauses linked by a conjunction will constitute a sentence, including example 4. Conjunctions also were not considered a part of a clause, meaning that cases where an FP directly followed a conjunction were analysed as clause initial.

Finally, regarding conjunctions, *so* was somewhat problematic, since it has several different functions, including the discourse marker function. As a DM, it usually appears at the beginning of a clause, which caused problems in the analysis. There was no ambiguity when it appeared at the beginning of each recording, because it clearly cannot function as a conjunction there. The rest of the occurrences were treated like conjunctions when connecting finite clauses.

The next issue has already been mentioned in the theoretical part – comment clauses such as *you know* or *I don't know*. These were treated as parts of a sentence if inserted in the middle of a sentence, and there were no occurrences at the edges in the sample. Because they are likely stored as chunks rather than planned in real time, they were analysed as sub-clausal units.

Not all FPs were produced within or before a finite clause. A total of 8 FPs occurred either within a non-finite clause lacking any syntactic connection to its surroundings, or a sub-clausal unit. For example:

EX.5 *which is nice so (er) second question* (ETC:EN009)

As mentioned earlier, *so* was only treated as a conjunction when linking two finite clauses, meaning this particular sub-clausal unit would not be a part of the previous sentence/clause. Thus the FP's position here is marked as *NF* (non-finite construction) – no distinction was made between non-finite clauses and such verbless constructions. The occurrences marked as *NF* were discarded (9 in total) from the calculations for the fourth

research question, though they still contain valuable information for the other research questions and as such will be used to answer them. A different approach was employed for multiple clause elements realised by a non-finite clause, such as objects and noun modifiers:

EX.6 *I wanna continue living over here* ***experiencing the things*** *I'm experiencing (em)* ***doing the things*** *and traveling to places and (eh) just really enjoying the[i:] experience and getting to be independent and I want that to continue* (ETC:EN007)

*Doing/experiencing the thing*s have no overt connection to the finite clause preceding them. Despite this, they were considered a part of the preceding clause due to the fact that *continue* is a phase verb, meaning it needs another non-finite verb as a complement, in this case the present participle. Since *experiencing* and *doing* have the same form as *living*, it is fairly safe to assume they depend on *continue*. Thus, both FPs in 7 were marked as sentence-medial and clause-medial. The approach was analogical in all similar cases. Let us now turn to data analysis and discussion of the results.

4. Data analysis and discussion

4.1 Overall frequency of FP production

After these changes to the data, the tokens and FPs were counted for each sample, after which the frequencies for each speaker (per 100 tokens) were calculated:

Table 2.: The number of tokens and FPs for each L2 speaker

Speaker (L2)	**001**	**002**	**003**	**004**	**005**	**010**	**013**	**018**	**020**	**025**	**Total**
Tokens	600	816	679	332	950	673	340	441	420	389	5640
FPs of these	39	18	96	11	15	27	24	18	14	4	266
Frequency per 100 tokens	6.5	2.21	14.14	3.31	1.58	4.01	7.06	4.08	3.33	1.03	4.72

Table 3.: The number of tokens and FPs for each L1 speaker

Speaker (L1)	**001**	**002**	**003**	**004**	**005**	**006**	**007**	**008**	**009**	**010**	**Total**
Tokens	872	382	436	392	445	417	510	437	756	828	5475
FPs of these	25	41	2	13	13	17	12	9	22	28	182
Frequency per 100 tokens	2.87	10.73	0.46	3.32	2.92	4.08	2.35	2.06	2.91	3.38	3.32

Coincidentally, both groups of speakers come very close in number of tokens (ratio of 376:365) produced, with L2s having produced slightly more. The L2s have, however, produced considerably more FPs (19:13). Running a Chi2 test suggests a statistically significant difference (Chi2=13.9187; P-value=0.0001909). This is very much in accord with what the previous research suggested – L2 speakers on average produce more FPs than L1s do.

However, it is worth noting that the differences between the individual speakers can be quite massive, even within their group. The most noticeable outlier here would be the speaker CZ003, with 96 total FPs produced – no speaker in either group produced at least half as many. Conversely, certain speakers barely produced any FPs whatsoever, though this does not necessarily mean that they are "more fluent", as our analysis only considers FPs, not the other types of disfluencies. Unsurprisingly, the speakers who produced fewer tokens also tended to produce fewer FPs, perhaps with the exception of EN002. These two speakers, CZ003 and EN002, were most certainly somewhat distorting the results, particularly CZ003. Two additional calculations were thus carried out, one excluding the subject CZ003, and one excluding both CZ003 and EN002.

For the first calculation, 679 tokens were removed (4961 tokens afterwards), as well as 96 FPs (170 FPs afterwards). The FP frequency per 100 words dropped to 3.43. Running a Chi2 test also showed that the statistical significance of this difference is much lower (Chi2=0.0839; P-value=0.772). This is at odds with the suggestions of previous research – the frequency of FPs is nearly the same for both groups: 3.32 in L1s and 3.43 in L2s.

Removing both outliers yields a much more interesting result. 382 tokens were removed (5093 afterwards), and 41 FPs (141 afterwards). The frequency in L1s dropped to 2.77 FPs per 100 words. Chi2 test suggests borderline significance (Chi2=3.6322; P-value=0.05667) – which is slightly below the threshold of acceptability. There are two ways of interpreting this – either we proceed "by the book", meaning no significant difference in frequency was found; or we ignore the results falling slightly short of the threshold. As the difference in frequency is rather hard to simply dismiss, the second approach seems more fitting. As such, our sample shows that L2s indeed produce FPs more often than L1s, though with the caveat that this difference is right on the border of statistical significance. Let us now proceed to FP clustering.

4.2 FP clustering

FP clusters appeared fairly sporadically in the sample – only 12 clusters were found across both speaker groups. 10 of these were clusters of 2 FPs, 8 by L2 speakers, 2 by L1s. Only 2 consisted of 3 FPs, and both of these were produced by L2 speakers.

Making any generalisations from these numbers is difficult. While the difference in frequency should be statistically significant (Chi2=5.1051; P-value=0.02386), a total of 6 (of these 1 three-FP) clusters were produced by a single speaker: the aforementioned, highly disfluent CZ003. Removing these balances the ratio down to 4:2, and makes the difference statistically insignificant (Chi2=0.8809; P-value=0.4351). As with the previous research question, FP over-production of the subject CZ003 greatly distorts the results, here especially so, given that the total number of clusters is quite small, meaning the speaker had to be removed from the equation. Thus, no true difference was found between the frequency of FP cluster production of L1 and L2 English teachers, at least within our sample. It could be interesting to inspect the second task in the corpus in particular – perhaps analysing conversational tasks could provide more insight into FP clustering due to the greater pressure and tempo. For the present, this question remains unanswered. The next section deals with the nature of words following an FP, where the results are far more interesting.

4.3 The nature of the words following FPs

LEXICAL UNITS

Of the total 448 FPs in the sample, 128 were produced before lexical words – almost one third. The number of occurrences for each type of words is represented in the table below:

Table 4.: The total number of FPs produced before each type of lexical words in each speaker group

Word	**Adj**	**Dadj**	**Iadj**	**Adv**	**Iadv**	**N**	**Dn**	**Lex**	**Lexinf**	**Idiom**	**Num**	**Total**
L1	10	1	1	6	0	10	1	16	2	1	0	48
L2	5	3	4	16	1	18	6	22	3	0	2	80
Total	15	4	5	22	1	28	7	38	5	1	2	128

As expected, L2 speakers have produced noticeably more FPs than L1s overall. While some of the categories, such as idioms and numerals, have minor representation, more can be interpreted from the others. The difference is perhaps the most noticeable in adverbs. Two possible explanations present themselves. The first is that not only retrieval is required, but also the real-time adverbial derivation from an adjectival form, or even that extra time is needed to identify the adverbial form if both forms are stored as discrete lexemes. The second explanation would be simply that many adverbs occupy (or can occupy) clause-initial positions, usually following a conjunction, such as *because (er) nowadays*. If L2s truly produce more FPs at clause boundaries, this could merely be a side effect of that – not an indication that adverbs take especially long to retrieve/derive.

Regarding adjectives, it is worth pointing out that several cases of an FP splitting an article from an adjective were observed, for example: *I was a (eh) a a new teacher*. This one is dubious at best, as this particular FP sounds the same as the indefinite article, meaning it could only be a case of quadruple repetition. The other cases were not immediate, some other words stood between the FP and the article. Many adjectives also occurred with either uncountable nouns, or nouns in plural (and without any articles). If there indeed is any relation between producing FPs more often before an adjective premodifying a noun of a particular definiteness (which does sound fairly strange now), the results do not seem to suggest it. Interestingly, ignoring the articles as a factor altogether shows that both groups

have produced the exact same number of FPs before adjectives – 12. Lastly, as stated earlier, nouns, gerunds, and participles in the premodifier position were tagged as adjectives. Several L1s produced fairly infrequent words in these positions (e.g., *methodical*, *database*, *mentoring*), more so than L2s, which could explain the (rather unexpected) lack of difference in the frequency of pre-adjectival FP production between the two groups.

In nouns and lexical verbs, the general tendency for L2s to produce more FPs shows yet again. Lexical verbs in particular are interesting, as a total of 9 cases (2 for L1s; 7 for L2s) have been observed, where the FP stood right between the subject (typically pronominal) and the finite verb. This could once again be caused by the need for syntactic planning. Because of the "fixed" word order of English (and it not being a pro-drop language), the subject is mostly in the first position, as the data is from a monologue (barely any imperative or interrogative clauses), only some adverbials could be placed before it. Due to these reasons, there is not much planning associated with the subject. Thus, these examples are probably rather similar to the "true" clause-initial FPs. If that really is the case, the higher FP frequency in L2s in these positions is in accord with what was expected.

GRAMMATICAL UNITS

281 FPs occurred before grammatical words – the majority, which is not surprising given their omnipresence in English:

Table 5.: The total number of FPs produced before each type of grammatical words in each speaker group

Word	**Pron**	**Prep**	**Conj**	**Aux**	**Mod**	**Cop**	**Copinf**	**Caus**	**Q**	**Det**	**Not**	**Total**
L1	48	8	51	2	1	3	0	1	2	1	0	117
L2	80	24	41	4	3	5	2	0	4	0	1	164
Total	128	32	92	6	4	8	2	1	6	1	1	281

Similarly to lexical units, most categories have too poor of a representation to make any conclusions. However, a similar trend of an overall higher FP production in L2s is observable as well. The grammatical types of verbs were rarely preceded by an FP, but are interesting precisely for this reason. Auxiliaries, modals, and copulars are abundant in English, and yet only a handful occurred after an FP. As discussed earlier, declarative

sentences and clauses dominate monologues, and they cannot begin with a verb. Thus, the much higher FP occurrence before lexical verbs (as opposed to grammatical ones) could support the theory that lexical units are more likely to be preceded by an FP. More insight could perhaps be gained from looking at all gerunds and infinitives as verbs – though another constraint here would be having to distinguish, e.g., regular gerunds from gerunds fully lexicalised into nouns. For example, a word like *teaching* could be very difficult to decide. Furthermore, it would be worth inspecting the finite verb forms in monologues to find out the frequency of verb forms which require auxiliares, and those which do not. The ratios could then be compared with the ratios of FP occurrences. But that lies beyond the scope of this thesis.

Let us now turn our attention to the words before which FPs occurred very frequently: pronouns, prepositions, and conjunctions. FPs preceded pronouns more frequently than any other type of words. Out of the 128 pre-pronoun occurrences, 118 occurred clause-initially (45 for L1s; 73 for L2s). This supports the theory that these cases have more to do with syntactic planning at the clause level, not with retrieval. The higher number of FPs produced by L2s is once again in accord with the previous findings.

There was a rather large difference in FPs preceding prepositions. L1s have almost exclusively produced them clause-medially (7 times), and only once clause-initially. L2s on the other hand produced 11 FPs in the initial position, and 13 clause-medially. Once again, the higher cognitive load at the beginning of a clause could be an explanation for the much higher initial occurrences for L2s. It is also worth pointing out that prepositions occurred clause-initially in various adverbials (*and **at** the beginning*). The clause-medial occurrences are also quite interesting, as it is not clear what triggers the disfluency. It could be the retrieval of the preposition itself, but it could be the retrieval of the following noun. The retrieval of the preposition here would not be so much about "remembering" the preposition, but rather about selecting the correct preposition for a particular meaning. For example, an L2 speaker produced an FP before *at that time* – the Czech equivalent *v té době* uses the preposition *v*, whose closest English equivalent would be *in*. In these cases, interference could also be a factor. Lastly, there was also one case where an FP was produced right before a self-correction with a preposition, such as:

EX.7 *flooded with PHD (erm) with PHDs and so I realised* (ETC:EN006)

Such cases of FP production cannot really be attributed to retrieval, though syntactic planning could be at play, given that another clause directly follows.

Conjunctions were the only word type which was noticeably more often preceded by an FP in L1 speakers.[13] L1s have produced roughly 25% more FPs before conjunctions (51:41). How can this be explained? It could be that native speakers simply produce more complex sentences, which translates to more conjunctions, and thus more opportunities and reasons for FP production. Looking at our data, we find that all but one occurrence (by an L1) appeared clause-initially – that is 50 for L1s, and 40 for L2s, one L2 case appeared in a non-finite construction. Regarding clause-initial, but sentence-medial occurring conjunctions, L1s have produced 46 FPs before these, whereas L2s 36. This means that sentence-initially, L1s have produced 4 FPs, and L2s also 4. The effect of overall more complex sentences in L1s might (help) explain the higher frequency of FPs in the clause-initial, sentence-medial position, though testing this would require analysing the syntactic complexity of both groups, the number of sentences and their clauses. Finally, while the number of sentence-initial occurrences is the same for L1s and L2s, the approach to conjunctions described earlier could certainly be distorting the results – hence the low number of FPs before sentence-initial conjunctions (and the low number of sentence-initial conjunctions in general).

OTHER UNITS

As for FPs, 13 FPs were produced before another FP, 2 by L1s, 11 by L2s. FPs appeared 26 times before a discourse marker, 15 times for L1s, and 11 times for L2s. Once again, generalisations are difficult to make from these numbers. The FPs preceding FPs reflect the higher tendency for FP clustering in L2s, though this has already been discussed. As for discourse markers, L1s surprisingly produced more FPs before them, though this is by no means a significant difference (Chi2=0.7418; P-value=0.3891). Thus, not much more can be said about these cases.

In summary, nearly everywhere where FPs were produced, L2s consistently produced more. This is true for both lexical and grammatical words. It is difficult to say whether FPs are truly more likely to occurr before lexical words than before grammatical words – this would require counting the total number of these words in the sample and comparing them

[13] Technically, besides adjectives without an article, though as discussed earlier, it probably makes more sense to ignore the article, and thus each group produced exactly 12 FPs before adjectives.

with the number of FP occurrences before each type. With that said, and as discussed in this chapter, there may be many other factors influencing FP production, and some of them may not even be related to the lexical/grammatical nature of the following word at all. The idea that lexical words take more time to retrieve sounds logical, but it is extremely hard to prove on this kind of data, due to the interfering factors. Let us now turn to the final research question.

4.4 Syntactic positionality of FPs

As mentioned earlier, 9 FPs occurred in non-finite consturctions, leaving 439 cases for positionality analysis. These cases were classified in one of three ways: clause- and sentence-initial (*CiSi*); clause-initial and sentence-medial (C*iSm*); or clause- and sentence-medial (C*mSm*). After counting them, the numbers look as such:

Table 6.: FPs produced in the three syntactic positions by each group of speakers

Position	**CiSi**	**CiSm**	**CmSm**	**Total**
L1	54	72	52	178
L2	53	106	102	261
Total	107	178	154	439

Overall, the most frequent occurrence was in the CiSm position, though once again, this might be a result of the way conjunctions were treated – sentences tended to be very long in the analysis. Because of this, there are also comparatively fewer CiSi occurrences. Interestingly, both groups exhibit almost the exact same behaviour in the CiSi position. This seems to suggest that regardless of the nature of acquistion (native X non-native), speakers tend to produce FPs when beginning a sentence. This makes sense, as they need to combine the processes of conceptualising the actual meaning, word retrieval, and then syntactic planning.

The CiSm position finally shows a difference in FP production – L2s have produced more FPs by about a third (Chi2=5.6153; P-value=0.0178). This position was also the most frequent position where FPs appeared for both groups. However, this does not necessarily mean that L2s struggle more with planning clauses – though it is probably also a factor. As discussed in the previous section, there are at least several other reasons why speakers might produce FPs, for example, the nature of the following word, or its frequency. Since L2s seem to produce FPs more often before nearly all kinds of words, the extent to which syntactic planning is responsible for the higher FP production in this position is unclear.

While these issues persist, the gap was the largest in the CmSm position, where L2s produced almost twice as many FPs (Chi2=14.9942; P-value=0.0001078). This is in accord

with the observations of de Jong (2016), as discussed in the theoretical section. With these numbers, it is farily safe to say that L2s encounter significantly more problems within finite clauses than L1s.

As with the first research question, the outliers could be distorting the results. A calculation similar to the one carried out earlier was done here as well – the speakers CZ003 and the speaker EN002 were removed. Removing CZ003 from the equation lowers the number of tokens to 4961, and the number of analysable FPs in L2s to 167 (by 94) – 15 CiSi, 45 CiSm, and 34 CmSm were removed. As for the speaker EN002, 382 tokens were removed, resulting in 5093 total tokens. Then, 41 analysable FPs were removed for L1s, 7 CiSi, 20 CiSm, and 14 CmSm. The numbers then changed as follows:

Table 7.: FPs produced in the three syntactic positions by each group of speakers (CZ003 and EN002 removed)

Position	**CiSi**	**CiSm**	**CmSm**	**Total**
L1	47	52	38	137
L2	38	61	68	167
Total	85	113	106	304

Curiously, L2s still produced fewer CiSi FPs than L1s, though they produced more FPs in the other positions, just as before – the differences, however, are much less drastic. The differences in the CiSm position bear no statistical significance (Chi2=0.98383; P-value=0.3213) themselves, though it is interesting that this was still the most frequent position for L1s, whereas not for L2s, as was the case before. The most significant difference (Chi2=9.3973; P-value=0.002173) was found in the CmSm position, which has not changed. This strongly supports the observation formulated above – in accord with the previous research – that L2s produce noticeably more FPs within finite clauses (and thus within sentences). It has to once again be reiterated that determining the precise explanation for this behaviour is not straightforward. Besides the aforementioned syntactic planning and word retrieval, morphological processes could be at play, or real-time derivation from a stored stem. For now, our data provides yet another proof of this tendency in the CmSm position. Let

us thus conclude the data analysis and proceed to the final discussion and summary of the findings.

5. Summary and final discussion

5.1 Summary of the findings

At the beginning of this thesis, a question whether L2 English teachers differ from L1 English teachers in FP usage was posed. The research thus far suggested a difference, particularly regarding frequency of FP production in various situations. An overall higher frequency in L2s has been reported almost unanimously. Our data (a monologue task) supports this, though the difference is of borderline statistic significance. There is a large variability from speaker to speaker, with some L2 speakers producing fewer FPs than most L1 speakers, though the opposite is more frequently the case.

L2s have also been reported to cluster FPs more often, and while this was also the case within our sample, an outlier who produced half of the total clusters was present. Removing this speaker also removed any significant difference in FP clustering between the two groups.

Another theory regarding FP production assumes that FPs are more likely to be produced before content words, more so in L2s. Our data shows that L2s produce more FPs before nearly all kinds of words in the sample, but concluding that this is caused by the nature of the following word is not possible, due to the reasons discussed earlier.

Finally, the hypothesis that L1s and L2s differ in FP production regarding finite clause and sentence positionality was tested. L2s in our sample exhibit a rather strong tendency to produce FPs inside finite clauses, both with outliers included and excluded. We do not find any significant differences in the other positions. Let us thus move on to the final discussion.

5.2 Final discussion

Reagrding FPs, the English of teachers who teach it is rather similar to the observations made in non-teacher speakers before. In other words, the generalisations suggested by previous investigation of FPs in L1s and L2s can also be largely applied to teachers. The general tendency for L2s to produce more FPs is also found in our sample. It is important to point out that this does not directly translate to lesser fluency, as FPs are only one of the several disfluency types. Nor is higher disfluency synonymous with the production of "worse" English, as it might help facilitate other aspects of "proper" English, such as accuracy and lexical richness.

The exact cause of higher FP production in L2s is very difficult to identify, though it is quite likely that multiple factors are at play. For example, de Jong (2016) attributes this to L1s producing language in larger chunks, which gives them fewer opportunities/reasons for disfluency (de Jong, 2016, p. 114). This is a plausible theory, and the findings in FP positionality of this thesis would support that, namely the finite clause-medial occurrences. Besides syntactic planning, the higher FP production in L2s is sometimes attributed to (greater) problems with retrieval (Boonsuk, Ambele, and Buddharat, 2019, p. 133). As discussed earlier, it is never quite clear which of these two factors causes a particular FP occurrence. Retrieving a word does not entail simply remembering the phonological form, but also its morphosyntactic properties, as well as syntagmatic properties (such as collocability). While it is fairly safe to assume that (at least most) L2s will encounter more problems with these processes, claiming L2s produce FPs more often before lexical units due to their lexical nature can be misleading. Not even a lexeme of a very low frequency is straightforward to analyse in this sense, since the low frequency can also translate to problems with syntactic planning, e.g., in terms of verbal/nominal valency. Even the higher pre-lexical FP occurrences in L2s observed herein could be misleading in this way, and thus cannot be taken as a proof of this theory. All of the FP usage parameters investigated in this thesis are nigh impossible to separate from one another.

Another limitation of our research is the poor representation of certain analysed parameters, such as FP clustering, or some of the specific types of words. Very little can be said about them, as they are not very frequent, and a much larger sample would be needed to provide more insight. Another issue stems from the positionality parameter. While the finite clause unit is fairly straightforward, the same cannot be said for sentences. The abundance of

ambiguous word types (such as *so*) in the sample makes it very difficult to reasonably segment the monologue into sentences. It has already been shown that the ASU does not address this problem. The utterance, while unfit for our data, could work quite well with the conversational task. Perhaps another solution could be offered by the approach of Kosmala and Crible (2022), who use the intonation unit for determining FP position. With this approach, they suggest that two kinds of FPs can be distinguished: *filler-as-signal* and *filler-as-symptom*, based on the position within the intonation unit (Kosmala and Crible, 2022, p. 237). Following this approach could help distinguish between signalling FPs and "true" disfluent FPs, which in turn could help better assess the degree of disfluency in L1s and L2s, instead of looking at every single FP in the sample.

Furthermore, other factors could be influencing the higher production. For example, L2s – teachers especially so – are under greater pressure during the recordings. In order to appear as advanced as possible, they could be consciously opting for more complicated morphosyntax and lexicon, resulting in greater cognitive load.[14] Most non-native teachers also less frequently encounter the need for more complex language production. Classroom English is usually simpler, and more importantly, more formulaic. Their being used to this type of language could make producing a more or less spontaneous monologue rather challenging, at least compared to L1s, who often use English in e.g., their household.

Finally, our sample was purely a monologue. This means that the speakers can avoid certain constructions/lexemes, both consciously and unconsciously. Both L1s and L2s could thus be more fluent in this task (Task 1) than in the dialogue (Task 2), where the interviewer asks questions which might demand a more complex answer. Conversely, the interviewer could also be priming the speakers, suggesting lexemes and morphosyntax. It would be very interesting to see the speakers' performance in Task 2 (or the others, for that matter), though that is a question for another day, another thesis. Let us now briefly conclude.

[14] From personal experience, two of the teachers I have recorded explicitly expressed nervousness beforehand. They also rather self-loathingly warned me that their English is not great. While this was not really the case, the added emotional stress could be influencing their disfluency.

6. Conclusion and final thoughts for future research

Within the framework of the present thesis, many of the initial questions cannot be answered with certainty, and more have arisen in the process. It seems that L2 English teachers in many ways resemble "regular" L2s when it comes to FP production. However, with the ETC expanding, more and more data is waiting to be analysed. As suggested earlier, the conversational task might hold much more insight into disfluency behaviour of English teachers. The relationship between FP production and word retrieval could perhaps be better elucidated by analysing the picture/story description task (Task 3), where the speakers must retrieve specific words for various objects in order to describe the story in the pictures. Analysing the frequency of words used there might also provide deeper understanding of FPs' role in retrieval. Finally, a cross-linguistic analysis of the L1 monologue (Task 5) by L2s could compare disfluency behaviour of L2s in their native language with that in English.

To reiterate that which has been said at the very beginning: understanding teacher language is vital for understanding second language acquisition in students. To what extent do students "copy" the linguistic behaviour of their teacher? Does the disfluency of their model speaker influence their disfluency production? Are the observed disfluency strategies (such as holding the floor) transmitted to the students? Answers to these and many other fascinating questions might very well be possible to find in the ETC corpus. For now, this thesis is only a small step towards understanding English teachers' disfluency, though hopefully, an important step.

Prohlášení o použití umělé inteligence

Během vypracovávání své bakalářské práce jsem nepoužil žádné nástroje generativní AI.

V Praze, dne _____

Filip Drábek

Declaration on artificial intelligence usage

In preparation of my bachelor's thesis, I have not used any generative AI tools.

In Prague, on the _____

Filip Drábek

7. Resumé

Tato bakalářská práce se zabývá vyplněnými pauzami v učitelské angličtině, a to na základě monologů nahraných pro korpus učitelské angličtiny ETC. V práci se srovnávají rodilí a nerodilí mluvčí, přičemž základní hypotézou je, že se tyto dvě skupiny budou v produkci lišit.

Teoretická část pojímá vyplněné pauzy a jiné typy disfluencí nikoliv jako „patologické" jevy, nýbrž coby přirozenou součást spontánní řeči. Předloženy jsou argumenty pro přirozenost disfluencí, například jejich všudypřítomnost i v jazyce rodilých mluvčích. V návaznosti na předchozí výzkum probírá možné příčiny produkce disfluencí a zejména vyplněných pauz, mezi něž patří třeba udržení si konverzačního tahu, potýká-li se mluvčí s problémy s plánováním své promluvy. Dále probírá podstatu těchto problémů, které literatura zpravidla vysvětluje obtížemi se syntaktickým plánováním a vybavováním slov, zejména pak málo frekventovaných, autosémantických slov. Na základě literatury formuluje první výzkumnou otázku, respektive první rozdíl v produkci vyplněných pauz mezi rodilými a nerodilými mluvčími, který bude později testován: celkově vyšší produkci vyplněných pauz u nerodilých mluvčích. Druhá výzkumná otázka se týká shlukování vyplněných pauz, tedy produkce několika vyplněných pauz v bezprostřední blízkosti, jež literatura pokládá za typičtější pro nerodilé mluvčí.

Dále teoretická část pojednává o dvou (komplikovanějších) dříve zpozorovaných rozdílech mezi rodilými a nerodilými mluvčími. Prvním z nich je údajná tendence nerodilých mluvčích produkovat relativně více vyplněných pauz před autosémantickými slovy, přičemž, jak už bylo řečeno, by tato tendence měla být způsobena obtížemi s vybavováním. Testování této tendence je třetí výzkumnou otázkou, přičemž je provedena detailnější slovnědruhová klasifikace než v předchozích výzkumech. Čtvrtá výzkumná otázka vyplývá ze zpozorované tendence nerodilých mluvčích produkovat více vyplněných pauz uvnitř promluv nebo vět. Této otázce je věnována samostatná sekce, poněvadž předchozí výzkumy volí různé, nejasně nadefinované jednotky, vzhledem k nimž je pozice vyplněných pauz posuzována, aniž by tuto volbu řádně vysvětlily. Tato samostatná sekce poukazuje na některé výrazné nedostatky těchto voleb a na příkladech demonstruje jejich problematičnost. Tak například jednotka ASU (Analysis of Speech Unit = jednotka analýzy řeči) je do velké míry definována arbitrárně a její definice nebere v potaz řadu syntaktických vlastností angličtiny, což při analýze

způsobuje množství problémů. Další z výzkumů jednotku, vzhledem k níž je pozice vyplněných pauz analyzována, dokonce ani nedefinuje.

Z těchto důvodů byla vytvořena experimentální metoda posuzování pozice vyplněných pauz: jejich relativní pozice je zkoumána vůči finitní klauzi a vůči souvětí. Tato volba je zdůvodněna objektivnější povahou finitních klauzí a souvětí při analýze. Následně jsou formulovány výzkumné otázky, tedy:

1) Produkují nerodilí mluvčí více vyplněných pauz než nerodilí?

2) Vykazují nerodilí mluvčí větší tendenci shlukovat vyplněné pauzy?

3) Existuje tendence produkovat více vyplněných pauz před (různými druhy) autosémantických slovy a je silnější u nerodilých mluvčích?

4) Produkují nerodilí mluvčí více vyplněných pauz uvnitř finitních klauzí a souvětí?

Následující část popisuje povahu dat, na nichž jsou tyto otázky zodpovídány. Data pocházejí z korpusu učitelské angličtiny ETC, jehož nahrávky sestávají z pěti částí (monolog, dialog, popis obrázků/příběhu, čtení, monolog v rodném jazyce). Pro účely této práce je použita pouze první část. Vzorek sestává z 20 mluvčích, 10 rodilých a 10 nerodilých, přičemž jednotliví mluvčí jsou vybrání náhodně. Tento vzorek obsahuje 448 vyplněných pauz, 182 pro rodilé mluvčí, 266 pro nerodilé mluvčí.

Dále je popsán proces zacházení s daty tak, aby mohly být výzkumné otázky zodpovězeny. Jde především o postup při počítání tokenů, slovnědruhovou klasifikaci pro účely třetí výzkumné otázky a zásady pro určování pozicionality vyplněných pauz. Celkový počet tokenů čítá 5475 tokenů pro rodilé mluvčí a 5640 tokenů pro nerodilé. Co se týče slovnědruhové klasifikace, slova, která následují vyplněné pauzy, jsou dělena na lexikální (autosémantické), gramatické (synsémantické) a ostatní (samotné vyplněné pauzy a discourse markers) jednotky. Uvnitř prvních dvou kategorií kategorií dochází k jemnějšímu členění slovních druhů, např. na plnovýznamová slovesa, respektive předložky, spojky apod.

V analytické části práce jsou postupně probírány jednotlivé výzkumné otázky. Tendence nerodilých mluvčích produkovat celkově větší množství vyplněných pauz je přítomná i v našem vzorku, v obou skupinách se však objevil vždy jeden mluvčí, jehož produkce vyplněných pauz (na 100 tokenů) byla výrazně nadprůměrná. Při zahrnutí těchto

mluvčích do výpočtů se tato tendence potvrdila, následně však byly provedeny dva další výpočty, přičemž pro první byl odstraněn nadprůměrně produkující nerodilý mluvčí a pro druhý též rodilý. První z výpočtů (tedy bez nadprůměrného nerodilého) už nevykazuje statistickou významnost, což poukazuje na zkreslování dat při inkluzi nadprůměrně produkujících mluvčích. Druhý výpočet (bez obou nadprůměrných mluvčích) opět vykazuje statisticky významný rozdíl, byť velmi hraniční.

Co se týče shlukování vyplněných pauz, vzorek jich obsahuje pouze 12 (2 pro rodilé, 10 pro nerodilé). To představuje statisticky významný rozdíl, 6 z těchto pauz však pochází od téhož nerodilého mluvčího – toho samého, který byl ve dvou předchozích výpočtech odstraněn. Odstranění shluků tohoto mluvčího již rozdíl nečiní statisticky významným. Činit závěry ohledně shlukování pauz tak není na našem vzorku možné.

V analýze pro třetí výzkumnou otázku se ukázalo, že obě skupiny nejběžněji produkují vyplněné pauzy před gramatickými jednotkami, což není vzhledem k jejich frekvenci překvapivé. Pro téměř všechny typy slov však platí, že nerodilí mluvčí před nimi produkují více slov. Přesto není jasné, zda existuje tendence produkovat pauzy spíše před lexikálními jednotkami, jelikož data poukazují i na působení jiných faktorů, jako je třeba syntaktické plánování. V tomto typu analýzy tak nelze činit závěry ohledně vlivu následujících jednotek na produkci vyplněných pauz.

V otázce syntaktické pozicionality se projevuje dříve popsaná tendence nerodilých mluvčích produkovat vyplněné pauzy častěji uvnitř syntaktických jednotek, oproti produkci na jejich hranách. Výsledky opět mohou být zkreslovány zahrnutím nadprůměrně produkujících mluvčích. Tato tendence se však nadále projevuje i po odstranění nadprůměrných mluvčích. Analýza tak podporuje předchozí závěr, že nerodilí mluvčí produkují více disfluencí uvnitř syntaktických jednotek, a nachází toto chování i u učitelů angličtiny. Významné rozdíly v ostatních pozicích nalezeny nebyly.

Práce následně tyto výsledky shrnuje a komentuje. Poukazuje na problémy s interpretací výsledků, zejména ohledně vlivu následujících slov na produkci vyplněných pauz a pozicionality pauz, jelikož segmentovat souvětí ve vzorku je obtížné díky vysoké frekvenci slov, která mohou být jak spojkami, tak discourse markery. Dále práce poukazuje na alternativní řešení předchozích výzkumů, která by mohla poskytnout lepší vhled do problematiky pozicionality. Upozorňuje také, že delší monolog může pro některé subjekty

představovat poměrně neobvyklý typ jazykové produkce, a tedy zde může být vyšší disfluence i z tohoto důvodu. Krátce je též probírán vliv extralingvistických faktorů na produkci vyplněných pauz, jako je nervozita či tlak.

V závěru práce naznačuje směr dalšího výzkumu disfluencí v učitelské angličtině na základě korpusu ETC, a to především na jeho ostatních částech, například na dialogových datech či na popisu obrázku.

8. References

AMBELE, E.A., BOONSUK, Y., and BUDDHARAT, Ch. (2019). Position of Hesitation Markers in Everyday, Informal Conversation in English. *ABAC Journal*, *39*(3), 129-140.

BERGMANN, C., SPRENGER, S. A., & SCHMID, M. S. (2015). The impact of language co-activation on L1 and L2 speech fluency. *Acta psychologica*, *161*, 25–35. https://doi.org/10.1016/j.actpsy.2015.07.015

CLARK, H. H., & FOX TREE, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, *84*(1), 73-111. https://doi.org/10.1016/S0010-0277(02)00017-3

DE JONG, Nivja H. (2016). Predicting pauses in L1 and L2 speech: the effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 113-132. https://doi.org/10.1515/iral-2016-9993

FOSTER, P., TONKYN, A., WIGGLESWORTH, G. (2000). Measuring spoken language: a unit for all reasons, *Applied Linguistics*, *21*(3), 354-375. https://doi.org/10.1093/applin/21.3.354

KOSMALA, L. (2021) A multimodal contrastive study of (dis)fluency across languages and settings: towards a multidimensional scale of inter-(dis)fluency. *Linguistics*. Paris: Université de la Sorbonne-Nouvelle.

RIAZANTSEVA, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*(4), 497-526. https://doi.org/10.1017/S027226310100403X

ROSE, R.L. (1998). *The communicative value of filled pauses in spontaneous speech*. MA Dissertation, University of Birmingham.

SCHMID, M. & FÄGERSTEN, K. (2010). Disfluency Markers in L1 Attrition. *Language Learning*. *60*(4), 753-791.

SKEHAN, P., and FOSTER, P. (2007). *Complexity, accuracy, fluency and lexis in task-based performance: a meta-analysis of the Ealing research.* in Van Daele, P.; Housen, S.; Kuiken, F.; Pierrard, M. and Vedder, I., (eds.) *Complexity, Fluency and Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussels: University of Brussels Press, 207-226.

WILLIAMS, S. (2022). *Disfluency and proficiency in second language speech production*. Cham: Palgrave Macmillan.

WILLIAMS, S.A., & KORKO, M. (2019). Pause behavior within reformulations and the proficiency level of second language learners of English. *Applied Psycholinguistic*, *40*, 723-742.

9. Sources

ANTHONY, L. (2014). AntConc (Version 3.4.4) [Computer Software]. Tokyo: Waseda University. Available from https://www.laurenceanthony.net/software

SCHMIDT, T., and WÖRNER, K. (2014), "EXMARaLDA", in *Handbook on Corpus Phonology*, 402-419. Oxford: Oxford University Press. https://www.exmaralda.org/

10. Appendices

10.1 Instructions for Task 1

The English Teacher Corpus
Task 1

Please choose one of the following topics. You can take a 2–3 minutes to think about what you are going to say but please do not take notes. You will be expected to talk about the topic for 3-5 minutes without being given any prompts or questions. After this time, we will switch into an informal conversation.

Topic 1:
Why did you decide to become an English teacher? Are you still happy with the choice? Have you ever thought about doing something else?

Topic 2:
Have you changed as a teacher in the course of your career. How? And why?

Topic 3:
Do you remember any critical incidents in your pedagogical career which had an impact on you as a teacher? What happened and how did it affect you? What did you learn?

10.2 Example of raw data from Task 1 (ETC: speaker EN004)

B: yes okay so I start okay so why did you dec= why did you decide to become a teacher (er) I didn't decide to become a teacher I kind of like no yeah I didn't decide to become a teacher I ended up becoming a teacher because it's one of the only options that would allow me to live abroad and also (er) stay in Czech for as long as I wanted to and I think after that I started to become an actual English teacher because I've seen that (er) it's something that I actually do enjoy and I think the main thing is I teach children so I'm I enjoy being around children and I think if I didn't enjoy being around children I wouldn't enjoy teaching aspect of it (er) and I would say yes so I'm happy in the position that I that I am a teacher I enjoy doing it but I would say as something that I would do for long term (er) that's hard to decide because of course (eh) we know teaching doesn't provide a lot of money and if you wanna have a family and do certain type of things I feel like (eh) traditional teaching isn't the best way to go (eh) in my case because since I'm an expat and it kind of limits me on how much I can make and how much time I would have to do other things besides just teaching and have I ever thought about doing something else (er) yes I have thought about doing many other things cause my background is in (er) logistics and (er) database management so: I<?> have this in the back of my head I do enjoy like I T and like (er) operations so I would love to get back in that (er) what I think I will <X> push and see if I can do more with teaching maybe go online or just expand it in a way where it can be something that I can do for the long term and the money makes sense and I can also incorporate operations and other logistics and stuff that I'd studied for when I was in university so it can all come together so I can just yeah have this fulfilment of work-life balance and yeah so I think that is it