



FILOZOFICKÁ FAKULTA  
UNIVERZITY KARLOVY  
V PRAZE

**ÚSTAV ANGLICKÉHO JAZYKA A DIDAKTIKY**

DIPLOMOVÁ PRÁCE

Bc. Tereza Štěpánová

*"I I erm I thought": Vybrané prvky řečového managementu českých pokročilých mluvčích angličtiny ve srovnání s mluvčími rodilými*

*"I I erm I thought": Selected performance phenomena of Czech advanced speakers of English in comparison with the native speaker norm*

Praha, srpen 2015

Vedoucí diplomové práce (supervisor): PhDr. Tomáš Gráf

## Declaration

*Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala 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 diplomové práce ke studijním účelům.  
I have no objections to the MA thesis being borrowed and used for study purposes.*

*V Praze dne*

*podpis*

## **Abstract**

The present thesis studies two types of performance phenomena, namely filled pauses and repeats. Both naturally occur in spoken discourse and do not hinder communication, their main function is to decrease the effects of online planning pressure. Since a speaker has to plan and produce language at the same time, non-native speakers are expected to experience greater difficulties when speaking a foreign language. The main aim of the present paper is to analyze English native speakers and Czech advanced speakers of English in terms of how their usage of these two phenomena differ. Attention is paid to overall frequencies as well as to their distribution, alternatively variation, and further performance phenomena in their proximity. The analysis comprises 3 non-native and 3 native speakers which were selected from the multilingual learner corpus LINDSEI and the monolingual corpus LOCNEC respectively.

**KEY WORDS:** spoken language, performance phenomena, repeats, filled pauses, learner corpus

## **Abstrakt**

Předkládaná diplomová práce se zabývá dvěma typy prvků řečového managementu, a to vyplněnými pauzami a opakováními. Oba tyto prvky se přirozeně vyskytují v mluvené řeči a nijak nenarušují porozumění. Jejich hlavní funkcí je zmírnění tlaku, který je na mluvčí při plánování jejich promluvy vyvíjen. Vzhledem k tomu, že daný mluvčí plánuje a produkuje řeč v témže okamžiku, předpokládá se, že nerodilí mluvčí se budou potkávat s většími problémy než mluvčí rodilí. Hlavním cílem této práce je analyzovat promluvy anglických rodilých mluvčích a českých pokročilých mluvčích angličtiny, a to vzhledem k tomu, jak se používání těchto dvou prvků liší v jejich promluvě. Pozornost je věnována celkové četnosti výskytu, ale také jejich distribuci, popřípadě jejich variaci a dalším prvkům řečového

managementu v jejich blízkosti. Pro potřeby analýzy byli vybráni 3 nerodilí mluvčí multijazykového žákovského korpusu LINDSEI a 3 rodilí mluvčí korpusu LOCNEC.

**Klíčová slova:** mluvený jazyk, prvky řečového managementu, opakování, vyplněné pauzy, žákovský korpus

## **Acknowledgment**

I would like to thank my MA thesis supervisor PhDr. Tomáš Gráf for his exceptional leadership, assistance and words of encouragement as well as the whole Department of English Language and ELT Methodology whose members will remain great mentors to me.

## **Table of contents**

Declaration	2
Abstract	3
Abstrakt	3
Acknowledgment	4
List of tables	8
List of figures	8
List of abbreviations	8
1 Introduction	9
2 Theoretical background	12
2.1 Spoken language	12
2.1.1. Features of spoken language	13
2.2 Finding a framework for performance phenomena	17
2.3 Performance phenomena in native and non-native speech	20
2.3.1 Repeats in native and non-native speech	23
2.3.2 Filled pauses in native and non-native speech	30
2.3.3. Concluding remarks	35
3 Material and method	37
3.1 Material	37
3.2 Method	39
3.2.1 Pilot study	41
3.2.1 Hypotheses	45
4 Analysis	48
4.1. Overall frequencies	49
4.1.1. Filled pauses	49
4.1.2. Repeats	50
4.2 Further performance phenomena	51
4.3 The distribution of filled pauses	56
4.4. Repeated word	59
4.3.1. Two-word repeats	62
5 Discussion of the findings	65
5.1 Czech advanced speakers in comparison with other L2 speakers	65
5.2 Implications for teaching	69
6 Conclusion	72
7 References	75
8 Sources	77

9	Resumé .....	78
10	Appendix.....	82

## List of tables

Table 1 - Filled pauses as used by non-native and native speakers.....	46
Table 2 - Repeats as used by non-native and native speakers.....	47
Table 3 - Filled pauses in close proximity of other performance phenomena.....	48
Table 4 - Repeats in close proximity of other performance phenomena.....	50
Table 5 - FPWCL and FPWCON in LINDSEI_CZ and LOCNEC.....	52
Table 6 - Position of a FP within FPWCL.....	53
Table 7 - Position of a FP within FPWCON.....	53
Table 8 - Repeated word in LINDSEI_CZ and LOCNEC .....	55
Table 9 - Repeats of function words in LINDSEI_CZ and LOCNEC .....	56
Table 10 - Combinations of word classes occurring in two-word repeats.....	58

## List of figures

Figure 1 - Framework of Fluency as defined by Götz (2013).....	17
Figure 2 - Clark and Wasow's Theory (1998).....	26
Figure 3 - Clark and Fox Tree's Filler-as-word hypothesis (2002).....	30

## List of abbreviations

*BNC* – British National Corpus  
*CIC* – The Cambridge International Corpus  
*COBUILD* – Collins Birmingham University International Language Database  
*CZ* followed by a number (e.g. CZ001) – numeric speaker identification in LINDSEI\_CZ  
*EN* followed by a number (e.g. EN001) – numeric speaker identification in LOCNEC  
*FP* - filled pause  
*ICLE* - The International Corpus of Learner English  
*L1* - first language, mother tongue  
*L2* - second/third language  
*LGSWE* – Longman Grammar of Spoken and Written English  
*LOCNEC* – The Louvain Corpus of Native English Conversation  
*LINDSEI* - the Louvain International Database of Spoken English Interlanguage  
*LINDSEI\_CZ* – The Czech component of the Louvain International Database of Spoken English Interlanguage

## 1 Introduction

In the recent decades, the interest in learner language has substantially increased as the number of speakers who use English as either their foreign or second language is rising steadily. Non-native speakers of English are estimated to outnumber native ones in ratio of at least 3:1 and this trend is expected to continue. The fact that English has become a global language can be undoubtedly attributed to the millions of students striving to learn the modern lingua franca every day. Including these or any other learners of a foreign or second language, the study of learner language particularly focuses on either the written or spoken production of these learners and the research has implications primarily for second language acquisition and foreign language teaching. With the employment of computer-based technology in linguistic research, analyzing learner language has been given a much needed boost. Since the 1980s large corpora specializing in non-native speaker data have been built and new ones are constantly appearing.

Initially, written language was the focus of attention as written corpora proved easier to assemble. However, negligence of the spoken discourse has been acutely felt and therefore, in the recent years, spoken learner corpora have also gained in size and scope. Nevertheless, designing and collecting data for a spoken corpus still represents an obstacle which modern technology helps to mitigate only to a certain degree as the compilation largely depends on the laborious task of manual transcription.

The present thesis aims to contribute to the analysis of spoken discourse and chooses to study selected performance phenomena, particularly focusing on repetitions and filled pauses. Formerly, linguists have worked with the assumption that these features hinder communication and that they represent disruptive dysfluency elements. However, recent studies suggest that repetitions and filled pauses occur naturally in spoken discourse and do not pose a challenge to understanding. Their main function is to ensure fluency of speech and they also help to minimize the effects of planning pressure since the speaker needs to plan and produce language at the same time.

One of the main aims of the present thesis is the comparison of usage of both studied phenomena in a corpus of native and Czech non-native speakers of English. Language communities, whose mother tongue is not represented by one of the major languages such as German or French, have yet not been given sufficient attention, especially those from the Slavic family, and the present thesis aims to shed some light on the use of English of these speakers. For the purposes of the present paper, they are represented by advanced learners; as such they have to deal with extensively more difficult operations since their planning involves all language dimensions and happens on more conscious level in comparison with native speakers. Therefore, a markedly different pattern of use can be expected. At the same time advanced learners usually consider native speakers as their natural models, which leads to the second aim of the present paper: differences in the pattern of use of the repetitions and filled pauses are to be used as the source for possible enhancement in the realm of teaching advanced speakers.

Drawing on grammars and studies based on corpus research, the theoretical part (see section 2) generally discusses the issue of spoken language including its features and presents an extensive description of repetitions and filled pauses within the framework of performance

phenomena (Biber et al., 1999). In addition, when relevant it gives an overview of findings of different use of the concerned phenomena in the native and non-native discourse.

The empirical part of the present thesis provides an analysis of 3 selected native and 3 non-native speakers. Repetitions and filled pauses are analyzed with respect to their distribution and frequency in spoken discourse, additionally, in the case of repetitions, variations of the particular language means are described. The data were obtained from the multi-lingual corpus LINDSEI, particularly its Czech component, and from the accompanying native speaker corpus LOCNEC. The specific handling of the data which was adopted for the purposes of the present thesis is described in detail in the methodological part (see section 3).

Findings of the above outlined analysis are reported in section 4, stating results for both native and non-native speakers. These are assessed and compared in the subsequent discussion (see section 5.1) and moreover, possible outcomes for language teaching of advanced learners are suggested (see section 5.2). A comprehensive summary taking into account the findings of both the theoretical and the empirical part is provided at the end of the present paper. Finally, an appendix including the complete set of analyzed data equipped with suitable coding is enclosed.

## **2 Theoretical background**

### ***2.1 Spoken language***

Until recently, studies aiming to comprehensively describe the English language has undoubtedly shown a bias towards written discourse. Only a small proportion of spoken language, such as pronunciation or prosody, has been subjected to analysis as these features require only limited amount of data. With the technological improvement in the past years, necessary tools needed for the study of spoken language, such as audio recording devices or voice recognition programmes, have reached such level of advancement that allowed to study spoken language on much larger scale. Since the 1990s, various spoken corpora, such as BNC, COBUILD or CIC, have been built and hereby linguists have been equipped with a resourceful tool for the study of spoken language.

Nevertheless, although supported by technology, the study of spoken discourse cannot dispense with its written form: the recording of a particular illustration of spoken word needs to be accompanied by a transcription, which is primarily used for the purposes of analysis itself. As Biber et al. (1999: 1042) aptly point out, transcriptions are highly conventionalized

models of speech since they are subjected to many of the practices employed in writing, e.g. marking of spaces between words or their spelling. In connection to the phenomena analyzed in the present thesis, we may wish to highlight the common practice of transcribing filled pauses either as *er*, *erm* or *uh*, *um* even though the actual pronunciation of these sets does not differ.<sup>1</sup>

Despite the fact that spoken discourse may have various forms, common everyday conversation has been selected as the most exemplary and frequent mode of expression. For the purposes of describing its features two comprehensive grammars, both based on corpus research, will be used in the following section. Primarily, we will draw upon the findings of Biber et al. (1999) which they ascertained from the *Longman Spoken and Written English Corpus*, and additionally, we will make use of Carter and McCarthy's grammar (2006) based on *Cambridge International Corpus*.

### ***2.1.1. Features of spoken language***

Conversation, or more generally spoken language, typically takes place in real time and does not allow any preparation or planning beforehand. It is characteristically interactive, i.e. involves at least two users, and happens face to face. These users share a common context which includes the immediate social and interpersonal situation furthermore determining the nature of their information exchange. Hence the spoken discourse shows a higher degree of grammatical reduction in the form of deictic expressions, situational ellipsis and manifests little elaboration since "greater precision would not only be superfluous, but it would also need more processing and delay the ongoing dynamic of the conversation" (Biber et al. 1999:

---

<sup>1</sup> The difference may lie only in the dialect - Biber et al.(1999: 1053) claim using the former set for British English whereas the latter one for American English.

1045). The dynamic character is also demonstrated by rather short communicative units where hesitation phenomena<sup>2</sup> such as repetitions, filled or unfilled pauses as well as self-corrections may occur since there is little time for planning and a continuous flow of speech needs to be ensured. The dynamism of spoken discourse has also inspired other authors to express its essential feature in a more metaphorical way: Halliday (1990: 81) compares it to a film, where the message is presented to you as a sequence of images which you have to conceive in its entirety. Carter and McCarthy (2006: 169-170) chooses a typically British metaphor when they characterize spoken language “as being more like the strung-together coaches of a train or links of chain rather than carefully constructed hierarchy of embedded structure.”

Given the presupposition that conversation happens in real time, it is a highly difficult procedural activity which to a large extent relies on immediate decisions of a particular speaker thus directly affecting the way it is constructed. Spoken language therefore shows a marked difference to the written discourse, where sentences are presented as a finished and carefully constructed product. In this context, Biber et al. (1999: 1066–1067) proposes 3 basic principles of online production. The first one called “keep talking” basically expresses the speaker’s intention to keep the conversation flowing using, if needed, various compensating strategies in order to avoid dysfluency. The second principle “limited planning ahead” involves the rather constrained time a speaker has to construct his or her message. Therefore, the content of spoken language shows little elaboration and is characteristic of having information with higher communicative dynamism towards the end. Last but not least, the third principle “qualification of what has been said” points to the pervasive need of the speaker to return or reformulate a message he or she mentioned earlier. Although it may not

---

<sup>2</sup> Hesitation phenomena will be discussed in detail later on as they represent the main focus of the present thesis.

logically fit into the sequence, the speaker may wish to express some afterthoughts or modify the utterance in order to further elaborate the message.

All of these principles clearly point to the problematic definition of a basic unit in spoken discourse as well as of its distinctive sub-units which help to organize the information structure both for the speaker and the listener. The following subsection presents an overview of various attempts aiming to find an equivalent term to the *sentence* of written discourse.

### ***2.1.1.1. Defining a basic unit in spoken language***

Looking at the authors studying spoken language in the years before corpus linguistics has reached its peak, Hunt's conception of the so-called *T-Unit* as the basic unit for spoken language was the most frequent choice.<sup>3</sup> In his definition, Hunt (1965: 37) highlights the inconvenience of the term *sentence* and proposes to use the word *unit* instead, yet his concept heavily relies on written discourse traditions: "These units might be christened 'minimal terminable units,' since they would be minimal as to length, and each would be grammatically capable of being terminated with a capital letter and a period." His choice of parsing spoken word is also determined by syntax as he has "defined [it] as one main clause plus any subordinate clause or non-clausal structure that is attached to or embedded in it." (Hunt, 1970: 4) Nevertheless, his definition is primarily based on written language and hence is only partially applicable to spoken discourse.

Taking into account Hunt's concept, Biber et al. (1999: 1069) comes with a similar unit, in their terms 'the maximal unit of conversational syntax'. Its most important kind - the clausal unit<sup>4</sup> "is a structure consisting of an independent clause together with any dependent

---

<sup>3</sup> based on Foster, Tonkyn and Wigglesworth's article (2000: 360)

<sup>4</sup>Their conceptualization mentions dysfluency phenomena yet removes them from the segmentation of a speech sequence.

clauses embedded within it” (Biber et al., 1999: *ibid.*). Furthermore, they admit that spoken discourse cannot be solely segmented into *clausal units* and add *non-clausal units*<sup>5</sup> into their system. In short, non-clausal units are all remaining structures that are neither clausal nor embedded into clausal units. Lastly, they separate coordinated structures into two individual types since a clause in speech may start with a coordinator and thus lack the first clause or “coordinated units need not be of the same type: a clausal unit can be coordinated with a non-clausal unit” (Biber et al., 1999: 1070).

The last point of view of the structure of spoken language follows the syntactic tradition of the previous two approaches, yet puts more emphasis on the importance of prosodic phenomena such as intonation and pauses and furthermore specifies the nature of sub-clausal units.<sup>6</sup> Foster, Tonkyn and Wigglesworth’s *Analysis of Speech Unit (AS-unit henceforth)* may consist of either an independent clause (minimally containing one finite verb) or “a sub-clausal unit, together with any subordinate clause(s) associated with either” (2000: 365). Sub-clausal unit may be an ellipted fragment of a clause fully recoverable from the context (based either on the discourse or the situation) or it may be an ‘irregular sentence’ as described in Quirk et al. (1985: 838-53). Subordinate clause needs to have a finite verb or a non-finite verb followed by at least one other clause element. Intonation and pause pattern play an important role in deciding where to put an end to the unit. A coordinated clause or a topicalized noun phrase will be treated as a separate AS-unit if there is falling or rising intonation and the clause element is followed by a pause of 0.5 second or longer. Dysfluency phenomena are also taken into account - they are transcribed and graphically

---

<sup>5</sup> Biber et al. (1999) uses the umbrella term C-units for clausal and non-clausal units.

<sup>6</sup> The term sub-clausal unit is in accordance with Biber et al.’s non-clausal unit.

separated by brackets.<sup>7</sup> It depends only on the focus of research if they will be included (and to what extent), nevertheless they are considered important features of spoken discourse.

In general, we may say that the prevailing treatment of spoken data seems to take the syntactic point of view, however the growing interests of spoken language features and rising refusal of written discourse convention contribute to a more detailed definition of the basic unit. The specific treatment of spoken data for the purposes of the present thesis will be described in the methodological chapter (see section 3).

## ***2.2 Finding a framework for performance phenomena***

Returning to the first of the three principles of online communication, i.e. what Biber et al. (1999: 1066) call ‘keep talking’, the foremost aim of a speaker is to maintain the course of communication and express his or her thoughts as precisely as possible given only a limited amount of time. If a speaker is successful in keeping the floor and getting his/her message across he is considered *fluent*.<sup>8</sup> In the ELT context, this term is generally used in order to describe speakers’ oral proficiency. Such a stance is accordingly projected in the *Common European Framework of Reference* (CEFR henceforth), which uses qualities such as ‘fluent’, ‘spontaneous’ and ‘effortless’ for the description of spoken performance. Taking into account the most proficient levels as represented in our sample, a proficient speaker “[c]an express him/herself spontaneously at length with a natural colloquial flow, avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.” (CEFR, 2001: 28)

---

<sup>7</sup> AS-unit boundaries are signaled by an upright slash and clause boundaries by a double colon.

<sup>8</sup> Chambers (1997: 536) aptly illustrates this point of view claiming that “[w]hen we say ‘She speaks French fluently’ in ordinary conversation, we mean that a person has a good command of the language and uses it with ease and efficiency.”

Although the above mentioned characterization regards fluency as an important skill, it conceives it from a rather global point of view based predominantly on the listener's impression. In contrast to this perception of fluency in the *broad sense*<sup>9</sup>, which takes the native speaker as a point of reference, Lennon (1990:389-390) describes fluency in the *narrow sense* as “presumably isolatable component of oral proficiency.” As such, it can be measured and thus allows a more objective judgement about one's oral proficiency. Additionally, Chambers (1997: 536) presents another definition which equals fluency with communicative competence. In the communicative language teaching, fluency consists in “effectiveness of language use within the constraints of limited linguistic knowledge.” As opposed to the CEFR, fluency is not necessarily bound to a particular level of proficiency and depends solely on the abilities of a speaker.

Despite the fact that different definitions of fluency coexist and have been used in various researches, a comprehensive framework which would take into account all individual factors that contribute fluency has been lacking in literature. Recently, Götz (2013: 9) in her book focusing specifically on this issue presents a framework that incorporates both the speaker's and the listener's point of view and operates with an abstract unit of fluency: termed *fluenceme*, it represents “[an] idealized feature of speech that contributes to the production or perception of fluency, whatever its concrete realization may be. Fluencemes build the basis

---

<sup>9</sup> Later in his work, he reformulates the term and uses ‘high order fluency’ instead (Lennon, 2000: 25)

for the distinction between productive, perceptive and nonverbal fluency.” Fluencemes and a

model speaker<sup>10</sup> then serve to build a framework for fluency (see Fig. 1).



**Fig. 1** Framework of Fluency as defined by Götz (2013)

In her concept of fluency, Götz distinguishes between features that can be objectively measured, the so-called *productive fluency variables*, and *perceptive fluency variables* which include features influencing the overall impression of a speaker that a listener may have. Finally, nonverbal fluency variables are also taken into consideration. As opposed to the second and third category of fluencemes, which involve various notions that a listener perceives and evaluates, the first category puts emphasis on the speaker since it “subsumes all the variables that are connected to speech production and the accompanying results of the

---

<sup>10</sup> In her wording “an idealized and prototypical (native) speaker” (Götz 2013: 8), which shows her rather cautious use of the term native speaker as it presents difficulties similar to defining the term fluency.

cognitive planning pressure that comes about while speaking” (Götz, 2013: 8). Hence, this group encompasses temporal variables, such as speech rate, mean length of runs, unfilled pauses and phonation time ratio; formulaic sequences which a speaker stores as prefabricated units and readily uses in communication in order to decrease the planning pressure; and finally, fulfilling a similar function as the last mentioned variable, fluency-enhancing strategies<sup>11</sup> which cover the use of speech management phenomena, discourse markers and small words.

As the present thesis strives to objectively measure and compare the use of selected performance phenomena of native and non-native speakers, Götz’s framework, which follows the same pattern, has been chosen as a point of reference.

### ***2.3 Performance phenomena in native and non-native speech***

In the previous section, we have somewhat readily used the term performance phenomena, yet the question remains what the term actually covers. Scholars have been rather divided on this topic - various terms including *repair fluency* (Foster and Tavakoli, 2009) *hesitations* (Biber et al., 1999) and others differs not only in dissimilar wordings but also in scope - some comprise more features than the others. Götz (2013) herself uses the term *speech management strategies* referring to a terminology oriented article by Rühlemann (2006), which adverts to the persisting use of writing-based terms for the description of spoken discourse. He argues that the inappropriate use of predominantly negative terms such as *dysfluency*, *error* or *repair* produces a contradicting notion: on one hand claiming that

---

<sup>11</sup> The present thesis chooses to use the term performance phenomena instead of fluency-enhancing strategies. Götz herself tends to use these terms interchangeably (see Götz, 2013: 32)

these features represent natural component of speech, yet on the other choosing a label that still has an undertone of inappropriateness.

Nevertheless, the prevailing consensus seems to be that due to the spontaneous nature of speech which allows only limited planning ahead and bounds the speaker to produce a continuous flow of conversation, performance phenomena have its rightful place in spoken discourse. As Götz (2013: 33) points out, contemporary research considers this assumption to be its premise and focuses on further study concerning the frequency and distribution of performance phenomena: “This approach describes the range and number of their occurrence as well as the preferred positions in which native speakers tend to use these speech management strategies in their utterances.” Furthermore, the research on non-native speakers anticipates an increased proportion of performance phenomena as they experience a higher planning pressure when speaking a foreign language. Götz (2013: 34) also alerts to the importance of detailed study of their frequency and distribution as well since non-native speakers have gained the reputation of sounding “bookish” or “pedantic” (Cutting, 2006: 172) in comparison to native speakers and the use of performance phenomena may also be a factor in distinguishing an L1 and L2 speaker.

For the purposes of an overview of performance phenomena, Biber et al.’s LGSWE<sup>12</sup> has been chosen as it covers the widest range of all possible strategies that might be used in order to cope with planning pressure. Apart from (un)filled pauses and repetitions, which will be dealt with in detail later on, Biber et al.’s grammar (1999: 1052–1066) mentions the so-called “retrace-and-repair” sequences, several types of incompleteness and syntactic blends.

---

<sup>12</sup> Cf. Foster and Tavakoli (2009) who include only repeats, (un)filled pauses and self-corrections (retrace-and-repair sequences in LGSWE terminology).

First, let us have a look at the first mentioned phenomenon: as a result of the constant pressure a speaker is exposed to, he may reconsider what he has just said (cf. Biber et al.'s constructional principle "qualification of what has been said") either because he has made a grammatical mistake, an option more probable for a non-native speaker, or because an inappropriate word has been used. In such a situation "the speaker retraces (or notionally 'erases') what has just been said, and starts again, this time with a different word or sequence of words" as in the following example (Biber et al., 1999: 1084)<sup>13</sup>:

- (1) Your face was down instead of up and so that's why **the**, [all my] labor was in my back 'cos you were pushing the wrong way **and**, uh, [but] luckily at the very end, just before I went into delivery, you flipped around somehow <...>

Retrace-and-repairs sequences may be furthermore accompanied by other performance phenomena as evidenced by the filled pause (underlined) following shortly the word the speaker decided to change. However, these self-corrections, as they are also termed (see Foster and Tavakoli, 2009; Götz 2013), show speaker's ability to become aware of his mistakes and as some studies prove (e.g. Lennon 1990 or Riggensbach 1991) they do not seem to pose a threat to speaker's fluency measures.

The next performance phenomenon in fact draws upon the above mentioned, the only difference being scope or intervention from another party. According to Biber et al. (1999: 1063–4), situations when a speaker leaves his utterance incomplete may follow the ensuing scenarios: a) the speaker abandons his utterance and starts anew, b) speaker's utterance is disrupted by another speaker or event, c) speaker's utterance is completed by another speaker, i.e. hearer of the conversation, and finally d) the speaker completely leaves his utterance

---

<sup>13</sup> In relation to other wordings used for retrace-and-repair sequences, Biber et al. (1999) criticizes using the term 'false start' completely in accordance with Rühlemann's argument (2006). On the other hand, Biber et al. continues to use the term 'dysfluency' and 'error' in his analysis of spoken conversation.

unfinished without attempting to reformulate it. The reasons for such actions might originate from the subject matter or might be conditioned by the extent of hearer's attention. Additionally, in the case of a non-native speaker the reason might consist in the lack of knowledge of the required word, perhaps combined with incomprehension of the hearer who would be otherwise able to finish his utterance. The last mentioned phenomenon can be perhaps attributed to the limits of short term memory: a syntactic blend is "a sentence or clause which finishes up in a way that is syntactically inconsistent with the way it began" (Biber et al., 1999: 1065) as demonstrated in the example:

(2) **You're** talking about a week and a half or something **aren't we**?

In conclusion, a speaker has a handful of strategies which can mitigate the planning pressure of online communication at his disposal, regardless of whether he uses them consciously or unconsciously, and applies them according to the particular situation of speaking and his language abilities. The following two sections take repeats and filled pause under scrutiny as they represent the core of the present thesis and thus deserve a more detailed analysis.

### ***2.3.1 Repeats in native and non-native speech***

First of all, it should be emphasized that repeats are to be distinguished from repetitions: the former being a feature of spoken communication used in order to reduce planning pressure, whereas the latter represents a purposeful reiteration of a particular word, phrase or a sequence of words for emphatic rhetorical effect (cf. Biber et al., 1999 or Foster, Tonkyn and Wigglesworth, 2000).

In general, repeats account for a strategy a speaker may use in order to gain time, as Carter and McCarthy (2006: 173) put it, the speaker is trying to “buy more time for thought.” In doing this, the speaker uses the same word or a combination of words repeatedly until he is finished with planning and able to carry on with the conversation. According to the findings of Biber et al.’s study for the LGSWE (1999: 1055), the most common scenario is the reiteration of one word and subsequently, the likelihood of multiple repeats decreases markedly with the number of words to be repeated: “[t]here are extremely few instances of three or more repeats.” The number reduces even more if we take into account two-word or three-word combinations (e.g. *I’ll* or *it was*).

In relation to what kind of words tend to be repeated, Biber et al.’s analysis (as well as others, see Clark and Wasow 1998, Götz 2013) shows undeniable preference for the function words with personal pronouns, possessive determiners, articles and conjunctions leading the list. The reason for such tendency seems to be twofold: on one hand, they rank the highest in the overall frequency of words used in the English language and on the other, there are grammatical reasons which require function words to be mandatorily expressed in the English sentence. Moreover, it is not coincidental that function words have the highest repeat index as they occur at a beginning of a clause or a noun phrase where the planning pressure reaches its top and the speaker still needs time to formulate his utterance. In such a situation, “it is easy for the speaker to utter a very frequent word, without having a clear plan for what words will follow it” (Biber et al., 1999: 1059). This hypothesis seems to be furthermore strengthened by the low repeat index of accusative personal pronouns which occur towards the end of syntactical units: in the case of the sequence of *you you*, Biber et al.’s research (1999: *ibid.*) has not found any “plausible instance of *you you* being a dysfluent repeat in object or prepositional complement position.”

As regards function words which have the inclination to be repeated, the personal pronouns, and especially the first person pronoun *I*, show the highest frequency, a finding easily justifiable by their almost invariable position at the beginning of a clause, where the speaker struggles with planning the most. The same accounts for conjunctions such as *if*, *and* or *when* which are to be found in a similar clause position as personal pronouns and hence represent further means of relieving planning pressure. Based on the findings of LGSWE (1999: 1059), possessive pronouns such as *my*, *your* and articles (*the*, *a*) show a similar tendency for frequency and distribution, “suggesting that the beginning of a full noun phrase (containing a noun head, a preceding determiner, and optional modifiers) is also a major point of planning pressure.” Nonetheless, the indefinite article *a* seems to be less frequent than the definite one and the prevocalic *an* is even less frequent. The explanation LGSWE proposes is that the prevocalic *an* requires the speaker to perform the selecting phase beforehand resulting in higher planning pressure. The data suggest that the speaker usually selects the ‘default form’, i.e. the indefinite article *a*, and only after that, when he has decided for a particular word, he opts for the grammatically correct form (cf. I would like **a**, **an** egg and a whole wheat English muffin.).

In contrast, prepositions and verbs represent the tokens with a rather low repeat index. Being function words, prepositions would be expected to rank alongside the above mentioned groups as they are frequent constituents and tend to introduce major syntactic units, i.e. prepositional phrase, however the data showed opposing tendency. The explanation might lie in the fact that prepositions are determined by their head noun and thus the choice might be lexical-oriented. Taking a look at the list of the last major category that shows relatively high incidence of repeats, i.e. verbs, we may observe that apart from the copula or auxiliary

*is*<sup>14</sup>, others do not seem to be a point of planning pressure as their recurrence is comparatively low. On that note Biber et al. (1999: 1060) suggest that the major planning pressure seems to lie on the initial noun phrase rather than on the verb phrase.

In connection to verbal repeats, LGSWE also includes a section on repeats with verb contractions. As has been mentioned in the overall characteristics of spoken discourse, the language used tends to be extremely economical in order to ensure constant flow of conversation and shows a great degree of grammatical reduction. Hence, the combination of a personal pronoun and an enclitic verb contraction ranks the highest on the frequency lists. Consequently, combined with their initial position in a clause, it also raises the likelihood of their repetition. Biber et al.'s research only confirms that in claiming "indeed, *it's*, *it'll* and *they're* are proportionally more common in repeats than any of the single words in Table 14.5.<sup>15</sup>" (1999: 1061). Based on their frequency analysis, combinations with the enclitic form of the verb to be (e.g. *I'm*, *you're*, *it's*<sup>16</sup>) are by far the most frequent with occurrence over 20,000 times per million words, followed by verb contractions with auxiliary *will* (*I'll*), auxiliary *have* (*you've*), and the enclitic *'d*, being either a contraction of *had* or *would*, with the occurrence of 1,000 times per million words.

If we take into account non-native speakers, first we may wish to emphasize that the planning pressure is considerably raised by the complexity of language levels a learner need to attend to when formulating his utterance. Apart from the actual content, the learner points his attention to syntactic, stylistic and lexical choices therefore the locus for repetitions does

---

<sup>14</sup> "The repeat of *is* appears to be particularly prevalent when the subject preceding *is* is a full noun phrase, particularly an utterance-launching fixed expression (or lexical bundle) such as ***the problem is***, and the predicative following it is a fairly heavy constituent" (Biber et al., 1999: 1060).

<sup>15</sup> Table 14.5. include all of the above mentioned groups, i.e. functional words and verbs.

<sup>16</sup> It includes contractions of the auxiliary *have* in third person singular.

not need to be only at the beginning of a clause (as it is with native speakers) but might be found further on in the clause. A pilot study on German advanced learners of the LINDSEI corpus (Götz, 2007) indeed shows that only 6% of the non-native speakers use repeats in a native-like way, i.e. a clause commenced with a repeat exhibits no other performance phenomena and continues fluently. Rather, the majority of learners tends to combine various performance phenomena together with repeats both at the beginning and later on in the clause, suggesting that an initial repeat is not enough to gain time for planning (Götz, 2007: 75)<sup>17</sup>.

A closer look at how non-native speakers use repeats has been carried out by the same author a few years later (Götz, 2013), which to a certain extent vindicates the previous findings. On the whole, non-native speakers show a significant underuse of repeats, yet the distribution among the investigated groups<sup>18</sup> of words reveals further particularities. Regarding the personal and possessive pronouns, German advanced learners of English show no deviation from the native speaker norm and repeat them actively at the beginning of a clause. On the contrary, Götz's findings (2013: 107) have revealed significant underuse of articles and determiners (excluding possessive pronouns). As she puts it, it "may be explained by the learners' need for a somewhat 'stronger' strategy in these sentence positions before a noun to give them more time to think about the correct lexical item, like, for example, a filled pause" (ibid: 108). Furthermore, English articles may still represent a problematic grammatical point even for higher level learners and as there is a correlation between frequency of a word and its repeat index, such an aspect might contribute to the

---

<sup>17</sup> Es scheint so zu sein, dass für die Lerner der Planungsdruck nicht unbedingt nach dem Beginn der Aussage nachlässt [...], sondern auch während des Weitersprechens konstant hoch bleibt. Folglich genügt eine Wiederholung am Anfang einer Phrase den Lernern nicht dazu, den Rest der Aussage zu Ende zu planen.

<sup>18</sup> Götz adopts the same groups as mentioned in LGSWE.

underrepresentation of articles in LINDSEI-GE. Additionally, conjunctions and verbal contractions have also been found to be underrepresented in the learner corpus. Such finding correlates with the comparatively low frequency of these function words as such. On this note, Götz (2013: 108) points to the undeniable fact that non-native speakers seldom use verb contractions and moreover remarks that “[f]or conjunctions, the same observation can be made, namely the learners use far fewer conjunctions than the native speakers and prefer shorter and less complex sentences without many ‘extensions’ to them.” Thus, the low frequency of the features is reflected in the low repeat index which the data only confirm. Corroborating the pilot study observations, LINDSEI-GE have also been found to overuse repeats of prepositions and verbs. Due to their prevailingly non-initial position in the clause, their higher occurrence in repeats might signal non-native speakers’ need for further planning even after having initiated their utterance.

One last issue that needs to be addressed is the question whether repeats account for a strategy that is employed intentionally or whether the speaker does not plan them at all. Being advocates of the latter approach, Biber et al. (1999: 1056) claim that “[r]epeats, as a form of dysfluency, are presumed to be unplanned or involuntary.” This assumption is presented in contrast with occurrences of verbal repetition which is used for the purposes of intensification or attracting attention. In direct contradiction to this view, Clark and Wasow (1998) argue for a language-production oriented approach, which takes into account Levelt’s (1989) works on language processing, and present an elaborate theory endorsing repeats as deliberate strategy that a speaker uses in order to gain time for planning (cf. Fig 2 below).

Stage	Speaker S's action	Spoken example
I. Initial commitment	S commits to a constituent	I
II. Suspension of speech	S stops vocalizing	{
III. Hiatus	S deals with potential delay	uh
IV. Restart of constituent	S restarts the constituent, restoring continuity to it	} I wouldn't be surprised at that

**Fig. 2:** Clark and Wasow's Theory (1998)

According to their theory, the first step a speaker needs to undertake is the initial commitment to a word and in pronouncing it he makes an obligation to produce the whole utterance. The second step involves the suspension of speech. Although, Levelt (1989) claims that a speaker can suspend his speech at almost any point, Clark and Wasow's research suggests that rising complexity of the constituent increases the likelihood of repeats. As function words tend to be readily available, present almost no pronunciation difficulties and most importantly tend to introduce major constituents, their repeat index is accordingly higher than the one of content words. Clark and Wasow (1998: 225) also claim that the initial commitment is not "merely premature, but preliminary: At the point at which speakers make the commitment, they are already expecting, at some level of processing, to suspend their speech." To prove their point, they focus on the prosody of repeats. In connected speech, function words hardly ever receive the full pronunciation and tend to be reduced. In the case of the first word in a repeat, the pronunciation however seems to be somewhat longer and fuller, e.g. *the* as /ði:/. Furthermore, the word(s)<sup>19</sup> do not undergo resyllabification. Having been nicknamed phonological orphans, they evidence the preliminary syntactic commitment both to the words themselves and to the constituent they initiate.

Following the suspension of speech, the speaker has to tackle the potential delay. The means a speaker has at his disposal are diverse: the range includes uninterrupted consecutive reiteration of the word (*I I*), both unfilled and filled pauses (*I . I ; I uh I*) and various discourse markers such as *you know* (*I you know I*). The reason for placing disruptions just before the to-be-repeated item stems from the continuity hypothesis. The continuity hypothesis

---

<sup>19</sup> In the case of contractions *I'm* does not resyllabify to "I.mem.ployed."

presupposes that “speakers prefer to produce constituents with a continuous delivery” (Clark and Wasow, 1998: 206). Restarting, i.e. repeating, a constituent thus restores the ideal delivery speech disrupted by suspension and hiatus. The reasons for such preference seems to be complex: it may be easier to formulate and produce a constituent from its beginning, it may be an act of thoughtfulness towards the addressee as it helps his orientation or it may be an attempt of the speaker at “present[ing] [himself] as prepared, thoughtful, and articulate” (Clark and Wasow, 1998: 207).

As may be inferred from the above mentioned studies, repeats “are not to be taken as a sign of sloppy or lazy performance” (Carter and McCarthy, 2006: 173) but should be regarded as a an effective and convenient tool which helps the speaker to temporarily deal with planning pressure without the necessity to disrupt his speech in a too prominent way. The research also suggests that the way speakers make use of this particular speech management phenomenon is by no means haphazard or coincidental and that speakers have adopted it as a partially automated strategy which is however employed at a precise point of speech having the primary function of gaining time for speech planning.

### ***2.3.2 Filled pauses in native and non-native speech***

The second performance phenomenon that both native and non-native speakers use, when dealing with planning pressure is represented by the so-called filled pauses<sup>20</sup>. As regards their functions and location in the course of continuous speech, they more or less correspond to repeats, i.e. speakers use them to cope with the increased demands they face when

---

<sup>20</sup> As with the term performance phenomena, there is no consensus about this term: labels as hesitators (Biber et al. 1999), fillers (Clark and Fox Tree, 2002) and others are used.

embarking on a new syntactic unit requiring them to compromise between the necessity of staying fluent and having enough time to plan their utterance. As opposed to the unfilled pause when a speaker is silent for a second or two, a filled pause is characterized by vocalization, usually accompanied by nasalization. In transcription these sounds are conventionally rendered as *er* and *erm* in British English and *uh* and *um* in American English<sup>21</sup>. Furthermore, Carter and McCarthy (2006: 172) also consider lexical forms such as *like*, *well* or *you know* as representatives of filled pauses.<sup>22</sup> If we take into account other performance phenomena, Biber et al. (1999: 1054) claim that filled pauses represent “an intermediate between unfilled pauses and repeat phenomena in marking a hold-up in the production process.”

Their intermediacy consists primarily in their distribution - although they are to be found at major syntactic boundaries as unfilled pauses, the frequency of their occurrence is higher at lesser or medial boundaries such as dependent clauses or coordinate constructions. A study conducted by Kjellmer (2003) shows that the latter lead the frequency list with the conjunction *and* at its top. Furthermore, his research has also discovered that the preferred position for the filled pause is the one following the conjunction (*and er* 76,5%, *and erm* 50,7%) rather than the reverse. Combined with the findings showing lower frequency for a filled pause collocating with a subordinating conjunction, Kjellmer (2003: 180) claims that “[t]his supports the idea that FPs [filled pauses] are used to signpost a new thought unit: subordinated clauses are obviously more intimately integrated semantically with their matrix clause and therefore less often constitute a completely new and separate thought unit.”

---

<sup>21</sup> However, the practices differ within the academic community - for instance, Totie (2014) claims using *eh*, *ehm* for British English and *uh*, *uhm* for American English. Moreover, Clark and Fox Tree (2002) use *uh* and *um* for the British English based London-Lund Corpus of Spoken English.

<sup>22</sup> It needs to be highlighted that the majority of grammars and other studies consider filled pauses to be realized by *er*, *erm*, alternatively by *uh* or *um*. The present thesis also adheres to such concept.

Moreover, the use of coordination refers back to one of the characteristic features of online speech: a speaker has a very limited time to plan and cannot produce elaborate constructions, therefore he rather opts for adding clauses to a linear sequence.

On the phrase level, filled pauses tend to occur more frequently with the increasing complexity of the phrase, be it a noun, adjectival or verbal one. In general, filled pauses tend to occur more frequently before the whole phrase rather than before a part of it, the only exception being modal verbs. In such case, the filled pause tends to come in-between the modal and the main verb. Suggesting that modality might intensify planning demands, Kjellmer (2003: 177) offers the following explanation: “Modals (...) have a more distinctive meaning of their own and represent a step in the semantic progression, as it were.” Finally, there are cases where filled pauses systematically occur inside phrases, usually preceding semantically ‘heavy’ words or important lexical choices thus suggesting that “the speaker takes time to select the appropriate expression” (Carter and McCarthy, 2006: 173).

The debate concerning the function of filled pauses in the spoken discourse traditionally accounts for two areas: on one hand for hesitations, such as the speaker cannot retrieve a word from his memory, searches for an appropriate one or highlights the following correction (Kjellmer 2003); on the other hand, filled pauses serve as signals which are motivated by other reasons than planning problems. These typically include the speaker either taking or keeping the floor (Maclay and Osgood, 1959)<sup>23</sup>, the speaker inviting their addressees to speak (Clark and Fox Tree, 2002) or speakers being polite (Fischer, 1999). As we may see, there is no consensus as to how the functions of filled pauses can be reliably described. In their article, Clark and Fox Tree (2002) propose a pragmatic approach based on Gricean

---

<sup>23</sup> Moreover, if the unfilled pause following the filled one exceeds certain time limit, the speaker might be suggesting that he has finished his turn.

theory (1975), particularly regarding the maxim of relevance, i.e. speakers and listeners share a given context and thus should be able to relate the content of the conversation to the particular situation and construct implicatures. Therefore, in their view filled pauses *uh* and *um* have the two following basic meanings and all the other functions are implicatures arising from the speaking situation. Clark and Fox Tree distinguish two basic meanings:

(a) <i>Uh</i> : "Used to announce the initiation, at $t('uh')$ , of what is expected to be a minor delay in speaking." (b) <i>Um</i> : "Used to announce the initiation, at $t('um')$ , of what is expected to be a major delay in speaking."
--

**Fig. 3** Clark and Fox Tree's Filler-as-word hypothesis (2002)

The distinction between *uh* and *um* as used for announcing a minor, respectively a major delay in speaking is based on data obtained from the London-Lund Corpus of Spoken English. The data suggest that *um* is followed by longer silent pauses than *uh* (0.68 vs 0.25 units) and also that *um* was more frequently followed by a period of silence (61% vs 29% of the time). These findings were consistent among various speakers thus indicating that "[s]tudents were able to estimate how long it would take them to retrieve the answer before they had retrieved it" (Clark and Fox Tree, 2002: 84). What is more, they argue that filled pauses should be considered interjections. Covering all language levels, they offer the following arguments in support of their filler-as-word hypothesis: in prosody, filled pauses conform to the English language system and more importantly are often cliticized onto preceding words and undergo resyllabification (*and uh* becomes *an.duh*); similarly to interjections they do not interfere with syntactic structures; semantically, speakers can distinguish a basic meaning (see Figure 3), and a range of implicatures conventionally oriented towards the rendition of the ongoing conversation. What goes against their proposal is the listener's point of view - several experiments manipulating the presence, respectively non-presence of filled pauses suggest that with natural quantity and distribution, i.e. the

particular speaker makes use of filled pauses in moderate frequency and at appropriate locations in the utterance, a listener does not consciously discern their presence in speaker's utterance (cf. Corley, MacGregor, Donaldson, 2007)

Since filled pauses are predominantly used to signal planning problems, their ubiquitous presence in the learner corpora, especially of lower proficiency levels, is hardly surprising. With the increasing level of language proficiency, the numbers capturing the overuse of filled pauses decrease, yet even the students who have reached the advanced level use filled pauses more frequently in comparison with native speakers. A pilot study (Götz, 2007) performed on German advanced speakers of English shows relatively substantial overuse of filled pauses and furthermore highlights the presence of other performance phenomena which learners tend to combine in order to gain time for planning. The necessity of combining performance phenomena has been described also by Tavakoli (2011: 75–77): her study on non-native speakers illustrates the use of filled pauses collocating with repetitions and reformulations. Another study carried out by Foster and Tavakoli (2009) supports these findings and furthermore emphasizes the influence of the task on speaker's fluency<sup>24</sup>: whereas native speakers did not seem to be affected, non-native speakers' performance has suffered due to task variation.

A more detailed study focusing on performance phenomena by Götz (2013) sheds more light on their use by non-native speakers and reveals some unexpected details. Although on average LINDSEI-GE speakers do indeed have tendency to overuse filled pauses, a closer look at individual speakers shows that it does not account for all of them - 16% of the speakers underuse filled pauses and even 6% of these do that to a highly significant degree.

---

<sup>24</sup> However, Foster and Tavakoli's (2009) research also showed that syntactic complexity and lexical diversity has been impacted both in the case of native and non-native speakers due to task altering.

Similarly on the phrase level, non-native speakers use filled pauses more frequently when compared to the native speaker mean. Yet, there are also 12% of speakers who do not conform to the general tendency and underuse filled pauses even within clauses. And finally, the overview of the use of filled pauses within constituents shows a resembling pattern: 82% of LINDSEI-GE speakers overuse them significantly, yet 8 learners do not divert from the native speaker norm and one even underuses filled pauses. Therefore Götz (2013:111) concludes that “[t]hese findings indicate that filled pauses do not constitute a general and highly problematic area for *all* German learners of English represented in LINDSEI-GE”. As she aptly remarks, such a statement however requires a complex study which takes into account other performance phenomena and/or speech enhancing strategies, which might reveal more accurately the way non-native speakers deal with planning pressure.

In general, we might conclude that filled pauses represent a resourceful tool which both native and non-native speakers use in order to tackle planning problems. Native speakers prefer using them primarily at syntactic boundaries of either major or of minor (e.g. subordinate clauses or coordinate structures) significance. Although we may find several exceptions, overall non-native speakers show stronger tendency to overuse filled pauses on all structural levels and furthermore employ other speech management strategies when constructing their utterance. In summary, the above mentioned research suggests that the use of one performance phenomenon does not suffice non-native speakers and in order to gain time for planning they make use of other strategies.

### ***2.3.3. Concluding remarks***

Looking jointly at the performance phenomena under scrutiny, we may wish to emphasize that both repeats and filled pauses are undeniably present in the spoken discourse,

be it a native speaker or a non-native speaker. The difference lies in their overall frequencies and distribution: what makes non-native speakers sound somewhat unnatural is the tendency to overuse performance phenomena and their location on lower syntactic levels and inside constituents which furthermore highlights their presence.

Considering all the potential variables that might come to play in speech production, Götz (2013: 138) distinguishes native and German speakers of English in the following manner: whilst native speakers do have their own very individualistic preference for combining formulaic language and performance phenomena in order to achieve fluency and show a great deal of variation among the speakers, “[t]he learners, on the other hand, are characterized by using either a comparatively high proportion formulaic language or a high proportion of all the strategies and they thus clearly lack the degree of variation displayed by the native speakers.” Although exploring a smaller sample of non-native speakers, the present thesis aims to test such a statement by closely analyzing the frequency and distribution of both repeats and filled pauses. The following section provides a description of the chosen sample and offers an overview of the methods used in order to map the use of the two selected performance phenomena.

## 3 Material and method

### *3.1 Material*

The data for the present thesis were obtained from the Czech component<sup>25</sup> of the multinational learner corpus *The Louvain International Database of Spoken English Interlanguage* (LINDSEI\_CZ henceforth), compiled at Université Catholique de Louvain. Being a spoken corpus of advanced learners of English with various L1 background (3rd- or 4th-year university students of English), it was originally designed as a counterpart to the International Corpus of Learner English (ICLE), a learner corpus of English comprising argumentative essays of higher intermediate to advanced learners of English. To date, 14 universities have contributed to LINDSEI, altogether collecting 1.4 million words of learner language, and other institutions are currently working on completing their national components.

Each national component of LINDSEI, i.e. a collection of c. 50 speakers, follows the same structure made up of three tasks: first, a student is asked to speak about a set topic having a short period of time to prepare, secondly, there is a free discussion concerned with the student's interests, study experiences or any future plans, and thirdly, the student is required to describe a series of pictures which constitute a story without any previous preparation. It needs to be pointed out that the nature of the tasks may influence a speaker's online production to a certain degree as stated by Yuan and Ellis (2003). A recent study by

---

<sup>25</sup> The Czech component has been compiled at the Department of English Language and ELT Methodology, Faculty of Arts, Charles University in Prague under the supervision of PhDr. Tomáš Gráf.

Gráf (2015) has shown that especially the last spontaneous task significantly affects speakers' production. Nevertheless, as the third task accounts for only a minor part of the approximately 15-minute long interviews, the present thesis chooses not to regard task influence in the following analysis.

Subsequently, the interviews are transcribed following LINDSEI transcription guidelines<sup>26</sup> and accompanied by a profile which provides information about the learner, the interviewer and the interview itself. Furthermore, LINDSEI is also complemented by a control corpus, the Louvain Corpus of Native English Conversation (LOCNEC henceforth), which comprises c. 50 native speakers of English performing identical tasks as those in LINDSEI. Such a corpus provides ample material for comparison of spontaneous spoken production of native and non-native speakers.

For the purposes of the present analysis 3 native speakers and 3 non-native speakers<sup>27</sup>, i.e. Czech speakers of advanced English, were chosen, altogether providing 223 instances of repetitions (108 in the subcorpus of native speakers and 115 in the non-native speaker subcorpus respectively) and 596 instances of fillers (225 in native speaker subcorpus and 371 in the non-native one respectively). The examples were semi-automatically retrieved from the selected subcorpora, manually checked and subsequently subjected to analysis. The speakers were chosen based on their use of filled pauses per hundred words and a representative sample has been selected.<sup>28</sup>

---

<sup>26</sup> For details see <<https://www.uclouvain.be/en-307849.html>>

<sup>27</sup> The number of speakers has been reduced in contrast with the expected number stated in the annotation in order to meet the requirements of a diploma thesis as regards the number of examples, i.e. 200 for a diploma thesis.

<sup>28</sup> English speakers EN051: 2,2 Filled pauses phw; EN020: 4 Filled pauses phw; EN001: 7,1 Filled pauses phw  
Czech speakers: CZ006 3,1 Filled pauses phw; CZ039: 7 Filled pauses phw, CZ034: 11,1 Filled pauses phw

### ***3.2 Method***

In order to obtain quantifiable (and hence comparable) data, any analysis of spoken language must start with the definition of a basic unit for analysis. In the present study, we will make use of Foster, Tonkyn and Wigglesworth's AS-Unit (2000) since it takes into account both syntactic structure of an utterance and also considers other prosodic phenomena as a factor in deciding where a unit of speech begins and where it ends. Section 2.1.1.1. provides a theoretical background to theory of units of speech of analytical purposes and establishes the definition of AS-Unit: "An AS-Unit is a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either" (Foster et al., 2000: 365). The following example is taken over from the pivotal article on AS-Units by the above mentioned authors (2000: 367) and shows some of the symbols used in marking:

- (1) | and they pinned er a notice to his front :: telling everybody :: what he had done (0.5) |  
| and marched him around the streets with a gun at his back |

An upright slash is used to mark an AS-Unit boundary (no distinction is made between the beginning and the end of a unit) , a double colon marks a boundary of a subunit within an AS-Unit, either a non-finite verb form followed by at least one more constituent (*telling everybody*) or a subordinate clause (*what he had done*). The figure in brackets signifies the length of an unfilled pause which is crucial for deciding whether a coordinative structure will belong to one AS-Unit or not. In the example above, the pause has reached the limit of 0.5 seconds in which case the two main clauses are treated as two separate AS-Units. Furthermore, the first clause has to be marked by falling or rising intonation. Therefore, we

may observe that AS-Unit parsing takes into account prosodic features which may significantly influence the transcripts as highlighted by Gráf (2015: 73).

As far as performance phenomena are concerned, Foster et al. (2000: 368) take into account false starts, repetitions and self-corrections marking them by braces as in the following examples:

- (2) | {what about} can you give me a credit slip |
- (3) | and {one person} er one person enjoy the canoe |
- (4) | I think :: {they're a very} they have good time |

As the present thesis is not concerned primarily with false starts or self-corrections, the transcripts included in the appendix will not follow this marking. However, both features will be taken into account when analyzing co-occurrence of other performance phenomena in the proximity of filled pauses and repeats.

As regards repeats, one of the two foci of the analysis, “it is necessary to distinguish between those repetitions which indicate dysfluency, and those which are used for rhetorical effect” Foster et al. (2000: 368). Hence, those used for rhetorical effect have been excluded from the sample, cf. the following example:

- (5) | and (erm) . {we} we went by bus :: which is . a little annoying :: because it was a **long long** way .. |
- (CZ006)

It has to be noted that the recording had to be used as well in order to determine with certainty whether it concerns a repeat or repetition for rhetorical effect.

Example (5) also provides the graphical marking of filled pauses in brackets that is taken over from the LINDSEI transcription system. This greatly facilitated the semi-automatic retrieval as it was possible to search for the symbol of the bracket without taking into account

the different variations of the filled pause. For the sake of better orientation and clear arrangement, all examples of repetitions and fillers are highlighted in bold both within the text of the present thesis itself and in the enclosed appendix. Each example illustrating the use of the selected performance phenomena will be supplied with a number that corresponds to coding of speaker provided by either LOCNEC or LINDSEI and the respective number of its AS-Unit as stated in the appendix. Compare the following example:

(6) | **(er)** well {**I**} . **I** don't know |  
(CZ039, AS-Unit 25)

### ***3.2.1 Pilot study***

In order to avoid misleading findings, a pilot study has been carried out on one native and one non-native speaker, giving significant insight into the analysis. However, it needs to be emphasized that owing to the low number of speakers, no conclusive evidence has been obtained from the pilot study and the hypotheses formulated for the research will necessarily have to rely on previous larger-scale research by other authors (e.g. Biber et al. 1999, Götz 2007, Tavakoli 2011, Götz, 2013). On the other hand, the pilot study proved to be helpful as regards the particular treatment of the data. Those findings are summarized in the following paragraphs.

First of all, our concept of the above mentioned *beginning of a clause* had to be reformulated as we have found numerous occurrences where filled pauses (FPs henceforth) did not directly introduce the clause but were preceded by other parts of speech, mostly conjunctions or adverbs, such as discourse markers (cf. the following example):

(7) 11. | **and**: . **(em)** . also .. it had to have some mountains :: but {not . not=} nothing too difficult :: because I didn't want :: to get into trouble . so . (P=0,35) |

As we are dealing with spoken corpora, our definition of a clause is more discourse-driven and therefore we have to take into account how clauses are structured in speech. Biber et al. (1999:1068) describe the so-called ‘add-on strategy’: in spite of complex underlying structure, the surface structure presents “a linear sequence of finite clause-like units [...] and within each chunk, the syntactic processing required is simple.“ Therefore the rather plain coordinating (*and*, *but*) or subordinating (*because*) conjunctions employed to create a continuous flow of speech are considered to be an inherent feature of the analyzed discourse and hence are not interfering with the clause structure. Nevertheless, when the analyzed element was preceded by an optional adverbial, these instances were treated as not appearing directly at the beginning of a clause but at the beginning of the respective sentence element.

Therefore, following the methodology of Götz (2013)<sup>29</sup>, the analysis of fillers includes *filled pauses within clause* (FPWCL henceforth), i.e. appearing directly at the beginning or the end of a clause, alternatively next to a conjunction. As the previous research suggests, FPs tend to occur at the beginning of subordinate clauses hence further differentiation between coordinating and subordinating construction has been included. Other occurrences of filled pauses have been treated on the level of phrases, using Götz’s (2013) terminology *filled pauses within constituents* (FPWCON henceforth). Only the basic sentence elements such as subject, verb, object, complement or adverbial were taken into account.

Furthermore, we may wish to highlight that assigning a filled pause either to a specific clause or a specific constituent is necessarily problematic. As some (both older and relatively recent) authors claim (see for instance McClay and Osgood, 1959 or Carter and McCarthy 2006), filled pauses may serve to keep the floor in order to prevent other speakers to interfere

---

<sup>29</sup> Her methodology primarily draws on Biber et al. (1999), i.e. primarily syntactic segmentation.

in the conversation. Hence some instances of FPs might occur seemingly amid clauses and it is only the recording that can disambiguate these. Although AS-Unit segmentation does take into account several performance phenomena and prosodic features, it is somewhat surprising that filled pauses seem to be omitted from the analysis. Being aware of the fact that such handling is essentially subjective, the present thesis chooses to assign FPs to particular clauses or constituents based on the speaker's intonation and pausing, relying on the principle of linearity of speech in ambiguous cases, i.e. to assign a particular filled pause to the constituent on the right.

Secondly, the pilot study has also shown that the analyzed hesitation phenomena also tend to co-occur with instances of overlap, i.e. when both speakers talk at the same time. In the following example, the interviewer (A) is trying to help the interviewee (B) to finish the utterance:

(8) <B> | she studies there {so . probably} | </B>

A: alright she <overlap /> can't have hated it that much then </A>

<B> | <overlap /> (em) no I hope . (P=0,75) | </B>

Based on these findings, the overlap has been integrated into the list of possible collocates.<sup>30</sup>

Thirdly, the pilot study has discovered two problematic groups of examples which have been eventually excluded from the analysis. First, we have found several instances of FPs which do not fulfill the primary function of these phenomena (see section 2.3.1 and 2.3.2.) but rather represent inserts, i.e. “stand-alone words which are characterized in general by their inability to enter into syntactic relations with other structures (Biber et al., 1999: 1082). The particular instances found in our sample belong to the category of response forms

---

<sup>30</sup> At this point, it has to be mentioned that accurate transcription of all overlaps cannot be guaranteed as the sound quality of the recordings does not allow hearing them properly.

which can be exemplified by affirmative forms such as *uhu*, *mm*, *mhm* used primarily for backchannelling. The following example demonstrates the use of such inserts by the native speaker (B) merely responding to the interviewer.

- (9) <A> I mean it's as if you were living here really </A>  
<B> **mm** </B>  
<A> because well when we: .. well in Belgium when we go back . we we sort of . we have two homes in a way </A>  
<B> **mhm** </B>  
<A> and we we . we still have a lot of activities in .. our home towns </A>  
<B> **mm** </B>  
<A> and a few on campus . so here I think you have all your activities here </A>  
<B> **mm** </B>

Keeping these in the sample would have substantially affected the analysis and they were therefore excluded from the sample.

Additionally, a close inspection of the data has also revealed an important feature that needs to be pinpointed, namely that repetitions frequently occur within larger structures which also helps decreasing time pressure during speech planning. To be precise, these repetitions occur within reformulating sequences, where the speaker notices his or her mistakes or suddenly realizes that he or she wants to change the utterance. In literature, these reformulations have been dubbed differently, let us mention Götz's (2013: 37) *self-corrections*, *retrace-and-repair* by Biber et al. (1999: 1062), or *false starts* by Foster et al. (2000: 368). The following example by the native speaker technically shows repetition but a detailed examination reveals reformulation of the mere personal pronoun *I* to the same personal pronoun and a contracted auxiliary 'm :

- (10) | erm . **I I'm** doing a . linguistics minor . (erm) as part of .. (er) |

Hence, all repetitions occurring within the frame of larger performance phenomena have been excluded from the analysis. They are only taken into consideration when analyzing other

performance phenomena in the proximity of either filled pauses or fillers, having the label *reformulations*. These would cover instances of both false starts and self-corrections.

Finally, a decision has to be made whether personal pronouns with enclitic verb contractions are to be treated as one word or two since the pilot study has showed several instances of these structures. Although from the syntactical point of view, they represent two compulsory sentences elements, “[i]t is likely that such contractions are processed by the speaker **and** hearer as single words, and therefore, for the purposes of studying dysfluency phenomena, that they should be treated as such” (Biber et al., 1999: 1061). The present paper chooses to adhere to such a claim, hence all contractions will be handled as one word.

In conclusion, the pilot study has shown that spoken discourse analysis requires a minute examination of both data and categories employed in the analysis. Despite the efforts trying to emancipate spoken language research from written discourse traditions, the lack of effective technological devices, insufficient research on spoken grammar and the dearth of comprehensive transcription methods necessarily contribute to its limitations and have to be taken into account.

### ***3.2.1 Hypotheses***

Even though there are only few studies concerned with the comparison of performance phenomena usage by both native and non-native speakers, they will serve as a source for the following hypotheses. We will mainly draw on the work of Sandra Götz (2007, 2013) and Biber et al. (1999), furthermore, we may turn to more narrowly specialized articles by such authors as Tavakoli (2011) . Additionally, it has to be stressed that the research questions will be aimed at rather qualitative features of the data due to its scale.

Nonetheless, the two subcorpora will be compared using the log-likelihood calculator in order to determine the proportion of both FPs and repeats in the speech of native and non-native speakers. The findings of Götz (2013: 110) suggest “that the great majority of native speakers can be characterized by a low number of filled pauses, whereas the great majority of learners highly overuse filled pauses.” On the other hand, repeats have been found to be significantly underused in the corpus of German advanced speakers of English (Götz, 2013: 107). Furthermore, the instances of both FPs and repeats will be compared with regard to their relative frequency per hundred words and per AS-Unit, excluding the figure for filled pauses per hundred words.<sup>31</sup> Therefore the first two hypotheses to be tested in our sample are formulated as follows:

1. Czech advanced speakers of English overuse FPs.

2. Czech advanced speakers of English underuse repeats.

The next feature, common both to FPs and repeats, is the occurrence of other performance phenomena in the proximity of either of them. In her article on pausing patterns, Tavakoli (2011: 75-76) has described how learners usually felt the need to employ reformulations and replacements together with (un)filled pauses and repeats in order to cope with the online planning pressure. Similar findings have been quoted by Götz (2007: 74), who claims that learners cannot make do with one performance phenomenon and use others to gain time for planning.<sup>32</sup> Unfortunately, any specification of the scope in which other phenomena occur is not given by either of these authors. However, based on the examples provided in the two above mentioned articles, the present thesis chooses to take into account other

---

<sup>31</sup> The figures signifying the use of filled pauses per hundred words have been used as the determining factor when selecting speakers for the analysis.

<sup>32</sup> (...) dass den Lernen die Wiederholungen nicht genügen, um sie zur Planung der gesamten Aussage zu nutzen, da in diesen Einheiten noch weitere Unflüssigkeiten vorkommen (...).

performance phenomena in the proximity of 4 words, i.e. approximately the length of one phrase. Hence the next hypothesis is as follows:

3. Czech advanced speakers of English employ further performance phenomena in the distance of 4 words of the respective filled pause or repeat.

Considering filled pauses only, previous research shows that the difference between native and non-native speakers consists not only in their frequency but also in their distribution. Based on Götz's findings (2013: 111-113), learners tend to overuse FPs also at lesser boundaries such as subordinate clauses or within phrases. Her previous research has also discovered significant overuse of FPs following conjunctions *and*, *because*, and preposition *with*. In comparison with native speakers, the figures have reached three or four times higher values. Based on these, the following hypothesis will be tested:

4. Czech advanced speakers of English overuse FPs both at higher and lesser boundaries, especially when preceded by conjunctions or prepositions.

Finally, a closer look at repeats will aim at the word or a cluster of words that is repeated in order to decrease planning pressure. As the likelihood of a repeat increases with the overall frequency of the particular word (Biber et al., 1999: 1056), repetitions of function words represent the most presumable outcome as regards both native and non-native speakers. Owing to their clause-initial position, they are the foci of the highest planning pressure. The analysis performed by Götz (2013) has showed no significant differences as regards pronouns and possessive determiners, on the other hand, the data showed significant overuse of verbs and prepositions, i.e. usually mid-clausal elements. As she points out (2013:108), "learners' higher need of planning phases even within clauses is clearly reflected by high repeat indices in these groups." Furthermore, advanced learners have been found to have a lower index of

subject-verb contractions therefore they do not repeat them as often as native speakers do.

Hence the two following hypotheses are to be tested:

5. As regards the repeated word, Czech advanced speakers of English overuse repeats of verbs and prepositions.

6. As regards the repeated word, Czech advanced speakers of English underuse repeats of subject-verb contractions.

These six hypotheses will be tested and discussed in detail in the following sections. Results showing the figures for both native and non-native speakers will be presented and compared. The subsequent discussion of findings will strive to determine basic differences in the usage of filled pauses and repeats and moreover infer particular implications for advanced learners' education.

## **4 Analysis**

The following sections will present the findings of the analysis. First, overall frequencies of both filled pauses and repeats will be compared in order to find the differences

in the usage of the selected performance phenomena by non-native and native speakers. Furthermore, attention will be paid to other performance phenomena that might occur in the proximity of filled pauses and repeats in the native and non-native subcorpora. Subsequently, both phenomena under scrutiny will be dealt with separately in order to ascertain more detailed information. First, we will have a look at the distribution of filled pauses within clause and clause constituents, and secondly, we will address the issue of which words tends to be repeated in the two subcorpora.

## ***4.1. Overall frequencies***

### ***4.1.1. Filled pauses***

As has been touched upon in the methodological chapter, our attention has been directed only to those instances of filled pauses which are used in order to cope with planning pressure. Those FPs that represented responses to the interviewer and hence did not take part in the planning process were discarded from the sample. Following the LINDSEI transcription system, they mostly involved FPs such as *mhm* and *uhu*. Instances of *mm* had to be individually evaluated by means of the recording where the speaker's intonation proved to be the decisive factor. Furthermore, no distinction has been made concerning the form of the filled pause itself.

The data provided 371 occurrences of filled pauses in the non-native speaker subcorpus (speaker CZ006 - 60 FPs, speaker CZ034 - 190 FPs, and speaker CZ039 - 121 FPs), and in the case of native speakers, the data included 225 instances (speaker EN001 - 109 FPs, speaker EN020 - 53 FPs, and speaker EN051 - 63 FPs). Subsequently, an FP rate was calculated as a ratio of FPs per hundred word (phw) and per AS-Units (per AS-Unit). The following table summarizes the results for both subcorpora.:

	LINDSEI_CZ	LOCNEC
<b>Corpus size in words</b>	6351	6743
<b>Number of FPs</b>	371	225
<b>phw</b>	5.842	3.337
<b>per AS-Units</b>	0.96	0.4

**Table 1.** Filled pauses as used by non-native and native speakers

As can be inferred from the table learners required more time for planning since they used almost one filled in pause in every AS-Unit, whereas native speakers integrated a filled pause into their speech approximately in every second AS-Unit. Finally, the speakers' use of filled pauses in LINDSEI\_CZ and LOCNEC were compared using log-likelihood (LL) calculation. With the LL value of 45.11, the analyzed Czech advanced speakers of English were found to greatly overuse filled pauses in their speech. This finding was furthermore highly significant ( $p < 0.0001$ ) and thus corroborated hypothesis number 1.

#### ***4.1.2. Repeats***

Taking into account the repetitive occurrences of a word or a cluster of words, the present analysis focused only on cases of unintentional reiteration, i.e. when speaker did not put particular emphasis on the sequence. Those cases when the speaker has deliberately laid considerable stress on a repeat were excluded from the sample. Again, the recording proved to be of invaluable assistance as the mere transcript would not suffice to determine the actual function of a particular repeat. Furthermore, repeats that came under larger performance phenomena such as false starts or self-corrections were also discarded from the analysis.

The data provided 115 instances of repeats as used by Czech advanced speakers of English (speaker CZ006 - 44 repeats, speaker CZ034 - 39 repeats, and speaker CZ039 - 32 repeats), the English native subcorpora contained 103 examples (speaker EN001 - 46 repeats,

speaker EN020 - 28 repeats, and speaker EN051 29 repeats). As with filled pauses, rates for repeats per hundred words (phw) and per AS-Units were calculated and are summarized in the following table:

	LINDSEI_CZ	LOCNEC
<b>Corpus size in words</b>	6351	6743
<b>Number of repeats</b>	115	103
<b>phw</b>	1.811	1.528
<b>per AS-Unit</b>	0.3	0.18

**Table 2.** Repeats as used by non-native and native speakers

Lastly, the relative frequencies of repeats in the two analyzed subcorpora were compared using log-likelihood statistical method. It has indicated that Czech advanced speakers of English overuse repeats which furthermore corresponds to the rates per AS-Unit. On the other hand, such a finding is in contradiction to hypothesis number 2. However, the result is not statistically significant ( $p > 0.5$ ) and therefore we cannot draw any reasonable conclusions from this finding. Given the size of the present sample, a detailed analysis of the repeated word may shed some more light on this issue as the qualitative approach could prove to be beneficial in revealing further particularities (see section 4.4.).

## ***4.2 Further performance phenomena***

Speaking a foreign language places higher demands on language planning, therefore it is not surprising that learners of English are expected to deal with the pressure by using other performance phenomena. Bearing this in mind, both filled pauses and repeats were examined from the point of view of whether further hesitation markers occurred in the proximity of 4 words. Further performance phenomena were categorized in the following manner: first, there

was no co-occurring element accompanying the particular filled pause or repeat; secondly, only an unfilled pause (UnFP) has been found; then, there was either one or two other performance phenomena besides an UnFP. Furthermore, the pause surrounding the performance phenomenon under scrutiny was vocalized hence an FP or an FP accompanied by other performance phenomena have occurred. The last category included other performance phenomena, such as reformulations, repeated or truncated words, being in the proximity of the analyzed FP or repeat.

As far as FPs are concerned, the analysis has not brought any conclusive evidence since the values of native and non-native speakers did not differ to a substantial degree. The only case where native speakers differed considerably involved the proportion of unfilled pauses suggesting that learners require more time for planning and the use of one type of a performance phenomenon may not suffice as demonstrated by Götz (2007). Moreover, the combination of two devices used in order to decrease planning pressure may not be enough to satisfy the learners' needs as there were multiple occurrences of three elements in close proximity. The following table summarizes the results for FPs, where considerably different values have been highlighted in grey:

Co-occurring element	LINDSEI_CZ	%	LOCNEC	%
No	84	22.6	49	21.7
UnFP	143	38.5	107	47.6
UNFP + 1 other	76	20.5	47	21
UNFP + 2 other	18	4.9	1	0.4
FP	24	6.5	4	1.8
FP + 1 other	5	1.3	1	0.4
Others	21	5.7	16	7.1
<b>Total</b>	<b>371</b>	<b>100</b>	<b>225</b>	<b>100</b>

**Table 3.** Filled pauses in close proximity of other performance phenomena

The considerably higher percentage of filled pauses occurring in the proximity of another filled pause can be evidenced by the proportion of filled pauses that have been repeated without any emphasis, i.e. when the speaker needed more time for planning and did not use the particular filled pause for rhetorical effect. While in the English subcorpus only 4 instances that account for 1.8% of the sample have occurred in close proximity, Czech advanced speaker of English repeated filled pauses in 8.4 %, i.e. in 31 instances. The following example produced by one of the Czech speakers illustrates this point:

- (1) | I I really think so :: and **(erm)** . **(er)** one of my friends :: who went on an on  
the[i:] Erasmus programme |  
(CZ034, AS-Unit 58)

It needs to be moreover stressed that it was Task 3, picture description, that included those four instances of a repeated filled pause as used by English speakers. It comes as no surprise since the recent study by Gráf (2015: 146) pointed out “that the design of Task 3 presents a greater challenge to spontaneous speech production even for native speakers.”

Furthermore, the data also provided the following example, produced by speaker EN051 in Task 3, which certainly corresponds to the above mentioned statement:

- (2) | ... **(erm)** .. (em) (em) (em) ... **(er)** I was going to say . :: it's .. it's  
Suzanne's .. :: Suzanne is the woman :: and it's **(erm)** . it's Suzanne's hu=  
husband's birthday . very soon (0,5) |  
(EN051, AS-Unit 25)

As the marking suggests, only the first and the last filled pause included in the string of filled pauses has been subjected to the analysis. This is due to the emphasis the speaker placed on the four in the middle which can be only attained from the recording. However, it is beyond doubt that the speaker was only trying to gain time for planning, in this case making up the

story, and did not want to convey any particular message. Such usage may support Clark and Fox Tree's (2002: 103) claims which regard filled pauses as interjections speakers use "to announce that they are initiating what they expect to be a minor or major delay before speaking."

Let us now consider repeats that have been found to co-occur with further performance phenomena. It needs to be highlighted that repeating a word or a cluster of words necessarily involves some limited planning ahead therefore it can be considered more demanding than producing a filled pause. Such presumption seems to be valid since the analysis carried out for repeats yielded more straightforward information. The data showed that undeniably repeats put higher demands on language planning for the learners as they avoided using further devices only in 46.1% whereas the native speakers did not need any other performance phenomena in 69 %. Similarly to filled pauses, Czech advanced speakers of English made use of further performance phenomena, including their combinations.

The following table presents the results of the analysis showing substantial value differences in grey:

Co-occurring element	LINDSEI_CZ	%	LOCNEC	%
No	34	29.6	35	34
UnFP	19	16.5	36	35
UNFP + 1 other	33	28.7	22	21.4
UNFP + 2 other	2	1.7	1	0.9
FP	21	18.3	6	5.8
FP + 1 other	0	0	0	0
Others	6	5.2	3	2.9
<b>Total</b>	<b>115</b>	<b>100</b>	<b>103</b>	<b>100</b>

**Table 4.** Repeats in close proximity of other performance phenomena

Particularly remarkable is the tendency of Czech advanced speakers of English to combine repeats with filled pauses. A closer analysis examining the collocates of all repeats suggests that learners incline to place an FP inside the sequence as illustrated by the following example:

(3) | and (eh) I always thought :: that the tragedies were {**much**} (er) **much** more  
difficult to stage |

(CZ039, AS-Unit 3)

They were found to do so in 8.7% as opposed to the native speakers who inserted an FP into a repeat in only 3%. Such a strategy demonstrates the need to lengthen the delay in speaking and it is therefore not surprising that it is more frequently sought-after by the learners.

Finally, learners may find themselves in situations where the planning pressure rises increasingly, especially when they produce more complex structures as in the case of the inserted subordinate clause in the following example:

(4) | and also (**er**) . (**eh**) some . :: they are called <foreign> Sudetendeutschen </  
foreign> :: {**some**} (**er**) . (**er**) {**some**} (**er**) . (**erm**) order it even from Germany  
mo= again mostly from <foreign> Bayern </foreign> (**erm**) . |

(CZ034, AS-Unit 52)

This particular instance uttered by one of the Czech speakers features multiple repetitions of filled pauses as well as a repeat. It is noteworthy that this chain of performance phenomena involves the subject where the planning pressure reaches its peak. What may also play a role is the presence of a foreign word requiring the learner to process yet another language<sup>33</sup> for a brief moment.

---

<sup>33</sup> Moreover, the recording revealed that the speaker pronounced the word in German accent which may further increase the demands on language planning.

In conclusion, we may say that the learners' need to buy more time for planning is reflected not only in the frequency of performance phenomena but also in their combination and frequent collocation. In comparison to filled pauses, repeats seem to require more time and effort from the learners since they occur more frequently in the vicinity of other performance phenomena than in the case of native speakers. Secondly, we may claim that these findings do not support the evidence for theories that regard filled pauses as words (cf. Clark and Fox Tree, 2002).

### ***4.3 The distribution of filled pauses***

As we have described in section 4.1.1., Czech advanced speakers of English have been found to significantly overuse filled pauses in comparison with native speakers. Yet the question remains whether the learners employ this particular performance phenomenon in the same fashion as their L1 counterparts do. Therefore, taking a qualitative point of view, a more detailed analysis has been performed. Our focus lay in determining the exact position of a filled pause within a clause with regard to the AS-Unit segmentation. All filled pauses have been assessed as to whether they occurred at a beginning or an end of a clause (FPWCL) or inside one (FPWCON). The data provided 187 instances of FPs used by Czech speakers of English at the beginning of clause, which accounts for almost a half of the examples (50,4 %). In the English native speakers corpus, filled pauses occurred at the beginning of a clause in 70,7 % concerning 159 examples. As regards lower structural levels of a clause, Czech speakers have made use of filled pauses in 184 instances whereas English speakers employed this particular performance phenomenon in only 66 instances. Subsequently, relative frequencies have been calculated per hundred words (phw) and per AS-Unit. The following table summarizes the results for both native and non-native speakers:

	LINDSEI_CZ	LOCNEC
<b>Corpus size in words</b>	6351	6743
<b>Number of FPWCL</b>	187	159
<b>phw</b>	2.94	2.36
<b>per AS-Units</b>	0.48	0.28
<b>Number of FPWCON</b>	184	66
<b>phw</b>	2.9	1
<b>per AS-Units</b>	0.48	0.12

**Table 5.** FPWCL and FPWCON in LINDSEI\_CZ and LOCNEC

Furthermore, the two corpora have been compared using the log-likelihood calculator and in both cases Czech advanced speakers of English have been found to significantly overuse filled pauses; in the case of FPWCL with the significance at  $p < 0.05$ ,  $G^2$  values of FPWCON have reached even higher significance at  $p < 0.0001$ .

As can be inferred from the data, Czech advanced speakers of English seem to struggle with online production both at the beginning and further in the course of an utterance, and require more devices that would enable them to gain time for planning. This finding partially corroborates hypothesis number 4.

Those instances which were found at the beginning of a clause were further distinguished as introducing either an independent or a subordinate clause. As the following table suggests both groups were found to behave in a similar way, i.e. the majority of FPWCL featured FPs at its very beginning. However, Czech advanced speakers have proportionally used more FPs at the beginning of a subordinate clause which gives further support to hypothesis number 4. Compare:

	LINDSEI_CZ	%	LOCNEC	%
<b>Bg</b>	130	70	120	75.5

	LINDSEI_CZ	%	LOCNEC	%
<b>BgSub</b>	31	16.6	16	10
<b>End</b>	25	13.4	23	14.5

**Table 6.** Position of a FP within FPWCL

Similarly, filled pauses within constituents were analyzed in relation to which sentence element they either introduced or whether they occurred in the middle of a phrase. Owing to the small size of data available for this partial analysis, the conclusions which can be inferred will be necessarily speculative. The following table summarizes the ratios:

	LINDSEI_CZ	%	LOCNEC	%
<b>Bg CON</b>	98	52.3	27	40.9
<b>Mid CON</b>	86	46.7	39	59.1

**Table 7.** Position of a FP within FWPCON

First of all we need to bear in mind that learners overuse FPWCON (2.9 phw for non-native vs 1 phw for natives) so the portion of FPs as produced by English speakers comprise rather sporadic uses. Nevertheless, the higher proportion of FPs produced within a phrase has been obtained from the data. In his study which included larger number of speakers, Kjellmer (2003: 174) claims that “although far less often, [FPs occur] in front of single words inside phrases, typically before semantically ‘heavy’ words.” Therefore we may assume that when a native speaker turns to using a FP within a constituent, his motives are most likely semantically oriented since the target word will not be a part of the everyday vocabulary. Furthermore, the particular FP may serve as a verbal signal for the listener since Kjellmer (2003: 188) also pointed out that “[q]uite often the FP in this way precedes an ironic and disapproving understatement.”

As regards Czech advanced speakers of English, their preference for the beginning of a constituent may be attributed to further planning problems within a clause as such since they have to take into account not only the semantics of the phrase but also its syntactical composition.

Finally, we should address the issues whether learners' overuse of FPs is related to their collocation with conjunctions and prepositions. As Czech advanced speakers of English they have used a FP following a conjunction in 55 cases and following a prepositions in 20 instances. In the case of English speakers, conjunctions preceded FPs in 41 examples and prepositions in only 5 instances. Such a result does not represent conclusive evidence hence hypothesis number 4 could be neither corroborated nor rejected in its full extent. Furthermore, Kjellmer (2003: 180) found also native speakers to use combinations of coordinating conjunctions and FPs (*and er, but er*) fairly frequently therefore the type of conjunction should be specified in detail in further research.

#### ***4.4. Repeated word***

As opposed to Götz (2013), whose research focused solely on taxonomy introduced by Biber et al. (1999) including mostly functions words, frequent verbs and subject-verb contractions, the present analysis - owing to deployment of the automatized process of the identification of repeated sequences - has not limited the set of words that the speakers have repeated since the data retrieval was conducted automatically. This proved to be of considerable importance as the research revealed that repeats consisting of a two-word sequence do not form a negligible portion of the sample.

The words or clusters of words that have been repeated were divided into three categories as they proved to be the most prominent and their frequency is given in the

following table:

	LINDSEI_CZ	LOCNEC
<b>Corpus size</b>	6351	6743
<b>Number of repeats</b>	115	103
<b>phw</b>	1.811	1.528
<b>Lexical Words</b>	9	10
<b>phw</b>	0.142	0.148
<b>Function Words</b>	91	74
<b>phw</b>	1.433	1.097
<b>Two Word</b>	15	19
<b>phw</b>	0.236	0.282

**Table 8.** Repeated word in LINDSEI\_CZ and LOCNEC

Confirming Biber et al.'s claims (1999), function words occurred most frequently in a sequence of two, alternatively three<sup>34</sup>, repeated words. The second most frequent group to be found consisted of two-word repeats. As section 4.3.1. will further specify, these also comprised primarily function words therefore their dominance over lexical words is not surprising.

Since our main interest consisted in which function words in particular have the tendency to be repeated by native and non-native speakers, a more detailed analysis has been carried out in order to pinpoint these differences. The following table presents the counts for selected groups of function words<sup>35</sup> :

	LINDSEI_CZ	%	LOCNEC	%
<b>Personal Pronouns</b>	45	49.4	17	23

<sup>34</sup> Although triple repeats have been found in the data, their number was rather negligible therefore they have been treated as double repeats (cf. Götz 2013).

<sup>35</sup> The group labelled *Other* included other pronouns, possessive determiners, negative particles or indefinite articles.

	LINDSEI_CZ	%	LOCNEC	%
<b>Definite Article</b>	13	14.3	3	4
<b>Prepositions</b>	17	18.7	3	4
<b>Conjunctions</b>	6	6.6	14	19
<b>Others</b>	10	11	11	15
<b>S-V Contractions</b>	0	0	26	35
<b>Total</b>	91	100	74	100

**Table 9.** Repeats of function words in LINDSEI\_CZ and LOCNEC

As can be easily inferred from the table above, Czech advanced learners of English have been found to underuse repeats of subject-verb contractions as there was no example of such an element. This finding was therefore highly statistically significant ( $p > 0.0001$ ) and corroborated hypothesis number 6. Owing to the fact that “the higher a word's frequency, the more likely it is to form repeats” (Biber et al., 1999: 1059), the ratio per hundred words (phw) for subject-verb contractions were calculated for both subcorpora giving 4.3 phw for English speakers and only 1.5 phw for Czech advanced speakers of English. Hence we may assume that the underrepresentation of repeated subject-verb contractions in LINDSEI\_CZ actually consists in the scarcity of their non-repeated forms in the Czech subcorpus. Additionally, the lower proportion of repeated personal pronouns in the native-speaker subcorpus may be connected to this issue the majority of the repeated contractions involved personal pronouns.

The second main interest of the present analysis concerned the use of repeated prepositions due to the increased planning pressure learners experience even further on in the course of an utterance. As has been expected, Czech advanced speakers of English have been found to overuse repeats of prepositions to a highly significant degree ( $p < 0.001$ ). This finding partially confirms hypothesis number 5. Unfortunately, the data have not provided enough examples of repeated verbs therefore hypothesis number 5 could not be corroborated

in its full extent. On the other hand, verbs mostly represented the second word in two-word repeats therefore they have been subjected to repetition to a certain degree (see the following section).

#### ***4.3.1. Two-word repeats***

Despite being a relatively scarce phenomenon, it is somewhat surprising that previous studies concerned with repeats as instances of performance phenomena (cf. Clark and Wasow 1998, Biber et al. 1999, Götz 2007 and 2013) have only seldom mentioned repetitive occurrences of a two-word sequence. Quoting sequences *I'll I'll*<sup>36</sup> and *it was it was*, LGSWE remains the only exception. Biber et al.'s research (1999: 1056) has found the former to be repeated 45 times per million words and the later 15 times per million words. On the other hand, such an approach can be justified by a highly demanding data retrieval that would require strenuous exercise since the data acquired from the present sample had to be obtained manually.

The present thesis have taken into account only those instances of a two-word repeat that were followed by other sentence elements and would not represent a complete structure. Hence the following example produced by speaker CZ039 and further similar ones were discarded from the analysis.

(5) | **I know I know** <laughs> |

The data provided 19 instances of two-word repeats in the English subcorpus, representing 18,4 % of all repeats, and 15 instances in the subcorpus of Czech advanced

---

<sup>36</sup> However, it needs to be pointed out that later on Biber et al. (1999: 1061) concedes that “[i]t is likely that such contractions are processed by the speaker **and** hearer as single words, and therefore, for the purposes of studying dysfluency phenomena, that they should be treated as such.”

speakers of English, which accounts for 13% of the sample. Furthermore, the two-word sequences have been analyzed with regard to which particular word class they belonged. The following table summarizes the results:

Two-Word Repeat	LINDSEI_CZ	LOCNEC
Pronoun + Verb	10	14
Prep + Det	4	1
Others	1	4
Total	15	19

**Table 10.** Combinations of word classes occurring in two-word repeats

It comes as no surprise that the majority of examples involved pronouns combined with verbs, constituting subject and predicate, since repeats in general have the tendency to occur at the beginning of an utterance where planning pressure reaches its highest point (cf. for example Carter and McCarthy, 2006: 173). Furthermore, personal pronouns represented the preferred choice (12 in the case of native speakers and 7 in the case of non-native speakers) which corresponds to their high repeat index as illustrated by Biber et al. (1999: 1057). Similarly, the choice of a verb in the analyzed two-word sequence usually involved the verb *to be* - we have found 8 instances in the English subcorpus and 7 in the Czech one. Otherwise, the sample included various lexical verbs such as *to have*, *to like*, *to find*, and additionally one instance of the modal verb *can*. Thus the following example seems to be the typical representative of two-word repeats:

(6) | and . **{it was}** . **it was** very interesting :: because :: although I've spent a lot of time at boarding school so away from my parents . :: I'd never been to America before . |

(EN051, AS-Unit 2)

Looking at the learners' other structures which involved a two-word repeat, the present analysis showed 4 instances which combined a preposition with a determiner. The determiners were either definite articles or possessive pronouns. The last example featured a conjunction combined with a personal pronoun, as can be seen in the following example:

(7) | so just when I saw the city :: **{I} I** just thought :: well it's not as amazing ::  
**{as I} as I** imagined it so . |

(CZ006, AS-Unit 5)

On the other hand, native speakers' two-word repeats did not show any preference for a specific combination of words as there were nouns, verbs, pronoun or prepositions to be found. Furthermore, one of the examples has even exceeded the boundaries of subject and predicate and included the determiner of the following object. It is noteworthy that the verb has been contracted and thus treated as one word in accordance with Biber et al. (1999: 1061).

Compare:

(8) | yeah yes :: **{there's a} there's a** lamppost :: <laughs> so everyone **{stands}** .  
**stands** behind it <overlap /> <laughs> |

(EN020, AS-Unit 21)

In the light of these findings, two-word repeats do not appear to be entirely dissimilar to one-word repeats as they mainly comprise function words and frequent verbs. When comparing the two subcorpora, we may claim that Czech advanced speakers of English do not deviate from the native speaker norm as they most frequently repeat combinations of pronouns and verbs at the beginning of an utterance. Concerning other two-word structures which might be subjected to reiteration, speakers of LINDSEI\_CZ showed preference for prepositional combinations which may support Götz's claims (2013: 108) for learner's overuse of repeats of prepositions. Nevertheless, the sample of two-word repeats as extracted

from the two subcorpora is considerably small, therefore no firm conclusions can be drawn and further research is required.

## **5 Discussion of the findings**

### *5.1 Czech advanced speakers in comparison with other L2 speakers*

Since the main purpose of the present thesis was to shed more light on the use of the selected performance phenomena by non-native speakers, i.e. by advanced speakers of English who have received substantial education primarily due to their university studies, it is therefore important to compare our findings to previous research in the area. As has been hinted at in the methodological chapter, the comparative research of native and non-native speakers has not provided many studies, hence our main point of reference was the fairly complex research of Sandra Götz (2007, 2013). She based her work on the German component of LINDSEI (LINDSEI\_GE henceforth) alongside with LOCNEC. Although being concerned with far larger number of spoken (inter)language features, her study has yielded several findings which can be compared with those of the present thesis and which may illustrate some tendencies that L2 speakers possibly share despite their different L1 background.

As far as filled pauses are concerned, speakers of both LINDSEI\_CZ and LINDSEI\_GE have been found to significantly overuse this particular performance phenomenon. Furthermore, a similar finding has been obtained when comparing FPs usage within clauses or components. It needs to be however pointed out that the difference between FPWCL and FPWCON has not been specifically described by Götz (2013) and therefore the

approach to assigning a filled pause to a particular section of an utterance may differ. When concluding her findings, Götz (2013: 113) mentions “clause-internal [or] constituent-internal” filled pauses which suggests that the methodology presented in the section 3.2.1. may not correspond to the one applied to the German-speaker data. In spite of this possible dissimilarity, both Czech and German advanced speakers were found to overuse filled pauses in their speech which moreover highlights that such a tendency is particularly strong for learners. Hence, filled pauses seem to be a very prominent means non-native speakers use in order to deal with online planning pressure.

Concerning repeats, Czech advanced speakers of English have shown no deviance from the native speaker norm as they mostly repeated function words thus confirming the assumption that the more frequent a word is, the more likely it is to be repeated (Biber et al., 1999). Taking into account the overall frequencies, the present analysis have found Czech speakers to overuse repeats, yet not significantly. On the contrary, speakers of LINDSEI\_GE have been found to underuse them. The reason for such a different result may lie in the small sample of the present thesis since the speakers were chosen according to their use of filled pauses not repeats. Furthermore, the contradictory results may be attributed to the different taxonomies the present analysis and Götz’s use; for the latter did not limit potential words to be repeated, whereas Götz (2013) has taken over Biber et al.’s one (1999).

Such an approach have revealed that two-word repeats seems to play a minor yet considerable role in dealing with planning pressure. Consisting mainly of function words and relatively frequent verbs, they were represented by similar proportion both in Czech and English subcorpora showing a tendency to primarily combine personal pronouns and verbs. The only aspect where Czech advanced speakers of English deviated from native speakers is their frequent use of combinations introduced by prepositions. Such a finding correlates with

the expected overuse of repeated prepositions alone in the learners' corpus, which have been confirmed. Therefore, we may say that two-word repeats do not differ from one-word repeats in what kind of word is repeated but in the length of the sequence. Therefore, both native speakers and learners have the opportunity to buy themselves even more time to plan their forthcoming utterance. Nevertheless, these claims have to be attested in a larger sample to gain credibility.

Similarly to Czech advanced speakers of English, speakers of LINDSEI\_GE have been found to overuse repeats of prepositions. When we consider the usually mid-clausal or end-clausal position of prepositions, we may come to a similar conclusion as with filled pauses: learners need to plan their utterance not only at the beginning of an utterance but require further planning stages before reaching the end of a clause. Hence repeating a preposition help them to decrease planning pressure.

Lastly, we have observed that speakers of LINDSEI\_CZ have showed significant underuse of subject-verb contractions. Similarly, German advanced speakers of English have been found to follow the same tendency. Furthermore, their speech has showed a lack of these structures in general contributing to the low index of their repeats. What may be deduced from such findings is the fact that learners seem not to have internalized this feature otherwise very common in native speech.

The third main area of interest of the present thesis concerned the occurrence of further performance phenomena in close proximity to both filled pauses and repeats. Comparing the two, repeats proved to cause greater difficulties to learners as they were surrounded by a larger proportion various performance phenomena than filled pauses. Therefore, filled pauses do not seem to have the same status as repeated words, i.e. they do not require a speaker to take into account grammatical and lexical aspects of the context and

may therefore be treated as verbal signals of delay. Nevertheless, learners of LINDSEI\_CZ showed undeniable preference for further performance phenomena in the vicinity of filled pauses or repeats when compared to native speakers. Such a tendency has been also observed by Tavakoli (2011) whose research was focused on mid-clause pausing. As she points out (2001: 75-76), native speakers have not been found to follow any regular pausing pattern whereas non-native speakers have felt the need to pause close to repetitions, replacements and reformulations. Although the present analysis took a rather quantitative point of view in this respect, i.e. it did not take into account the exact position of the further performance phenomenon, learners have been observed to overuse filled pause within constituents which, to a certain degree, may be in accordance with Tavakoli's findings (2011).

To conclude, it is beyond doubt that speaking a foreign language represents a strenuous exercise even for non-native speakers who have reached a considerably high level of proficiency. As opposed to native speakers, learners prefer those performance phenomena which do not require retrieval of further lexis as reflected mainly in their overuse of filled pauses. Such behaviour implies limited repertoire of those fluency-enhancing strategies that require further language planning, for instance discourse markers or formulaic sequences. For a native speaker, these prefabricated chunks of language, be it a short phrase or a relatively complex clause, are readily available and require little effort as regards planning. Hence, learners seem to lack the automatization and promptness of native speakers. As a result, learners' avoidance of formulaic sequences should be reflected in their instruction in a language classroom. The following section will discuss various approaches to teaching spoken grammar and offer possible methods of raising the awareness of various fluency-enhancing strategies.

## ***5.2 Implications for teaching***

The results of the above presented study have brought several insights, yet from the pedagogical point of view, the main outcome is that the speech of advanced speakers of English cannot be considered equivalent to the native one. Although they have received substantial education and their language competence complies with the requirements of CEFR<sup>37</sup>, they should be treated as learners nevertheless. As Cobb aptly points out: “It appears that even advanced learners are unlikely to discover very quickly on their own all of the relevant features of a second language that make it native like” (quoted in Götz, 2013: 139).

As the findings of the present study suggest, learners are not aware of the fact that they give preference to selected performance phenomena and therefore their knowledge of native-like strategies should be actively supported by awareness-raising activities. In this relation, Timmis (2005) suggests including various noticing tasks that would alert learners’ attention to the features of spoken language that native speakers naturally use. As he points out (2005: 120), such tasks would allow learners to familiarize themselves with these strategies yet would not expect to adopt them since “it is at least questionable whether we want learners to produce these forms at any stage.” Furthermore, he stresses the fact that the rules governing spoken language are often very intricate and cannot be comprehensively communicated. On the other hand, advanced learners should be aware of the fact that spoken grammar has a lot of features that are fundamentally distinct from written grammar. Hence, noticing tasks could be an effective means of conveying these differences without the necessity to formulate rules. According to Timmis (2005: 121), the main aim of such activities consist in “encourag[ing]

---

<sup>37</sup> C2 Level: A speaker can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in the most complex situations.

learners to compare their expectations of native-speaker English and the reality of native-speaker English.”

A similar approach that supports implicit teaching of spoken language features is advocated by McCarthy and Carter (1995). Furthermore, they highlight the importance of adjusting classroom methodology to the findings of spoken language analysis. In the above mentioned article (McCarthy and Carter, 1995: 217), they propose a revised methodology attempting to replace the traditional Presentation-Practice-Produce model. Being nicknamed ‘three Is’, such methodology would first focus on *illustrating* the use of authentic data in their natural use and context. Secondly, through *interaction* stage, learners’ awareness would be aimed at observing and discussing features that accentuate interpersonal uses of language and the negotiation of meanings. Thirdly, learners would be encouraged to take a step further and “draw conclusions about the interpersonal functions of different lexico-grammatical options, and develop a capacity for noticing such features” in the *induction* phase (McCarthy and Carter, 1995: 217).

In this context, Götz (2013) suggests using LOCNEC for noticing tasks at the university level since the corpus represents a ready-made resource which can be used to illustrate various features of spoken language and can be thus incorporated into language analysis courses. The authentic data gathered from native speakers with similar background can serve as a source of “topic-related chunks and phrases [which] may also turn out to be very useful for the learners and their anticipated future communicative encounters” (Götz, 2013: 141). These would include discourse markers, fixed phrases, formulaic sequences or other lexical chunks that would enable learners to gain more time for planning. Taking into account Tavakoli’s study (2011) which proved that learners are less prone to pause in the middle of formulaic sequences, encouraging advanced learners to actively make use of these

prefabricated sequences of language could exponentially increase their temporal fluency and allow them to focus more on the content of their utterance rather than its form.

In contrast to university-level courses which can be updated relatively easily in order to include recent scientific research findings, ELT materials for general public have lagged behind and as a result, the language classroom has been “producing speakers of English who can only speak like a book, because their English is modelled on an almost exclusively written version of the language” (McCarthy and Carter, 1995: 207). Features of spoken language are only slowly entering the syllabi of textbooks and in order to overcome such shortcomings, teachers can reach for specialized resource materials designed for advanced learners. For instance, Alan Maley’s *Advanced Learners* (2009) provides some useful and prepared activities that help learners notice the differences between spoken and written grammar<sup>38</sup> frequently referring to LGSWE and Carter and McCarthy’s grammar (2006). Furthermore, it also encourages them to use pre-fabricated construction and highlights their frequent use in spoken texts.<sup>39</sup>

To conclude, it is beneficial to acquaint advanced learners with the various features of spoken grammar, especially with those that are extremely common. In our case it would concern subject-verb contractions which proved to be very frequent in the English-speaker corpus but which learners were shown to markedly underuse. Nevertheless, the degree of conscious acquisition of fluency-enhancement strategies and their consequent active usage remain questionable. As Götz herself points out (2013: 127), some of them may be too subconscious to teach (as in the case of repeats for instance). Moreover, we may consider to what extent a speaker’s L1 influences the make-up and use of fluency-enhancing strategies

---

<sup>38</sup> 4.2 Spoken grammar, p 52-54.

<sup>39</sup> 3.9 Fixed phrases, p 49-50.

and whether s/he does not transfer them to L2. For it is possible that some features of spoken language may be particular to each language and then when speaking another one, these might be applied to L2 at least to a certain extent. A research of such scope however exceeds the focus of the present thesis and will require large-scale research in the future.

## **6 Conclusion**

The aim of the present thesis was to analyze selected performance phenomena, namely filled pauses and repeats, in the speech of native speakers and Czech advanced non-native speakers of English. The data have been excerpted from the Czech component of a multilingual corpus LINDSEI and its accompanying corpus LOCNEC. Both selected phenomena frequently occur naturally in native speech and do not impede understanding. Their main function is to decrease online planning pressure and allow a speaker to gain more time for planning. Our main interest lay in the relative frequencies and distribution of these two performance phenomena in the learner and the native corpus. Czech advanced speakers were expected to experience more difficulties when planning their utterance since they need to convey their message in a language that is not their mother tongue. As regards filled pauses, their exact position in an utterance has been the main focal point of the analysis and

repeats were analyzed in detail with regard to the repeated word or cluster of words they included. Moreover, further performance phenomena occurring in the vicinity of filled pauses and repeats were examined in order to highlight further planning difficulties that speakers may experience. These would be employed to tackle continuing planning pressure and enable speakers gain even more time to formulate their utterances.

As far as filled pauses are concerned, the frequency analysis showed that Czech advanced speakers of English significantly overuse this particular fluency-enhancing strategy in comparison with native speakers. To be precise, Czech advanced speakers were found to require a filled pause roughly for every eighteenth word (5.842 phw) they produced, whereas English native speakers were in need of one only approximately for every thirtieth word (3.337 phw). Since the present analysis took a rather qualitative view due to the size of its sample, a further analysis revealed that the learners overused filled pauses both at higher structural boundaries (FWPCL) as well as lesser ones (FPWCON). Filled pauses occurring within constituents especially highlighted the increased amount of planning pressure in the learner's speech. Such findings confirm the results of larger-scale studies on English learners with different L1 backgrounds - for instance German speakers have been found to significantly overuse filled pauses as well (see Götz 2013). Hence, we may assume that such verbal behaviour may not be peculiar to Czech speakers only and may represent a feature that is common to all non-native speakers. However, such a claim requires further research carried out on learners with various L1 backgrounds.

If we take into account the sequences of words that speakers repeated in the course of their utterance, the analysis showed overuse in the case of non-native speakers, similarly to filled pauses. However, the comparison of the two corpora displayed no significant result and therefore no conclusions may be drawn from the findings. The insufficiency of data has been

most likely caused by the method employed in order to select speakers for the present analysis. If it were not the ratio of filled pauses which was the main indicator for the selection of speakers, the analysis might have showed different results. On the other hand, due to the presence of two phenomena one had to be designated as the decisive one. Nevertheless, the qualitative perspective has yielded some interesting results. In accordance with the findings of Götz (2013), Czech advanced speakers of English have been found to underuse repeats of subject-verb contractions to a very significant degree and similarly to German advanced speakers of English, Czech learners showed a very strong tendency to overuse repeats of prepositions. Even more interestingly, the present analysis has discovered that a fairly prominent portion of all repeats accounted for sequences including two words, a finding which has not been discussed in the previous research in this area. Following a detailed analysis, two-word repeats have proved to be mostly composed of function words and frequent verbs, thus complying with the assumption that the more frequent a word is the more likely it is to be repeated. When comparing the two subcorpora, native and non-native speakers seemed to be using two-word repeats in a similar fashion. However, Czech advanced speakers of English have shown a slight inclination towards repeating a two-word sequence introduced by a preposition which furthermore highlights their planning difficulties amidst an utterance.

When assessing the environs of filled pauses and repeats in search of other performance phenomena, Czech advanced speakers of English have shown an increased tendency to employ these in combination. It comes as no surprise since all of the previous findings suggest that learners feel the need to use hesitation markers even later on in an utterance which gives evidence to their planning difficulties. Furthermore, the present analysis has discovered that repeats place higher demands on the learners' planning skills

since repeated sequences were more frequently followed by other performance phenomena in contrast to filled pauses. Consequently, we may wish to highlight the ambiguous status of a filled pause. Although some authors (Clark and Fox Tree, 2002) claim that they should be considered words, these findings suggest the contrary.

All in all, the present thesis has gathered more information that contributes to the description of an advanced learner. It is beyond doubt that reaching a high proficiency in English does not equal to using the language in a native-like manner. As regards the two performance phenomena under scrutiny, Czech advanced speakers of English have showed that they rely on including filled pauses and repeats in their speech to a substantial degree. Such findings highlight the permanent presence of planning pressure which seems to persist further on in an utterance of even advanced learners. Moreover, the findings of the present analysis correspond to a considerable degree to those of speakers with different L1 backgrounds and therefore further research that would test the above mentioned assumptions is desirable. It needs to be however emphasized that a detailed analysis of spoken language is inevitably time-consuming and requires considerable technical skills.

## 7 References

- Biber, D. et al. (1999) *Longman Grammar of Spoken and Written English*. Harlow; New York: Pearson Education ESL.
- Carter, R. and M. McCarthy (2006) *Cambridge Grammar of English: A Comprehensive Guide*. Cambridge: Cambridge University Press.
- Chambers, F. (1997) "What Do We Mean by Fluency?". *System* 25 (4), 535–44. doi:10.1016/S0346-251X(97)00046-8.
- Clark, H. and J. E. Fox Tree (2002) "Using Uh and Um in Spontaneous Speaking". *Cognition* 84 (1), 73–111. doi:10.1016/S0010-0277(02)00017-3.
- Clark, H. and T. Wasow (1998) "Repeating Words in Spontaneous Speech". *Cognitive Psychology* 37, 201–242

- Corley, M., L.J. MacGregor and D.I. Donaldson, (2007). "It's the way that you, er, say it: Hesitations in speech affect language comprehension". *Cognition*, 105, 658-668.
- Council of Europe (2001) *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Cutting, J. (2006) "Spoken grammar: Vague language and EAP", in R. Hughes (ed.) *Spoken English, TESOL and Applied Linguistics: Challenges for Theory and Practise*, pp. 159-181. Houndmills: Palgrave Macmillan.
- Fischer, K. (1999). From cognitive semantics to lexical pragmatics: the functional polysemy of discourse particles. Unpublished PhD dissertation, Universität Hamburg.
- Foster, P. and P. Tavakoli (2009) "Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity". *Language Learning* 59(4), 866-896.
- Foster, P., A. Tonkyn, and G. Wigglesworth (2000) "Measuring Spoken Language: A Unit for All Reasons". *Applied Linguistics* 21 (3), 354–75. doi:10.1093/applin/21.3.354.
- Götz, S. (2007) "Performanzphänomene in Gesprochenem Lernerenglisch: Eine Korpusbasierte Pilotstudie". *Zeitschrift Für Fremdsprachenforschung* 18 (1), 67–84.
- Götz, S. (2013) *Fluency in Native and Nonnative English Speech. Studies in Corpus Linguistics, volume 53*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.
- Gráf, T. (2015) *Accuracy and fluency in the speech of the advanced learner of English*. PhD Thesis, to be published. Charles University in Prague.
- Grice, H. P. (1975) "Logic and conversation", in P. Cole and J. L. Morgan (Eds.) *Speech acts Syntax and semantics*, pp. 113–128. New York: Seminar Press.
- Halliday, M. A. K. (1989) *Spoken and written language*. Oxford: Oxford University Press.
- Hunt, Kellogg (1965) *Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3*. Champaign, IL: National Council of Teachers of English.
- Hunt, K. (1970) *Syntactic Maturity in School Children and Adults*. Monograph of the Society for Research into Child Development.
- Kjellmer, G. (2003) "Hesitation. In Defence of Er and Erm." *English Studies* 84 (2), 171–98.
- Lennon, P. (1990) "Investigating Fluency in EFL: A Quantitative Approach\*". *Language Learning* 40 (3), 387–417. doi:10.1111/j.1467-1770.1990.tb00669.x.
- Lennon, P. (200). "The Lexical Element in Spoken Second Language Fluency", in H. Riggenbach (ed.) *Perspectives on Fluency*, pp. 25–42. Ann Arbor: University of Michigan Press.

- Levelt, W. J. M. (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Maclay, H. and C.E. Osgood (1959) "Hesitation phenomena in spontaneous English speech". *Word* 1, 19-44.
- Maley, A. (2009). *Advanced Learners*. Oxford: Oxford University Press.
- McCarthy, M. and R. Carter (1995) "Spoken grammar: what is it and how can we teach it?" *ELT Journal* 49, 207–218.
- Quirk, Randolph, S. Greenbaum, Sidney, G. Leech, and J. Svartvik (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Riggenbach, H. (1991) "Toward an Understanding of Fluency: A Microanalysis of Nonnative Speaker Conversations." *Discourse Processes* 14, 423–41.
- Rühlemann, C. (2006) "Coming to Terms with Conversational Grammar: 'Dislocation' and 'Dysfluency.'" *International Journal of Corpus Linguistic* 11 (4), 385–409.
- Tavakoli, P. (2011) "Pausing patterns: differences between L2 learners and native speakers". *ELT Journal* 65, 71–79. doi:10.1093/ELT/ccq020
- Timmis, I. (2005) "Towards a framework for teaching spoken grammar." *ELT Journal* 59, 117–125.
- Yuan, F. and R. Ellis (2003) "The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production". *Applied Linguistics* 24 (1), 1-27. doi: 10.1093/applin/24.1.1

## 8 Sources

Gilquin, G. et al. (eds) (1995– ) The Louvain International Database of Spoken English Interlanguage. Centre for English Corpus Linguistics at UCL in Louvain.<sup>40</sup>

De Cock, S. (2004) The Louvain Corpus of Native English Conversation. Centre for English Corpus Linguistics at UCL in Louvain.

---

<sup>40</sup> The present thesis have used only the Czech component of LINDSEI which was not available on the CD ROM published in 2010.

## **9 Resumé**

Předkládaná diplomová práce se zabývá vybranými prvky řečového managementu, a to konkrétně vyplněnými pauzami a opakováními. Tyto dva prvky se běžně vyskytují v mluveném projevu a nijak nenarušují tok komunikace. Naopak představují praktický nástroj na to, kterým se mluvčí vyrovnává s tlakem, jenž je na něj vyvíjen ve chvíli, kdy musí

zároveň plánovat a produkovat řeč. Nerodilý mluvčí má z tohoto pohledu situaci ztíženou ještě více než mluvčí rodilý, neboť jeho produkce není tak zautomatizovaná. Očekává se tedy, že nerodilý mluvčí bude tento problém řešit použitím prvků řečového managementu. Práce je rozdělena do několika částí - po krátkém úvodu následuje část teoretická, která zasazuje vyplněné pauzy a opakování do širšího kontextu mluveného slova a nastiňuje možné rozdíly mezi mluvčími rodilými a nerodilými. Poté následuje metodologická kapitola, která přibližuje zvolenou metodiku přístupu k datům a data samotná. Hlavní analytická část představuje výsledky výzkumu, které byly získány na základě korpusů LINDSEI a LOCNEC. Tyto poznatky jsou posléze diskutovány v širším kontextu výzkumu těchto prvků a jejich možné pedagogické důsledky jsou představeny v další části.

Teoretická kapitola nejprve nastiňuje stručnou charakterizaci mluveného jazyka, která naznačuje jak problematická je definice jeho základní stavební jednotky a překládá různé přístupy, které se ve výzkumu používají (např. Hunt 1975, Biber et al. 1999, Foster et al. 2000). Dále pak zasazuje opakování a vyplněné pauzy do teoretického rámce, který byl vytvořen Sandrou Götz (2007, 2013) pro účely jejího výzkumu německých nerodilých mluvčích angličtiny. Její studie slouží jako referenční bod pro tuto diplomovou práci – část této studie totiž replikuje. Teoretická část pak také poukazuje na úskalí terminologie pro mluvený jazyk a okrajově nastiňuje psychologické pozadí produkce opakování a vyplněných pauz. Na závěr překládá získané poznatky o rodilých a nerodilých mluvčích angličtiny v souvislosti s prvky řečového managementu. Ty pak slouží jako hlavní zdroj pro formulaci hypotéz testovaných v empirické části.

Metodologická kapitola nejprve představuje využitý materiál pro účely této diplomové práce, který byl získán ze dvou propojených zdrojů sestavených na univerzitě v Lovani. Prvním materiálem je multijazykový žákovský korpus LINDSEI, a to sice jeho český

komponent, z něhož byli vybráni 3 reprezentativní mluvčí. Druhý korpus, který byl využit pro účely této práce je LOCNEC, který byl sestaven jako referenční korpus k LINDSEI a který obsahuje nahrávky a přepisy rodilých mluvčích. Těm byly zadány identické úkoly jako nerodilým mluvčím v LINDSEI a díky tomuto přístupu jsou korpusy porovnatelné. Z LOCNECu byli opět vybráni 3 reprezentativní mluvčí. Vzhledem k tomu, že předmětem zkoumání je mluvený jazyk, bylo tedy zapotřebí stanovit základní jednotku. Pro tyto účely byla využita teorie AS-Unit autorek Fosterové, Tonkynové and Wigglesworthové (2000), která bere v potaz nejenom syntaktickou strukturu promluvy, ale také její prozodie. Metodologická kapitola pak dále prezentuje shrnutí poznatků z pilotní studie, které poukázaly na některé zvláštnosti popisu mluveného jazyka. Závěrem je stanoveno šest hypotéz, které jsou posléze testovány na vybraných mluvčích.

Výsledky analýzy jsou nejprve prezentovány společně pro oba vybrané prvky - empirická kapitola podává informace o celkových frekvencích vyplněných pauz a opakování jak v korpusu rodilých mluvčích, tak v korpusu žakovském. Podobně jsou prezentovány poznatky o výskytu dalších prvků řečového managementu v těsné blízkosti opakování a vyplněných pauz. Následně tato kapitola pohlíží na oba prvky odděleně a zkoumá jejich specifické vlastnosti v rámci obou korpusů. Co se týče vyplněných pauz, pozornost se vztahuje na jejich distribuci v rámci promluvy, a to jak na úrovni klauzí, tak i na úrovni větčných členů. Samostatná podkapitola o opakování se zabývá slovy nebo slovními kombinacemi, které daní mluvčí produkovali za účelem získání více času pro plánování jejich promluvy.

Následná diskuze porovnává získané poznatky této práce s výsledky předchozích výzkumů na podobné téma. Ukazuje se, že nerodilí mluvčí (a to nejenom čeští) mají výrazně větší problémy s tím, jak se vypořádat s tlakem, který je na ně během produkce vytvářen. To

pak reflektuje vyšší frekvence vyplněných pauz a jejich častý výskyt i na nižších syntaktických úrovních, tedy hlavně na začátku nebo uvnitř větného členu. Co se týče opakování, překládaná diplomová práce potvrdila domněnky o tom, že právě onen zvýšený tlak na plánování promluvy se projevuje častým opakováním předložek, což poukazuje zpět k vyplněným pauzám uprostřed klauzí. Naopak, nerodilí mluvčí projevují tendenci neopakovat stažené tvary, což je dáno jejich nízkou frekvencí v rámci jejich promluvy. Analýza také zjistila, že obě dvě skupiny opakují nejenom samostatná slova ale i dvojice slov. S ohledem na typ slov, v obou případech byla nejčastěji zastoupena slova gramatická, popřípadě frekventovaná slovesa. I zde se projevila náchylnost k opakování předložkových vazeb u nerodilých mluvčích, což dále poukazuje na obtíže s plánováním, se kterými nerodilí mluvčí potýkají i po začátku promluvy.

Analýza dalších prvků řečového managementu, které se vyskytovaly v těsné blízkosti vyplněných pauz a opakování, prokázala, že nerodilí mluvčí mají tendenci tyto prvky řetězit, a získávat tak více času na plánování. Závěr práce nastiňuje možné dopady na výuku a vnímání pokročilých mluvčích angličtiny a překládá několik přístupů, které mohou být uplatněny v jazykové třídě. Otázkou však nicméně zůstává, do jaké míry je možné nerodilé mluvčí strategie mluvčích rodilých naučit a zda je to vůbec jejich cílem. Jasně však je, že nerodilí mluvčí se i na pokročilé úrovni musí vypořádávat s různými obtížemi během plánování své promluvy a použití opakování nebo vyplněných pauz na jejím začátku je nedostačující.

## 10 Appendix

The appendix contains all AS-Units that were examined for the purposes of the present analysis. Each table represents one speaker as referenced in the heading, first the data for filled pauses are presented starting with Czech advanced speakers of English in chronological order based on LINDSEI coding. AS-Units containing filled pauses of English native speakers then follow. The same pattern is applied to repeats.

### Filled pauses

CZ006	AS-Unit
1	I decided :: to talk about . <b>(erm)</b> . my visit to another country :: which has impressed me
2	well I think :: it's . partially :: because I study English and American Studies :: and I . love the country :: and <b>(eh)</b> . I love the people there :: so it's just really <b>(eh)</b> . >> I think << a heart thing <overlap /> really .
3	but <b>(eh)</b> . well I it's been >> I think << two years ..
4	and <b>(erm)</b> . we we went by bus :: which is . a little annoying :: because it was a long long way ..
5	<b>(erm)</b> we we went to: France :: took the ferry on the other side . :: then went straight to Wales :: and then from there another ferry to Ireland .
6	we spent a day in Dublin . <b>(eh)</b> .
7	but <b>(eh)</b> Dublin at night was really beautiful :: or rather it was . something different
8	and there was <b>(eh)</b> live music :: playing there :: and people and dancing and drinking and singing :: and it it was . >> I think . << nice experience
9	<b>(erm)</b> well after visited Dub= after we visited Dublin :: we went to: <b>(mm)</b> . we went west to the other part of Ireland :: and <b>(eh)</b> . we spent a night or two >>I think<< in Galway . <b>(mm)</b> . (P=0,85)
10	and from Galway we <b>(er)</b> took . a trip or two . <b>(eh)</b> to: the mountain of Saint Patrick . :: which is a kind of . holy place or :: how should I . say that . (P=0,9)
11	and <b>(em)</b> ma= many people go there to the mountain :: and go to the top :: and <b>(em)</b> . have it as a sort of like . I don't know .
12	Muslim people go to Mecca :: so <b>(em)</b> >>I don't know<< some Irish people go there
13	<b>(eh)</b> at first it was really nice the weather was nice
14	but then it got . the weather got quite terrible :: and started raining :: so it was .. quite exciting going to the top :: because the way . was quite rusty . (P=0,45) :: and <b>(eh)</b> . not very . not very pleasant really
15	. <b>(em)</b> . so after we . visited this . mountain of Saint Patrick . :: we went <b>(eh)</b> to: . >>I think<< Connemara is the national park there
16	<b>(eh)</b> . we we had some nice time in Galway :: and then we visited Limerick :: which is . not not very far from Galway . (P=0,83)
17	and <b>(em)</b> there was . quite a nice castle there and things :: so we spent a day in Limerick .. (P=1,1)
18	and then <b>(eh)</b> we returned <b>(eh)</b> . to: I think :: Sixmilebridge the village . was named .

CZ006	AS-Unit
19	(eh) . the the stay in (eh) Sixmilebridge was quite nice .
20	the: . lady :: who was organizing the trip said :: that the people there . still remember us :: and the (mm) when we visited :: because we (eh) we went to some local pub :: and there was a band playing also :: and we joined :: and sang with them :: and . (P=0,8)
21	(eh) and . yeah they said :: that . it was the most lively evening <starts laughing> :: probably they had there for for years <stops laughing> <overlap />
22	and when I was there with a friend of mine :: we . (eh) spent some time sitting on a bench in a park :: and . watching other people :: and just enjoying being there <laugh>
23	and the= there came (eh) . gentlemen in suits and wi= with (eh) . (erm) cases and things :: and they . came to the park sat on the bench then took a sandwich out of the case :: and ate it :: and then just .. (eh) <XX> went back :: and it was really nice
24	I think it was :: when I was really young (erm) .
25	and (er) I really loved the music :: and I wanted :: to know what they sing about :: so I found the lyrics :: and . tried :: to translate them :: and understand what they're singing about . (P=0,6)
26	. (er) . I think :: it's one of the main reasons :: why people don't speak very well English here
27	but . she could understand :: because she . (erm) . knew the sound of the language <overlap />
28	but (em) . I actually . do experience a sort of .. teaching received at a grammar school :: because I teach . grammar school kids
29	and (eh) when I checked . the . grammar topics :: they had in a textbook :: I say hey we we did that twice already :: so why is it here again . (P=0.58)
30	and it wasn't really . difficult or . something (eh) .
31	but yeah back to my . my experience then at (eh) . <lip sound> (er) secondary school ..
32	and (eh) we didn't change the teacher throughout all . eight years >>I think<< <overlap />
33	(eh) . it was <foreign> Gymnázium Jaroslava Heyrovského </foreign>
34	. (eh) well
35	and you go by (eh) by the underground :: take (eh) . one one stop to <foreign> Můstek </foreign> . :: and change the trains :: and it's at a stop . <foreign> Hůrka </foreign>
36	. (eh) I . I couldn't decide :: which school to choose .
37	yeah there were just things :: I didn't like :: such as (eh) chemistry
38	<B> so (eh) . but (eh) then I . just thought :: that I really like . the language and . the literature . as well . :: so I thought . :: I should give it a try <laugh>
39	(eh) . well . maths <laughs>
40	(eh) I actually teach maths as well now
41	. but (er) I didn't really want :: to . go to . mathematical schools
42	.. so (eh) . there's a lady . ::possibly . either the lady wanting to have a picture of herself :: or or the guy just trying to . perhaps hit on the lady :: by drawing a picture of her .. (P=1,1)
43	. but she's angry probably yelling at the painter :: that . it's not me :: it doesn't look like me . :: so . she she wants him :: to . (er) . redo the picture ..
44	(eh) then the painter . happily .. is re= repainting the picture . (P=0,75)

CZ034	AS-Unit
1	(eh) they <starts laughing> no not silly <stops laughing>
2	(er) when I was in America :: there was (eh) a joke going around :: and the people were calling me silky <overlap /> . like the material silk silky
3	but (eh) . it's actually my English nickname is derived from the Czech one :: because (er) in Czech (er) my friends and family call me . <nickname of the interviewee>
4	and (er) therefore I told everybody in (erm) America :: that (eh) they can call me <nickname of the interviewee>
5	yeah I tried to explain :: that it's not S[es] <overlap /> at the beginning :: but it's (eh) <a native sound> [ts] </a native sound> as in pizza for example :: but
6	(er) I have chosen the topic number two . (er) :: that is a country you have visited :: which has impress= . impressed you . :: and (eh)
7	the country would be Germany . :: and (eh) I would probably start . (er) . with by telling you :: how I started :: (er) learning German
8	(er) so there was (eh) .. there was nothing as (er) as (eh) as an excitement to tell the truth at the beginning
9	(er) so there was (eh) .. there was nothing as (er) as (eh) as an excitement :: to tell the truth at the beginning
10	it was . (eh) it's usual in our country :: that the second language is German .
11	and: (er) .. in fact I was not very happy: about it
12	(er) but we couldn't choose :: so: it was simply like this
13	but when I look at it now: :: I am very happy :: that (erm) . I I learned it :: and (eh) .. (erm) .
14	so during (eh) my schooling days there was nothing special about it
15	but then . I: decided one summer :: to: (eh) go to Germany :: to improve my (eh) knowledge of the language :: and (er) I chose . (er) . Munich
16	. and that was (er) such an amazing experience
17	(er) I think :: that . part of that experience was of course my expectations
18	(eh) . I had been before: in America :: and therefore: I (eh) was (eh) very scared :: to go to another foreign country
19	in in America for me (erm) . with English (P=0.56)
20	and I was (eh) so surprised :: that (eh) . it was not actually so difficult
21	and surprised me :: and (eh) gave me a lot of confidence (P=0.51)
22	and (eh) one reason for this is probably . :: that (er) . (eh) surprisingly enough .. (er) German culture is way more similar to Czech one (er)
23	and (eh) . just (eh) it's those little things (eh) such as food . :: or: <laughing> or: . simply .. (er) it's very hard <overlap /> :: to describe it
24	and (eh) . just (eh) it's those little things (eh) such as food . :: or: <laughing> or: . simply .. (er) it's very hard <overlap />:: to describe it
25	but it's just this way of thinking (er) :: that I (eh) that it's really hard to describe .

CZ034	AS-Unit
26	but it's similar <b>(erm)</b> .. (P=1.4)
27	and of course <b>(eh)</b> part of this experience is . <b>(eh)</b> was my family :: and I was very lucky in having a great family :: <b>(er)</b> which took a great care of me: :: and <b>(eh)</b> took me: for example to Berlin as well for for two weeks :: and: <b>(er)</b> I had one week off: and: . (P=-0,8)
28	and just <b>(er)</b> . <b>(eh)</b> a lot of fun .
29	what I admired about about <foreign> München </foreign> and <b>(eh)</b> the[i:] area down south <b>(er)</b> was <b>(er)</b> . <b>(eh)</b> were also traditions <b>(erm)</b> .
30	<b>(eh)</b> as you probably know :: they have <foreign> Oktoberfest </foreign> :: which I am not very fond of
31	but <b>(er)</b> simply the fact :: that <b>(er)</b> when you <b>(er)</b> . <b>(erm)</b> .. accompany your child <b>(eh)</b> . <b>(eh)</b> to a school on your on his or her first day :: <b>(eh)</b> you dress up in the traditional clothes: <overlap />:: and:
32	<X> and even the the[i:] adults everybody . :: and actually it is it is weird :: when <b>(eh)</b> someone is not dressed up like that so:
33	<b>(er)</b> .. yeah the= they are really aware of their traditions
34	<b>(erm)</b> . I went to Dresden many many years ago
35	but <b>(eh)</b> . that was that was <b>(eh)</b> <b>(eh)</b> actually also very nice
36	we were just the pupils at <b>(eh)</b> at a high school :: and .. <b>(eh)</b> again they were they were so similar <b>(eh)</b> to us
37	but <b>(eh)</b> to tell the truth :: we spoke in English :: because that was better for all of us
38	<b>(erm)</b> and then I of course visited <b>(eh)</b> the[i:] Alps
39	<b>(eh)</b> we went with a<?> family we went on a couple of trips :: <b>(erm)</b> and actually we <b>(eh)</b> . accidentally we went on Sunday :: and what I witnessed there was . :: <b>(eh)</b> that at the . one of the . mountains <b>(eh)</b> . on on a top of one :: <b>(er)</b> there was a group of musicians :: who were playing musical instrument :: and it was . a part of a Sunday . mass
40	<b>(eh)</b> they told me :: that <b>(er)</b> it it didn't use to be like that <b>(em)</b> . twenty years ago
41	I would say so . I would say so . <b>(erm)</b>
42	<b>(er)</b> actually <b>(eh)</b> no :: it didn't end
43	there it was another another .. <b>(erm)</b> inspiration another . bit of inspiration . <b>(erm)</b> :: and: <b>(eh)</b> after I came back . <b>(erm)</b> after a year or two :: I was very lucky :: in . getting <b>(eh)</b> admitted <b>(eh)</b> for a job in a German newspaper :: because <b>(er)</b>
44	<b>(er)</b> but it was just okay
45	and <b>(er)</b> . <clears her throat> and I had to <b>(eh)</b> communicate mostly in Czech so: <b>(eh)</b>
46	of course I had to communicate wi= with my boss in German . <b>(er)</b>
47	<b>(eh)</b> writing rather than speaking
48	<b>(erm)</b> . so <b>(eh)</b> that was that :: and <b>(erm)</b> why why why is there even a German newspaper in our country <b>(erm)</b> ..
49	before the second world war and: . <b>(er)</b> . after that <b>(eh)</b> about three million of people were sent away from this country:
50	and <b>(er)</b> but some of them <b>(eh)</b> were not (P=0,6)

CZ034	AS-Unit
51	and: this newspaper was founded after the Velvet Revolution :: and: <b>(eh)</b> its goal is <b>(eh)</b> again :: to keep the sense of community within the remaining members
52	and also <b>(er)</b> . <b>(eh)</b> some . :: they are called <foreign> Sudetendeutschen </foreign> :: some <b>(er)</b> . <b>(er)</b> some <b>(er)</b> . <b>(erm)</b> order it even from Germany mo= again mostly from <foreign> Bayern </foreign> <b>(erm)</b> . :: so: it was very interesting and:
53	<b>(er)</b> . <b>(eh)</b> w= what do you mean <b>(eh)</b> now by <b>(er)</b> strong background
54	this this this <b>(er)</b> contact with Germany it . started . three years ago
55	so it . and that's another point :: why: <b>(eh)</b> I wanted :: to talk about this topic :: because I <b>(mm)</b> I never planned it .
56	and as <b>(eh)</b> as I told you before . :: I was not very happy :: that <overlap /> I had to study it at school
57	. and suddenly <b>(er)</b> . my life has taken me into this direction :: and I found it very <b>(eh)</b> . enriching
58	I I really think so :: and <b>(erm)</b> . <b>(er)</b> one of my friends :: who went on an on the[i:] Erasmus programme
59	of course <b>(eh)</b> these Erasmus programmes are full: of <b>(er)</b> various European nations :: and it was so curious <b>(er)</b> .
60	it's . and <b>(eh)</b> again something :: that I didn't expect
61	<b>(er)</b> that was <b>(er)</b> probably connected with the fact :: that . <b>(er)</b> I: . spent ten months <b>(er)</b> in the U S
62	and: <b>(er)</b> I .. <b>(er)</b> I dunno I well
63	but I would say :: that . I am simply <b>(er)</b> that kind of person :: who: is happy at the Faculty of Arts
64	in <b>(er)</b> literatures . <b>(er)</b> . and this kind of stuff
65	definitely <b>(er)</b> I appreciate so much :: even the fact :: that the . <b>(er)</b> our programme was divided into the bachelor .
66	part and the <b>(er)</b> master .
67	<b>(erm)</b> .. I I really appreciate the fact :: that I <b>(eh)</b> chose now only the[i:] linguistic branch
68	<b>(eh)</b> much more and
69	<b>(er)</b> .. I have a slight idea
70	<b>(er)</b> probably it would be something in the in the direction of of language teaching
71	<b>(erm)</b> . in . very . in general I like <b>(erm)</b> . <b>(er)</b> I like :: to work <b>(er)</b> with languages :: that are in contact :: so <b>(er)</b> .. I would not be opposed :: to trying translation for example
72	<b>(er)</b> . so: we will see <overlap /> we will see
73	(mhm) so <b>(er)</b> there is a man in the picture :: and this man is . <b>(erm)</b> painting .. a . a picture of a woman
74	<b>(er)</b> this woman is sitting on a chair :: and is looking in <b>(mm)</b> . a different direction ..
75	<b>(er)</b> then . the woman <b>(er)</b> . stands up :: and: she is looking at . the picture of herself .
76	<b>(er)</b> and sh= she is pointing <laughing> at her at her neck ..
77	she is probably not satisfied :: because in <b>(eh)</b> the subsequent picture <b>(er)</b> she is sitting again . <b>(er)</b> on the chair :: and: the painter continu= continues painting her :: and . <b>(er)</b> i= in the last picture <b>(eh)</b> it's probably . <b>(er)</b> .. in the exhibition (P=0,6)

CZ034		AS-Unit
78	and: <b>(er)</b> . the woman :: who was previously <b>(er)</b> the model . :: is showing <b>(er)</b> the portrait of herself to her <b>(er)</b> woman friends . :: who are looking at the picture	
79	<b>(er)</b> ... yeah the hair is different <b>(erm)</b> .. :: because in the second picture her <b>(eh)</b> hair is straight . :: while in the . third one :: it it has curls <b>(er)</b> . at the end . :: and: so probably what what the model said to the painter was :: that she . didn't like . :: what her hair looked like :: so: therefore he changed it	
80	<b>(er)</b> ... yeah the hair is different <b>(erm)</b> .. :: because in the second picture her <b>(eh)</b> hair is straight . :: while in the . third one :: it it has curls <b>(er)</b> . at the end . :: and: so probably what what the model said to the painter was :: that she . didn't like . :: what her hair looked like :: so: therefore he changed it	
81	<b>(erm)</b>	
82	oh: well <b>(eh)</b> really depends on what kind of a painter he is .	
83	<b>(er)</b> if he has to earn money by painting :: then he . couldn't say anything :: <b>(eh)</b> what he thought .	

CZ039		AS-Unit
1	<b>(eh)</b> . I think :: I can get over that	
2	<b>(er)</b> we've got . two departments :: one is <overlap /> department of the literature :: and <b>(er)</b> it's the	
3	their headquarters is <b>(er)</b> on the fourth floor	
4	<b>(er)</b> I was thinking about . <b>(erm)</b> . talking about a film :: I liked	
5	but then I realized :: it's much <starts laughing> easier :: to talk about <stops laughing> something :: we don't like really :: so I . <b>(er)</b> I . <overlap /> remembered	
6	<starts laughing> maybe <stops laughing> :: but <b>(er)</b> I think the case	
7	<b>(er)</b> it's because :: when you don't when you like something . :: <b>(er)</b> you just feel :: that it's good .	
8	obviously you should have reasons for <b>(er)</b> positive . emotions and positive opinions as well	
9	<overlap /> :: because I went to this festival . :: which takes place in Prague <b>(eh)</b> every summer	
10	<b>(er)</b> . it takes place in Prague castle .	
11	<b>(er)</b> it's called the . Shakespearean	
12	<b>(er)</b> I thought it was good idea :: to . <b>(er)</b> go :: to see a comedy instead of a tragedy .	
13	but . it turned out :: <b>(ehm)</b> that it <starts laughing> wasn't such a good idea <stops laughing> :: because . <b>(er)</b> it seems to me :: that <b>(er)</b> . nowadays :: when somebody wants to . <b>(er)</b> stage Shakespeare . :: <b>(er)</b> . the . the [i:] actors and the the director on the people all the people :: who . create the the performance . think :: they must change it somehow :: to make it att= more attractive for the <overlap /> for contemporary people and contemporary audience	
14	but <b>(er)</b> I saw . or I've seen <b>(er)</b> a very . nice performance of King Lear	
15	and <b>(eh)</b> I always thought :: that the tragedies were much <b>(er)</b> much more difficult to stage	
16	because they are . full of these grand . <b>(er)</b> words and grand visions and honour and stuff like that . :: and people . usually nowadays don't care much about that . (P=0,6)	
17	and . <b>(eh)</b> comedies on the [i:] other hand . are full of <b>(er)</b> sexual jokes	

CZ039	AS-Unit
18	and all this stuff . :: but . <b>(er)</b> somehow the case is :: that <b>(er)</b> tragedies
19	<b>(eh)</b> I've seen several tragedies by Shakespeare staged in several Prague theatres :: and they were . better than the comedies . by Shakespeare :: so this particular comedy . :: <b>(er)</b> where .. <b>(er)</b> when where . <b>(em)</b> <foreign> Tatiana VilhelmovÃ¡ </foreign> and other .. well . good . actors
20	play <b>(er)</b> . was a disaster for me :: because it was .
21	I think that Shakespeare is funny . <b>(er)</b> just <starts laughing> by himself <stops laughing> <overlap /> :: and
22	or contemporary audience to be . <b>(er)</b> amused :: so <b>(er)</b> I thought it . too much :: because they changed it into . a hilarious <b>(er)</b> . contemporary .. performance full of contemporary music <overlap /> :: and
23	<b>(er)</b> to go and see <b>(em)</b> Shakespeare's plays <overlap /> >> I suppose . so . <<
24	no I didn't <b>(er)</b> . :: because . <b>(eh)</b> when I bought the tickets :: I was preparing for my finals :: and I was dealing with Shakespeare :: and <overlap /> <starts laughing> I was like <stops laughing>
25	well <laughs> they are so great books <b>(er)</b>
26	so I went to the[i:] internet :: I booked <b>(er)</b> I booked Midsummer Night's Dream . :: and I was looking forward to <overlap /> seeing it
27	no: <b>(er)</b> no :: I've seen
28	<b>(er)</b> . I said <overlap /> I've seen
29	because the catharsis at the[i:] end was so <b>(er)</b> . overwhelming
30	I just <b>(er)</b> I was si= sitting there . :: and . <b>(er)</b> I think :: they . they did <b>(er)</b> great thing with the with the space on the stage :: because there was d= dark . :: and there was . <b>(er)</b> . suddenly . just <b>(er)</b> . the[i:] actors . playing
31	<b>(er)</b> . <b>(er)</b> <foreign> ježiš </foreign> .. :: who was her .. <starts laughing> :: the[i:] youngest daughter <stops laughing>
32	<b>(er)</b> not Goneril not Regan but . Cordelia <overlap /> yes so:
33	Cordelia just came <b>(er)</b> swinging
34	and it was just . out of the blue :: and <b>(eh)</b> everybody realized :: that she was dead
35	and they used . <b>(er)</b> the space of the stage well >>>I think<<<
36	and it was it was . well . maintained :: and it . I think :: it's quite difficult :: to . <b>(er)</b> to achieve this nowadays :: because people are just . so bored :: because they see . so much action in in movies <overlap /> :: so it's hard to
37	. <b>(er)</b> create some . some <b>(eh)</b> powerful . <b>(eh)</b> feelings in in in the theatre :: when you've . got only stage and <overlap /> some
38	<b>(er)</b> <overlap /> well
39	<b>(er)</b> I think :: the contrastive view is very <b>(er)</b> very interesting :: and I <b>(er)</b> . studied . both subjects <overlap /> as on my B A studies as well
40	<b>(er)</b> but also the the dep= both or all of the departments have different approaches :: so <overlap /> it's interesting
41	<b>(er)</b> to compare this :: and I think . :: the only disadvantage is that <b>(er)</b> .
42	well now I'm <b>(er)</b> at the[i:] . <b>(er)</b> English language <overlap /> :: so there is <starts laughing> no literature anymore <stops laughing> :: but
43	. we for example we had the phonetics <overlap /> <b>(er)</b> at the English . <b>(er)</b> studies

CZ039	AS-Unit
44	and then we had Mr . <name of teacher> <overlap /> at the <b>(er)</b> Czech programme
45	but it was <b>(er)</b> maintained very differently :: and I must say :: that the[ic] English department . <b>(er)</b> relied mo= more . on <b>(er)</b> <name of department> :: and I think :: it was better :: because we had this seminar .
46	but <b>(er)</b> on the Czech language and literature we had only the the lecture <overlap /> :: which was
47	and <b>(er)</b> . this the seminar was what was . the most useful <overlap /> . here
48	but <b>(er)</b> . no I miss it <starts laughing>
49	but but as many of my colleagues said <stops laughing> :: <b>(er)</b> the the . <b>(er)</b> the BA studies . somehow . <b>(er)</b> destroyed <starts laughing> our love for literature <stops laughing> a bit <overlap /> a bit :: because
50	>>you know<< when you have to do <overlap /> something :: when you have to read . <b>(er)</b>
51	when you have to read <b>(er)</b> . the Lord of the Flies . just during . three hours <overlap /> :: you don't get to enjoy it
52	but . >>you know<< it's <b>(er)</b> . it is an advantage as well :: because you can you can see thr= through it :: and ordinary people can't :: so . it's .
53	but <b>(er)</b> . of course the the studying literature opens your mind to <overlap /> several interpretations
54	well <b>(er)</b> currently I I work as a teacher of <b>(eh)</b> Czech language for foreigners and <b>(er)</b> English as well . (P=0,7)
55	but I must say :: that it's <b>(er)</b> a bit . <b>(er)</b> it's . great :: but <b>(er)</b> it takes lot of energy <overlap /> :: and
56	<b>(er)</b> I just . I don't know :: if I could . if I could do that for . whole of my life <overlap /> :: and if I could
57	because I . when I m= when myself when I was at sec= secondary schools :: I <b>(er)</b> just . had . some argument with other students :: because I just . couldn't understand <starts laughing> :: why they are <stops laughing> behaving the way <overlap /> or they were
58	and when I saw <b>(er)</b> this then . even
59	I just . <b>(em)</b> didn't . get on with my with people . as old as I'm
60	when I was <b>(er)</b> at the secondary school :: so I think it would be much worse even
61	but . when I came to to the university :: I was a bit disappointed about the . <b>(er)</b> about the things like lectures and seminars :: because seminars were usually great
62	but I just . can't get over the lecture stuff :: because it's . >>you know<< even . when the . topic is extremely interesting . :: <b>(er)</b> listen . to: . person's . talking for one hour and half without <b>(eh)</b> any activity from your from your own is extremely difficult <overlap /> :: and I
63	. <b>(eh)</b> I was . starting to get bored :: even if I was <overlap /> interested
64	I just think :: that . <b>(er)</b> it's . there is not enough space :: to work with the students
65	I think :: you can <overlap /> <b>(eh)</b> do it pretty well
66	because you've got this . great <b>(er)</b> .. management of the seminar :: but I think :: that's because of your experience from the secondary school :: that somehow you
67	or and <b>(er)</b> you must be <starts laughing> tired after our seminars <stops laughing> . :: because it's <overlap /> >>you know<<
68	but >>you know<< sometimes people are . <b>(er)</b> they are great academics
69	but they just are not good . <b>(er)</b> pedagogians <overlap /> psychologist all this
70	. well <b>(er)</b> so should I describe the pictures

CZ039		AS-Unit
71	(er) and then the lady . (er) looks at the picture or the panting painting :: and she doesn't like it .	
72	and (er) then . the lady . takes the painting :: and . shows it to her friends .	
73	(er) well I . I don't know	
74	(er) . n= no I think :: I wouldn't say anything	
75	(er) I think :: the painter must be used to that . :: because in past (er) often the painters (er) changed the (er) .. the looks of people for example :: when they painted (er) . kings and queens	
76	. (er) differently in real life :: so I think	

EN001		AS-Unit
1	I thought :: I'd talk about a f= a film :: which I have seen recently :: which is (erm) . Dead Poets' Society	
2	and it's about (erm) . life in a[ei] . (eh) public school in America	
3	(erm) and .. it . it shows :: how (erm) . a a group of boys :: inspired by a sort of . unauth= . unauthentic sort of English t not unauthentic but (erm) . <overlap /> unorthodox English teacher . :: (erm) form a group called Dead Poets Society :: where they (erm) .. read poetry and and treat it as a . a living thing . (0,5)	
4	and (erm) .. one of the boys is particularly inspired by this (erm) .	
5	(erm) his parents are very much against this :: because they've had lots of . plans . for his life :: how he's going to become a doctor	
6	(erm) because he's	
7	and managed to drag himself up and (erm) .. (1,0)	
8	and so they're very keen that (erm) . they should do .. or that<?> he should do what they want him to	
9	(erm) but he decides he wants to become an actor . (0,6)	
10	and (erm) he takes part in a play against his (erm) . father's wishes (0,5)	
11	and (er) . his father's furious with him :: and eventually the boy commits suicide :: <overlap /> and it's very sad at <begin_laughter> the[i:] end	
12	<X> <end_laughter> but (erm) I thought :: it was very interesting :: because it (erm) .. highlighted the sort of gap :: that can exist between . generations	
13	and and how (erm) .. sometimes parents want :: to fulfil all their wishes	
14	and (erm) .. I just thought :: it was very interesting :: because of that	
15	it was (erm) .. it was<?> quite a it was<?> quite a sad and touching <begin_laughter> <XX> <end_laughter>	
16	(erm) . I don't think I think I get on . well with my parents	
17	and (erm) . and they're quite accepting of >>>you know<<< :: whether I wanted :: to (er) go to university or not go to university	
18	(erm) so it's not a problem :: I've ever had but I have seen it in . you know some of my friends	

EN001	AS-Unit
19	but I have seen it in . >>you know<< some of my friends :: where <b>(erm)</b> .. >>you know<< they've been . they've felt :: forced into going to university by their parents
20	and then maybe other people <b>(erm)</b> .. have wanted :: to go to university :: but because their parents thought :: it was . <begin_laughter> useless and pointless <end_laughter> :: they haven't felt supported
21	<b>(erm)</b> ... one was this <b>(er)</b> .. quite a practical reason really
22	this <b>(erm)</b> . the job market was <begin_laughter> so very <end_laughter> bad :: when I left school
23	<b>(erm)</b> . <X> I I didn't think :: I would have much chance :: of . <b>(er)</b> getting a job without any further qualifications .
24	and also I thought . :: I would like . I would like :: to do more learning :: than I had . done at school <b>(erm)</b> .
25	<b>(erm)</b> . I I'm doing a . linguistics minor . <b>(erm)</b> as part of .. <b>(er)</b>
26	.. <b>(erm)</b> .. well
27	<b>(er)</b> just the one in fact <b>(er)</b>
28	I I've always been . <b>(erm)</b> . very keen on reading
29	<b>(erm)</b> .. and I found it hard to decide :: which . course to carry on with :: cos I didn't really want :: to . drop French <begin_laughter>:: but <b>(er)</b> <end_laughter>
30	because <b>(erm)</b> I would have . if I'd done it . along with English :: cos I didn't want to drop English . :: <b>(erm)</b> . I probably would have ended:: not not doing the[i:] year abroad for the French :: and so maybe I'd have got behind the[i:] other people
31	<b>(erm)</b> . I don't think :: I want to carry on . :: studying <b>(erm)</b> . an academic subject as such
32	but <b>(erm)</b> .. I'm quite interested in :: doing something like occupational therapy . :: and that would involve :: doing a . another course at . <b>(er)</b> university
33	. it it's to do with the[i:] idea :: that <b>(erm)</b> people's <b>(erm)</b> mental and physical health can be improved <b>(erm)</b> . with suitable activity
34	like people who have suffered strokes and that sort of thing :: or <b>(er)</b> who have <b>(erm)</b> been ill :: or have some sort of disability :: and <b>(erm)</b> . helping them :: to learn to do things :: that they used to do :: before they were ill
35	maybe writing <b>(erm)</b> <overlap /> or or whatever
36	<b>(erm)</b> .. if you do it as a first degree it's . three years
37	<b>(er)</b> n= no :: I <b>(er)</b> . I live in town
38	. <b>(er)</b> well being a second year there are= there aren't any places on campus <overlap /> for <XX>
39	<b>(erm)</b>
40	it is it is quite nice living off campus :: because <b>(erm)</b> .. you don't feel <b>(erm)</b> . stuck on campus all the time
41	<b>(erm)</b> . yeah . <b>(erm)</b> and <b>(er)</b> ... generally rent works out <?> more expensive
42	but <b>(erm)</b> .. no :: it is nice :: <b>(erm)</b> .. living . living in a house
43	<b>(erm)</b> I live with <b>(erm)</b> .. two other students and . the[i:] fiancée of one of those <overlap /> students
44	<b>(er)</b> yes
45	. yeah <b>(erm)</b> one of . one of them was already living in the house with some other friends:: who are now living elsewhere

EN001	AS-Unit
46	. yeah yes I do like it (erm) ...
47	(er) yeah I'm I'm in the Choral Society
48	(erm) ... well we have a we have a concert about . every . every term .
49	we did (erm) .. some Haendel and also a piece by . Vaughan Williams :: and it's it's mostly that sort of more serious sort of music
50	(erm) but<?> I really enjoy that (eh)
51	. it . it was (erm) .. Sunday first<?>
52	I used to have (erm) . singing lessons :: when I was at school
53	<overlap /> but (er) I can't really afford to <begin_laughter> now <end_laughter>
54	but (er) . but I'd like to again some time
55	. (erm) .. I suppose sometimes there are . problems fitting everything in . timewise
56	but (erm) . but having some hobbies . :: as I do :: it it sometimes makes it easier :: to work if I >>you know<< spend an evening doing something else
57	<laughs> (erm) no I have I have been home once this term
58	but (er) . I don't normally go home very much (erm) .
59	well I go home for the hol= holidays (erm)
60	(erm) . yeah I think :: quite quite a lot of people . (erm) stay here for most of the term
61	(erm)
62	yeah it it is quite odd :: when you (erm) . spend the whole term . here :: and then . go home :: and you're there for a month
63	but now :: that (erm) I know more people :: and they come to visit me at home
64	.. (erm) ... yeah it it is it is quite . difficult to
65	(er) .. yeah I I like going home
66	but (erm) yes I d= I do . feel yeah I do have to . tell them . :: where I'm going and <overlap /> and that sort of thing
67	but (erm)
68	(er) <overlap /> she's . she's older than me
69	<overlap /> she's .. (erm) she's doing an M A at the moment in Manchester
70	(erm) .. and she's doing
71	but (erm) . (erm) . I suppose it's been this situation for quite a while :: because .. last year :: I was here :: while she was living well she was living at home and working :: and then before that she was doing her first degree :: while I was still at school
72	but (erm) . yeah it's nice :: to see her :: when I do see her
73	... (erm) ... the man is drawing a . a picture of maybe it's his wife or his <overlap /> his girlfriend :: and (erm) and she's she's posing for him :: to do the picture :: and then she sees :: what he's drawn ::: and .. she doesn't like it :: and she's she's cross with him :: and .. it it does look quite like her

EN001	AS-Unit
74	(erm) .. so then it looks like :: he's drawn the picture again
75	but (erm) it draws it :: making her look more beautiful :: and she's<?> he's made a . smile :: where<?> she's really sort of frowning
76	(erm) .. and then she s= she seems pleased with that picture . :: and she's showing it off to all her friends

EN020	AS-Unit
1	(erm) number three
2	(erm) no it's just I thought :: it was really good (erm) ..
3	yes it is (erm) .. (P=1,0)
4	but .. it was actually . I mean :: I'm quite interested in films artistic :: and (erm) there was a lot in that that sort of it was quite a dark film
5	(erm) . and it was really good
6	but you don't actually see them (erm) the murders taking place
7	(erm)
8	.. (erm) .. I think :: it was the[i:] atmosphere actually .
9	(erm) .. but .. yeah I don't know really
10	(erm) part of the reason yes <overlap /> <laughs>
11	before I came up to university :: I didn't really . go that much :: because :: (er) where . I live :: it's it's fairly quiet so
12	to the to the cinema but (er) . well to to a town (erm) ..
13	it's . I enjoy it :: (erm) cos I'm very interested in films anyway so
14	no idea (erm) .. possibly
15	(erm) .. I'm doing culture and communication :: which sort of involves .. films to a certain extent but (er) .. not a great deal
16	(erm) .. I don't really know <laughs>
17	I'm . I've done (er) I did some summer work a couple of years ago :: and I was doing layouts for magazines
18	and (erm) .. that really really interested me
19	(erm) so .. maybe that sort of area . but involving films so ..
20	(erm) .. maybe involved in that
21	it's just lots of sort of . (erm) .. areas :: that I could look into
22	(erm) .. at the time :: when I lived on campus :: I thought :: I don't want to live out . :: because it's so nice being on campus

EN020	AS-Unit
23	<b>(erm)</b> you can't ever get away from . people you know :: which I'm so used to
24	<b>(erm)</b> and I've actually really enjoyed it . :: because you've got the freedom of your own house :: and then you can so you've got some people in your house
25	<overlap /> yeah . you've got your own space . <b>(erm)</b> .. (P=0,6)
26	I think it's more of more independent yeah the first year's really good staying on campus :: because you're .. <b>(erm)</b> you're in= involved in everything
27	yes :: and it's it's a bit more relaxing . <b>(erm)</b> ..
28	yes .. <b>(erm)</b> yeah :: so I think :: I I prefer .. I prefer town now
29	<b>(erm)</b> . I have a few .
30	<b>(erm)</b> I'm hitching in most of the time
31	no no it's not too bad <b>(erm)</b> ..
32	we've had <b>(er)</b> one of my friends was picked up by the vice-chancellor once ..(P=1,1)
33	and he didn't actually realise :: until <b>(erm)</b> . the vice-chancellor said something :: and he was like<?> oh no <overlap /> <laughs>
34	<b>(erm)</b> I don't know
35	it was really strange :: cos I put I mean I s= filled out all the[i:] <b>(er)</b> UCAS forms :: to <overlap /> decide where to go
36	I was just putting down universities :: that I thought .. would be would be good :: and most of them were miles away you know . <b>(erm)</b> ..
37	<b>(erm)</b> every time we go up and have a <XX> round (P=0,6)
38	and <b>(erm)</b> .. <overlap /> I just put it down
39	<b>(erm)</b> it's near a place :: called Tunbridge Wells
40	<b>(erm)</b> <overlap /> I
41	<b>(erm)</b> .. because it's just too far to go
42	<b>(erm)</b> it depends on <starts laughing> :: how fast my dad drives <stops laughing>
43	<b>(erm)</b> .. I haven't actually been by train
44	<b>(erm)</b> .. well there's a man . a man :: painting a picture of a woman .. (P=1,1)
45	and <b>(erm)</b> .. oh I don't know :: what to say . <b>(erm)</b> ..
46	and doesn't look very pleased :: with what he's .. with what he's drawn <b>(erm)</b> .. :: so he starts again :: and completely changes <starts laughing> :: what she looks like <stops laughing> .. :: <b>(erm)</b> making her look much prettier and .. <b>(erm)</b> more attractive :: than she does normally
47	and then she: . <b>(erm)</b> shows the picture to all her friends :: and . saying oh look at me <laughs>

EN051	AS-Unit
1	(erm) . so do you want me to . just
2	I . :: when was it :: about two years ago I went to the States for (erm) for three months
3	(erm) and so: I found it really interesting :: getting to know the people . :: and having to live in a completely foreign country on my own .
4	<overlap /> (erm) having to work and having to get up having to yeah stay out really late :: cos we were working in this restaurant .. (erm) .. (P=0,9)
5	and (erm) .. <overlap /> I
6	well the problem was . was :: that . the the the contract :: that I'd signed to work with this man in this fish restaurant :: had (erm) ..
7	(er) and I moved out
8	(erm) . and although I really enjoyed it at the time :: I didn't I was I just wanted :: to come back to England :: and I came back early :: because I was just so . :: cos when you're so far away from home and you have something going<?> wrong it's just it's very depressing .
9	(erm) .. so I suppose :: that if I've learnt anything from it it's :: that . you should be more careful about where you .
10	I just felt .. <overlap /> (erm)
11	she (erm) she: . she (em) :: what did she do .
12	(erm) and so .. we just thought >>you know<< :: this is not out of order
13	(erm) . and so she: threw a bottle at one girl's head
14	and >>you know<< I should have called the police :: when that actually happened :: cos <XX> to have someone :: to back us up . (erm) ..
15	(er) . eighteen
16	and once I got (er) . once I thought . :: cos at the time :: when that actually happened :: I thought :: I've gotta keep my job :: I've got to earn this <overlap /> money
17	(erm) . but no :: it was good fun
18	<overlap /> he teaches (erm) . he teaches teachers :: how to teach English
19	(erm) he actually did his linguistics ling= (eh) degree here <overlap /> or M A here
20	yeah . (erm) .. and so . my parents met in the Middle East .
21	(erm) and then when :: I w= I was born in England :: but . then went to Sudan for four years (P=0,5)
22	<overlap /> (erm) it's really <X> (eh) it's not very helpful actually <laughs>
23	(erm) .. yeah no it was I can remember some things I can remember a lot more about Malaysia cos I was older :: when I was living there
24	(erm)
25	(erm) . but whereas Ethiopia :: where they live now :: . you can't walk in the streets
26	yeah though it's not very often . (erm) <overlap /> (erm)
27	and (erm) and the Ethiopians used to confuse us
28	(erm) . I mean at the moment :: it's not too bad at all

EN051	AS-Unit
29	(em) . I much prefer . I'm . I just love Africa I really enjoy . going there
30	(erm) . not :: I haven't <overlap /> done that much
31	<XX> and they think :: it was the[i:] (erm) the tribesmen
32	<overlap /> (erm) I can't re= . :: whatever that one's called :: that's where we went .
33	(erm) but that was good as well .
34	cos . my mom went there for a[ei] (erm) .. a school .
35	(erm) . she: she works for the[i:] American school in Ethiopia and they've got American schools all over Africa and they have these huge conferences where all teachers come together
36	so (erm) so she went she really really liked it
37	she said :: it's very . >>you know<< . (eh) it's it's very very touristy .
38	<overlap /> they (erm) the just they . tourists have never been there :: so there's these beautiful things in Ethiopia :: that no one's ever seen :: which is quite nice .
39	<overlap /> but (erm)
40	there's (em) .. these huge churches . carved out of rock
41	(erm) .. so >>you know<< it's it's difficult :: to see some of the things sometimes .
42	(erm) and it's just I miss :: seeing the sky :: and . >>you know<< it's so overcast and <overlap /> dull over here
43	I don't know about . (erm) . mammals ..
44	we saw: we went (er) at Christmas .
45	we were coming from .. (eh) the ho= the restaurant
46	there's this hotel resort on this (eh) on a lake
47	and then there's a[ei] (em) . an ostrich farm just down the road so
48	.. (erm) I love Malaysia because . I grew up there .
49	but Chinese tend to be very sort of (erm) .. like
50	(erm) . and so whereas in Ethiopia you might you >>you know<< you can actually <overlap /> get a joke out of Ethiopians
51	(erm) and going to school with Chinese . (erm) ..
52	you can't go and play cos they don't play they just work (erm)
53	(erm) . I used to just go and play with them .
54	(erm) . I really liked it it was
55	... (erm) .. (em) (em) (em) (em) ... (er) I was going to say . :: it's .. it's Suzanne's .. :: Suzanne is the woman:: and it's (erm) . it's Suzanne's hu= husband's birthday . very soon (0,5)

## Repeats

CZ006	AS-Unit
1	describing . <b>{the} . the</b> visit :: and why I liked it
2	and (erm) . <b>{we} we</b> went by bus :: which is . a little annoying :: because it was a <b>long long</b> way ..
3	(erm) <b>{we} we</b> went to: France :: took the ferry on the other side . :: then went straight to Wales :: and then from there another ferry to Ireland .
4	so just when I saw the city :: <b>{I} I</b> just thought :: well it's not as amazing :: as I as I imagined it so .
5	and there was (eh) live music :: playing there :: and people and dancing and drinking and singing :: and <b>{it} it</b> was . >> I think . << nice experience
6	<b>{I} I</b> wouldn't dare doing that again <laugh>
7	but then it got . the weather got quite terrible :: and started raining :: so it was .. quite exciting going to the top :: because the way . was quite rusty . (P=0,45) :: and (eh) . <b>{not very} . not very</b> pleasant really
8	but . <b>{we} we</b> made it to the top :: and it was really nice
9	(eh) . <b>{we} we</b> had some nice time in Galway :: and then we visited Limerick :: which is . <b>{not} not</b> very far from Galway . (P=0,83)
10	<b>{the} the</b> stay in (eh) Sixmilebridge was quite nice too .
11	the: . lady :: who was organizing the trip said :: that the people there . still remember us :: and the (mm) when we visited :: because <b>{we} (eh) we</b> went to some local pub :: and there was a band playing also :: and we joined :: and sang with them :: <b>and</b> .
12	<b>(eh) and</b> . yeah they said :: that . it was the most lively evening <starts laughing> :: probably they had there <b>for for</b> years <stops laughing> <overlap />
13	yeah and th= <b>{they} they</b> actually asked the lady . :: whether we are coming again someday <laugh>
14	and yeah the next day <b>{we} we</b> took the ferry back to England :: and . we went to London :: and spent <XX> . some nine hours break . <b>{for} for</b> the bus drivers ..
15	<b>{I} I</b> went to the National Gallery . <overlap /> also
16	well I think :: it's the people :: and as I said :: Britain for me <b>{is} is</b> a h= is a matter of heart really
17	<b>{I} . I</b> just love staying there :: and watching the people .
18	<b>{I} . I</b> started :: listening to the Beatles
19	yeah and then . watching movies :: <b>{I} I</b> really like watching movies in English
20	<b>{I} I</b> really hate the dubbed movies :: they do here . :: so I prefer watching that in English :: and . well you catch phrases from that and .. well
21	and <b>{he} he</b> was really surprised :: that we have everything dubbed here . (P=0,67)
22	and I have to say :: <b>they they</b> repeat things a lot
23	but not <b>{I} I</b> wouldn't say in a good way .
24	and (eh) when I checked . the . grammar topics :: they had in a textbook :: I say hey :: <b>{we} we</b> did that twice already :: so why is it here again . (P=0,58)
25	<b>{I} I</b> think :: it was . things like . can and can't . :: so <overlap /> really <b>{basic} basic</b> stuff
26	but yeah back to <b>{my} . my</b> experience then at (eh) . <lip sound> (er) secondary school ..
27	it was (eh) . well <b>{we} we</b> were quite lucky :: because we had a good teacher .. (P=1,2)

CZ006		AS-Unit
28	and you go <b>{by}(eh) by</b> the underground :: take (eh) . <b>{one} one</b> stop to <foreign> Můstek </foreign> . :: and change the trains :: and it's at a stop . <foreign> Hůrka </foreign>	
29	. (eh) <b>{I} . I</b> couldn't decide :: which school to choose .	
30	<b>{I} I</b> actually tried it . for about a month .	
31	so yeah <b>{I} I</b> dropped out from the other school . after about a month or so <laugh>	
32	so <b>{I} I</b> really enjoy teaching it right now	
33	.. so (eh) . there's a lady . possibly . either the lady wanting :: to have a picture of herself :: <b>{or} or</b> the guy just trying :: to . perhaps hit on the lady :: by drawing a picture of her .. (P=1,1)	
34	. but she's angry probably yelling at the painter :: that . it's not me :: it doesn't look like me . :: so . <b>{she} she</b> wants him :: to . (er) . redo the picture ..	
35	and although . the . picture does not look very much like the lady . :: <b>{the;} the</b> lady seems :: to be pleased with the picture . :: when she shows it to other people . possibly in her house	
36	or I don't know :: if it's in a gallery maybe . :: so <b>{she} she</b> seems :: to be very proud of :: how she looks in the picture	
37	. I think :: it depends very much on :: how good of a friend I would be with her :: because I tend :: to be quite honest . so <b>{I} I</b> might . >>I don't know<< . mention things <laugh>	
38	<b>{is that} . is that</b> really you	

CZ034		AS-Unit
1	(er) so there was (eh) .. there was nothing <b>{as} (er) {as} (eh) as</b> an excitement to tell the truth at the beginning	
2	but when I look at it now: :: I am very happy :: that (erm) . <b>{I} I</b> learned it and (eh) .. (erm) .	
3	<b>{in} in</b> America for me (erm) . with English	
4	and <b>{that was} that was</b> the first: thing :: which amazed me really	
5	and (eh) . just (eh) it's those little things (eh) such as food . :: <b>{or;} &lt;laughing&gt; or:</b> . simply .. (er) it's very hard <overlap /> :: to describe it	
6	and of course (eh) part of this experience is . (eh) was my family :: and I was very lucky in having a great family (er) :: which took a great care of me: :: and (eh) took me: for example to Berlin as well <b>{for} for</b> two weeks :: and: (er) I had one week off: and: .	
7	what I admired <b>{about} about</b> <foreign> München </foreign> and (eh) the[i:] area down south	
8	<X> and even <b>{the} the[i:]</b> adults everybody . :: and actually <b>{it is} it is</b> weird :: when (eh) someone is not dressed up like that so:	
9	<b>{and} and</b> I think :: that's simply very important	
10	but (eh) . <b>{that was} that was</b> (eh) (eh) actually also very nice	
11	we were just the pupils at (eh) at a high school :: and .. (eh) again <b>{they were} they were</b> so similar (eh) to us	
12	we were just the pupils <b>{at} (eh) at</b> a high school and .. (eh) again they were they were so similar (eh) to us	

13	(eh) we went with a<?> family we went on a couple of trips (erm) :: and actually <b>{we}</b> (eh) . accidentally <b>we</b> went on Sunday :: and what I witnessed :: there was . (eh) :: that at the . one of the . mountains (eh) . <b>{on}</b> on a top of one (er) :: there was a group of musicians :: who were playing musical instrument :: and it was . a part of a Sunday . mass
14	(eh) they told me :: that (er) <b>{it}</b> it didn't use to be like that (em) . twenty years ago
15	there it was <b>{another}</b> <b>another</b> .. (erm) inspiration another . bit of inspiration . :: (erm) and: (eh) after I came back . (erm) after a year or two :: I was very lucky in . getting (eh) admitted (eh) for a job in a German newspaper :: because (er)
16	but it was mainly <b>{my}</b> <b>my</b> boss :: so it was .
17	(erm) . so (eh) that was that :: and (erm) <b>{why}</b> <b>{why}</b> <b>why</b> is there even a German newspaper in our country (erm) ..
18	and also (er) . (eh) some . :: they are called <foreign> Sudetendeutschen </foreign> :: <b>{some}</b> (er) . (er) <b>some</b> (er) . (erm) order it even from Germany mo= again mostly from <foreign> Bayern </foreign> (erm) . :: so: it was very interesting and:
19	<b>{in}</b> <b>in</b> German you mean or
20	<b>{this}</b> <b>{this}</b> <b>this</b> (er) contact with Germany it . started . three years ago
21	and <b>{as}</b> (eh) <b>as</b> I told you before . :: I was not very happy :: that <overlap /> I had to study it at school
22	so it . and that's another point :: why: (eh) I wanted to talk about this topic :: because <b>{I}</b> <b>(mm)</b> <b>I</b> never planned it .
23	<b>{I}</b> <b>I</b> really think so :: and (erm) . (er) one of my friends :: who went on an on the[i:] Erasmus programme
24	but those simil= similarities <b>{are}</b> <b>are</b> really huge . :: as you said
25	and: (er) <b>{I}</b> .. <b>(er)</b> <b>I</b> dunno I well
26	<laughs> because I am interested <b>{in}</b> <b>in</b> languages
27	<b>{in}</b> <b>in</b> cultures
28	(erm) .. <b>{I}</b> <b>I</b> really appreciate the fact :: that I (eh) chose now only the[i:] linguistic branch
29	(er) probably it would be something <b>{in the}</b> <b>in the</b> direction <b>{of}</b> <b>of</b> language teaching
30	(erm) . in . very . in general <b>{I like}</b> <b>(erm)</b> . <b>(er)</b> <b>I like</b> :: to work (er) with languages :: that are in contact :: so (er) .. I would not be opposed :: to trying translation for example
31	but I re= <stops laughing> I really think :: that (er) <b>{it is}</b> <b>it is</b> hard job <laughs> ..
32	(mhm) so (er) there is a man in the picture :: and this man is . (erm) painting .. <b>{a}</b> . <b>a</b> picture of a woman
33	(er) and sh= she is pointing <laughing> <b>{at her}</b> <b>at her</b> neck ..
34	yeah the hair is different (erm) .. :: because in the second picture her (eh) hair is straight :: . while in the . third one <b>{it}</b> <b>it</b> has curls (er) . at the end . :: and: so probably <b>{what}</b> <b>what</b> the model said to the painter was :: that she . didn't like . :: what her hair looked like :: so: therefore he changed it
35	that's questionable <overlap /> <b>{and}</b> <b>and</b> :: so I dunno

1	but then I realized :: it's much <starts laughing> easier :: to talk about <stops laughing> something :: we don't like really :: so <b>{I}</b> . (er) <b>I</b> . <overlap /> remembered
2	but . it turned out :: (ehm) that it <starts laughing> wasn't such a good idea <stops laughing> :: because . (er) it seems to me :: that (er) . nowadays :: when somebody wants to . (er) stage Shakespeare . :: (er) . <b>{the}</b> . <b>{the}</b> . <b>the[i:]</b> actors and <b>{the}</b> <b>the</b> director on the people all the people :: who . create <b>{the}</b> <b>the</b> performance . think :: they must change it somehow :: to make it att= more attractive for the <overlap /> for contemporary people and contemporary audience
3	and (eh) I always thought :: that the tragedies were <b>{much}</b> (er) <b>much</b> more difficult to stage
4	they are there so <overlap /> <b>{the}</b> <b>the</b>
5	so I went to the[i:] internet :: <b>{I booked}</b> (er) <b>I booked</b> Midsummer Night's Dream . :: and I was looking forward to <overlap /> seeing it
6	<b>{no:}</b> (er) <b>no</b> :: I've seen
7	I just (er) I was si= sitting there . :: and . (er) I think :: <b>{they}</b> . <b>they</b> did (er) great thing with the with the space on the stage :: because there was d= dark . :: and there was . (er) . suddenly . just (er) . the[i:] actors . playing
8	so this was <b>{the}</b> <b>the</b> catharsis >>I think<< . should be there <overlap /> in tragedies
9	and <b>{it was}</b> <b>it was</b> . well . maintained :: and it . I think :: it's quite difficult <b>{to}</b> . (er) <b>to</b> achieve this nowadays :: because people are just . so bored :: because they see . so much action <b>{in}</b> <b>in</b> movies <overlap /> :: so it's hard to
10	. (er) create <b>{some}</b> . <b>some</b> (eh) powerful . (eh) feelings <b>{in}</b> <b>{in}</b> <b>in</b> the theatre :: when you've . got only stage and <overlap /> some
11	yes (er) I think :: the contrastive view is <b>{very}</b> (er) <b>very</b> interesting :: and I (er) . studied . both subjects <overlap /> as on my B A studies as well
12	but also <b>{the}</b> <b>the</b> dep= both or all of the departments have different approaches :: so <overlap /> it's interesting
13	but it was (er) maintained very differently :: and I must say :: that the[i:] English department . (er) relied mo= more . on (er) <name of department> :: and I think :: it was better :: because we had this seminar . :: but (er) on the Czech language and literature we had only <b>{the}</b> <b>the</b> lecture <overlap /> :: which was
14	<b>{but}</b> <b>but</b> as many of my colleagues said <stops laughing> :: (er) <b>{the}</b> <b>{the}</b> . (er) <b>the</b> BA studies . somehow . (er) destroyed <starts laughing> our love for literature <stops laughing> a bit <overlap /> a bit :: because
15	<overlap /> so I had real struggle <b>{to}</b> <b>to</b> decide
16	yes <overlap /> <b>{I}</b> <b>I</b> say :: that we are deformed <laughs>
17	but . >>you know<< it's (er) . it is an advantage as well :: because <b>{you can}</b> <b>you can</b> see thr= through it :: and ordinary people can't :: so . it's .
18	<overlap /> but sometimes <b>{you}</b> <b>you</b> just people read something :: and they just . don't .
19	but (er) . of course <b>{the}</b> <b>the</b> studying literature opens your mind to <overlap /> several interpretations
20	well (er) currently <b>{I}</b> <b>I</b> work as a teacher of (eh) Czech language for foreigners :: and (er) English as well .
21	but . when I came <b>{to}</b> <b>to</b> the university :: I was a bit disappointed <b>{about the}</b> . (er) <b>about the</b> things like lectures and seminars :: because seminars were usually great
22	but I just . can't get over the lecture stuff :: because it's . >>you know<< even . when the . topic is extremely interesting . :: (er) listen . to: . person's . talking for one hour and half without (eh) any activity <b>{from your}</b> <b>from your</b> own is extremely difficult <overlap /> :: and I
23	to stand there :: and just talk for half an hour <overlap /> :: and <b>{I}</b> <b>I</b>
24	they are <b>{they}</b> <overlap /> <b>they</b> know lot of stuff

CZ039		AS-Unit
25	(er) well <b>{I}</b> . <b>I</b> don't know	

EN001		AS-Unit
1	it's a film :: which <b>{I}</b> . <b>I</b> had never seen before :: but everyone else seemed to have seen <laughs> :: and so I was interested :: to see it . (0,5)	
2	(erm) and .. <b>{it}</b> . <b>it</b> shows :: how (erm) . <b>{a}</b> a group of boys :: inspired by a sort of . unauth= . unauthentic sort of English t not unauthentic but (erm) . <overlap /> unorthodox English teacher . :: (erm) form a group called Dead Poets Society :: where they (erm) .. read poetry <b>{and}</b> <b>and</b> treat it as <b>{a}</b> . <b>a</b> living thing . (0,5)	
3	yes and he and well <b>{he's}</b> <b>he's</b> having chances that his parents never had	
4	his father had . worked from being <b>{very}</b> . <b>very</b> poor and humble and . (0,6)	
5	<b>{and}</b> <b>and</b> how (erm) .. sometimes parents want :: to fulfil all their wishes	
6	<b>{it was}</b> (erm) .. <b>it was</b> <?> quite a it was<?> quite a sad and touching <begin_laughter> <XX> <end_laughter>	
7	er I've never really had <b>{that}</b> <b>that</b> problem .. myself	
8	<b>{and}</b> (erm) . <b>and</b> they're quite accepting of >>you know<< :: whether I wanted :: to (er) go to university or not go to university	
9	(erm) . <X> <b>{I}</b> <b>I</b> didn't think :: I would have much chance :: of . (er) getting a job without any further qualifications .	
10	<overlap /> and .. <b>{I}</b> <b>I</b> dropped it rather than rather than do it as a minor	
11	(er) n= no <b>{I}</b> (er) . <b>I</b> live in town	
12	<b>{it is}</b> <b>it is</b> quite nice living off campus :: because (erm) .. you don't feel (erm) . stuck on campus all the time	
13	and <b>{you}</b> <b>you</b> go home to a different place at the[i:] end of the day .	
14	but (erm) .. no :: it is nice :: (erm) .. <b>{living}</b> . <b>living</b> in a house	
15	.. <b>{who}</b> <b>who</b> do I live with	
16	(er) yeah :: <b>{I'm}</b> <b>I'm</b> in the Choral Society	
17	(erm) ... well <b>{we have a}</b> <b>we have a</b> concert about . <b>{every}</b> . <b>every</b> term .	
18	it .. it's nice :: to go into the town from campus sometimes :: because the campus is so modern . :: and <b>{in}</b> . <b>in</b> the town <overlap /> there's a lot of older buildings :: <overlap /> which is nice<?> <laughs>	
19	we did (erm) .. some Haendel and also a piece by . Vaughan Williams :: and <b>{it's}</b> <b>it's</b> mostly that sort of more serious sort of music	
20	<b>{but}</b> (er) . <b>but</b> I'd like to again some time	
21	<b>{but}</b> (erm) . <b>but</b> having some hobbies . :: as I do :: <b>{it}</b> <b>it</b> sometimes makes it easier :: to work if I >>you know<< spend an evening doing something else	
22	.. yeah well <b>{the}</b> <b>the</b> holidays <b>{are}</b> <b>are</b> quite .. long	
23	<b>{my}</b> <b>my</b> parents tend to come up once or twice a term :: to: see me :: cos it's not all that far to come	
24	.. (erm) ... yeah it <b>{it is}</b> <b>it is</b> quite . difficult to	
25	<b>{I do}</b> <b>I do</b> like going home	

EN001	AS-Unit
26	<b>{it has} it has</b> advan= some advantages over being here and being here
27	(er) <overlap /> <b>{she's} . she's</b> older than me
28	(er) .. yeah <b>{I} I</b> like going home
29	I do get on with my parents and <b>{they're not} they're not</b> very . strict
30	<overlap /> <b>{she's} .. (erm) she's</b> doing an M A at the moment in Manchester
31	<laughs> yes <b>{I} I</b> do miss her
32	... (erm) ... the man is drawing <b>{a} . a</b> picture of maybe it's his wife or <b>{his} &lt;overlap /&gt; his</b> girlfriend :: <b>{and} (erm) and {she's} she's</b> posing for him :: to do the picture :: and then she sees :: what he's drawn :: and .. she doesn't like it :: and <b>{she's} she's</b> cross with him :: and .. <b>{it} it</b> does look quite like her
33	<overlap /> but she thinks :: <b>{it's} .. it's</b> too ugly
34	it's a film :: which <b>{I} . I</b> had never seen before :: but everyone else seemed to have seen <laughs> :: and so I was interested :: to see it . (0,5)

EN020	AS-Unit
1	<b>{it's} .. it's</b> . it one of the sort of film :: I wouldn't normally go for :: because <b>{it's} it's</b> scary <laughs>
2	but .. it was actually . I mean :: I'm quite interested in films artistic :: and (erm) there was a lot in <b>{that} that</b> sort of it was quite a dark film
3	<XX> lots of powerful images <b>{and} and</b> that ..
4	so .. <b>{it's} it's</b> a different approach really
5	no <b>{there's} there's</b> a bit at the[i:] end so .
6	>>you know<< <b>{what's} what's</b> going on
7	before I came up to university :: I didn't really . go that much :: because :: (er) where . I live :: <b>{it's} it's</b> fairly quiet so
8	<b>{to the} to the</b> cinema but (er) . well <b>{to} to</b> a town (erm) ..
9	but since I've been up here >> you know<< :: <b>{it's} it's</b> so easy :: just to walk down to the cinema :: so I've <b>{seen} .. seen</b> loads <laughs>
10	<overlap /> <b>{yeah} . yeah</b> it probably does
11	<overlap /> <b>{yeah} yeah</b> <laughs>
12	<b>{yes} yes</b> :: I was
13	yes :: and <b>{it's} it's</b> a bit more relaxing . (erm) ..
14	yes .. (erm) yeah :: so I think :: <b>{I} {I prefer}</b> .. I prefer town now
15	<overlap /> whereas >>you know<< <b>{you} you</b> do sort of ..
16	yeah :: <b>{it's} it's</b> ..
17	<b>{no} no</b> :: it's not too bad (erm) ..

EN020	AS-Unit
18	<overlap /> {no} no so far :: touch wood :: I haven't been <overlap /> <laughs>
19	so no :: {it's} {it's} it's okay
20	but normally {there's} there's a big queue :: so well or a couple of people whatever so
21	yeah yes :: {there's a} there's a lamppost :: <laughs> so everyone {stands} . stands behind it <overlap /> <laughs>
22	<overlap /> it yeah it is yes :: {which is} which is good <laughs>
23	{so} so he didn't sort of say very much <laughs>
24	yeah :: {it's} . it's interesting
25	I was just putting down universities :: that I thought .. {would be} would be good :: and most of them were miles away you know . (erm) ..
26	I can call on them {which is} . which is quite good <overlap /> yeah
27	<overlap /> it is very different yes . :: {which} which makes it quite nice >>you know<< s= coming to . somewhere completely different
28	(erm) .. well there's {a man} . a man :: painting a picture of a woman .. (P=1,1)

EN051	AS-Unit
1	I . :: when was it :: about two years ago I went to the States {for} (erm) for three months
2	and . {it was} . it was very interesting :: because although I've spent a lot of time at boarding school so away from my parents . :: I'd never been to America before .
3	well the problem {was} . was :: that . {the} {the} the contract :: that I'd signed to work with this man in this fish restaurant :: had (erm) ..
4	well the problem {was} . was :: that . {the} {the} the contract :: that I'd signed to work with this man in this fish restaurant :: had (erm) ..
5	but then everything got really difficult between . me and him :: {because} because I wasn't living and I wasn't giving him money for my rent any more ..
6	{she} (erm) {she:} . she (em) :: what did she do .
7	but they thought :: that we were just over the top :: that {it} . it wasn't really that big a deal .
8	and <first name of a girl> this . weird woman was sleeping with the boss :: so of course he wouldn't do anything about it :: because he was just like oh {no} no :: <first name of a girl> wouldn't do that :: <first name of a girl> wouldn't do that
9	and I just thought :: everyone just thinks :: I'm a liar basically :: and {I} I wasn't <laughs>
10	but I thought :: >> you know<< {I'm} I'm here :: to have a nice time <overlap />:: and make a bit of money
11	<overlap /> {he teaches} (erm) . he teaches teachers :: how to teach English
12	but you get >>you know<< . if you're <overlap /> {a} a white
13	well the problem used to be . :: that they have funny number plates :: {that} that mark :: what country you're from

EN051	AS-Unit
14	and some of them <b>{are} are</b> diplomatic :: and some of them are aid .
15	and so during the cold war :: when all the Russians were supplying the military arms to Ethiopia :: <b>{they} they</b> hated the Russians
16	<b>{it's} it's</b> nice
17	<b>{and} and</b> I . I'm really happy living like that
18	and there'll be forty different types of chocolate and you just think <b>{why} why</b> do we need
19	<b>{I find} .. I find</b> coming back to England where people are so aggressive and so you know insulated and <overlap /> so cut off from everyone else
20	for foreign . >>I think<< <b>{or} or</b> may= maybe for everybody well maybe for certain Kenyans as well
21	she said :: it's very . >>you know<< . (eh) <b>{it's} it's</b> very very touristy .
22	(erm) .. so >>you know<< <b>it's it's</b> difficult :: to see some of the things sometimes .
23	sometimes they escape and you just see herds of ostrich sort of stampeding across the . s= <b>{the} the</b> savannah <overlap /> so
24	and . yeah <b>{it was} it was</b> untouched then
25	... (erm) .. (em) (em) (em) (em) ... (er) I was going to say . :: <b>{it's} .. it's</b> Suzanne's . :: Suzanne is the woman:: and <b>{it's}(erm) . it's</b> Suzanne's hu= husband's birthday . very soon (0,5)
26	so she gets up :: and she looks :: and she's absolutely flabbergasted :: and she says :: .. <b>{but} {but} but</b> I don't look like that .